

# 多尺度特征融合的铁轨异物入侵检测研究

王楠, 侯涛, 牛宏侠

(兰州交通大学自动化与电气工程学院, 730070, 兰州)

**摘要:** 针对铁路轨道异物检测中不同尺度异物目标的检测易受复杂环境的影响, 导致出现检测精度低及检测速度慢等问题, 提出一种多尺度特征融合的铁轨异物检测(RMF-YOLO)算法。首先, 设计并引入改进的卷积注意力模块(ICBAM), 结合 YOLOv7 特征提取网络, 以增强复杂场景下的特征提取能力。其次, 在所有高效层聚合网络模块中采用 GhostConv 替代普通卷积层, 以降低计算复杂度, 提高特征输出效率; 设计一种改进的加权双向特征金字塔网络 N-BiFPN 结构, 加强多尺度特征融合能力, 平衡不同层级特征信息, 提高多尺度检测能力。最后, 为进一步提升检测精度, 采用 WIoU 损失函数结合动态非单调聚集机制, 有效应对低质量锚框产生的梯度, 提高模型对不同尺度异物检测的整体性能。实验结果表明: 在自制的铁轨异物数据集上, RMF-YOLO 算法减少了原网络模型的参数量, 有效提升了模型的检测精度与检测速度, 改善了漏检与误检问题, 平均精度提升了 5.5%, 检测速度提升了 5.88%, 计算量减少了 12.25%, 能满足铁轨入侵异物检测中对检测精度和实时性的需求。

**关键词:** 铁轨异物检测; 特征融合; 多尺度; YOLOv7; 损失函数

**中图分类号:** TP391.41; U298 **文献标志码:** A

**DOI:** 10.7652/xjtuxb202409014 **文章编号:** 0253-987X(2024)09-0139-15

## Research on Railway Track Foreign Object Intrusion Detection Based on Multi-Scale Feature Fusion

WANG Nan, HOU Tao, NIU Hongxia

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

**Abstract:** To address the issues of low detection accuracy and slow detection speed caused by the influence of complex environments on detecting foreign objects of different scales on railroad tracks, this paper proposes a multi-scale feature fusion RMF-YOLO (railway multi-scale fusion YOLO) method for railroad track foreign object detection. Firstly, ICBAM, an improved convolutional attention module, is designed and introduced in combination with the YOLOv7 feature extraction network to enhance feature extraction capability in complex scenarios. Subsequently, GhostConv is adopted in all efficient layer aggregation network modules instead of regular convolutional layers to reduce computational complexity and enhance feature output efficiency. An improved weighted bi-directional feature pyramid network (N-BiFPN) structure is introduced to enhance multi-scale feature fusion, balance feature information across different levels, and improve

收稿日期: 2023-12-12。 作者简介: 王楠(1997—), 女, 硕士生; 侯涛(通信作者), 男, 教授, 硕士生导师。 基金项目: 甘肃省重点研发计划-工业类资助项目(23YFGA0049); 甘肃省自然科学基金资助项目(21JR7RA321; 22JR5RA358)

网络出版时间: 2024-04-03

网络出版地址: <https://link.cnki.net/urlid/61.1069.T.20240402.1344.006>

multi-scale detection capabilities. Lastly, to further enhance detection accuracy, the WIoU loss function combined with the dynamic non-monotonic aggregation mechanism is used to effectively deal with the gradient generated by the low-quality anchor frames and to improve the overall performance of the model for the detection of foreign objects at different scales. The experimental results indicate that on the self-made railway track foreign object dataset, the RMF-YOLO algorithm reduces the parameter count of the original network model, effectively enhancing the model's detection accuracy and detection speed of the model and addressing issues related to missed and false detections. On average, the precision increased by 5.5%, the frames per second (FPS) improved by 5.88%, and the computational load decreased by 12.25%. These enhancements meet the requirements for both detection accuracy and real-time performance in railway track foreign object detection.

**Keywords:** railroad foreign object detection; feature fusion; multi-scale; YOLOv7; loss function

铁路在全球运输系统中扮演着重要角色,随着铁路运输的快速发展,铁路运输安全也越来越受到人们的关注。对于铁路运输安全,铁轨附近出现的入侵异物是威胁铁路系统安全的原因之一,典型的铁路非法入侵有行人、车辆、动物、轨旁基础设施坍塌等,特别是在开放式铁路环境下,行人的行为是不可预测的<sup>[1]</sup>。列车前方异物的入侵会对高速列车的安全构成重大威胁,因此及时发现非法入侵异物,确保列车安全至关重要。

有效的铁路异物入侵检测方法对铁路安全具有重要意义。目前,铁路入侵检测主要有两种方法:接触式和非接触式。接触式检测需要大量的工程工作,主要是利用传感器技术判断异物是否与传感器接触,这种物理保护方法受到外界因素影响较大,不适合大规模的使用。基于深度学习的异物目标检测是一种典型的非接触式检测方法,具有易于维护和准确性较高的特点,目标检测方法可分为两阶段和单阶段。两阶段方法涉及生成候选区域,然后进行分类和回归,虽然具有较高的检测精度,但是由于计算复杂而难以进行实时检测<sup>[2]</sup>。单阶段方法需要生成位置和类别概率<sup>[3-5]</sup>。近年来,涌现出基于 Transformer 的目标检测算法,将目标检测任务视为序列到序列的预测问题,并通过引入自注意力机制推理目标与全局图像的上下文关系,如端到端目标检测器(DETR)<sup>[6]</sup>等。

目前,已经有很多学者开展基于深度学习的铁路异物入侵检测研究。Guo 等<sup>[7]</sup>通过改进的 SSD 网络进行高速铁路间隙入侵检测,将高级特征反卷积为低级特征,并与原始的低级特征融合,以增强语

义信息。汪世财等<sup>[8]</sup>根据障碍物的点云位置特征和时间特征去除背景点云数据,以提高障碍物关联可靠性。徐岩等<sup>[9]</sup>针对铁路异物检测问题,提出一种替代全连接层的方法,然而这种方法在复杂环境下的异物检测性能相对较弱。张剑等<sup>[10]</sup>采用可切换的空心卷积提取特征,降低了目标检测的误报率,但每秒帧数(FPS)较低,不能满足实时检测的要求。Hussain 等<sup>[11]</sup>提出使用反向注意力机制进行目标对象的清晰边界检测,使用二元交叉熵(BCE)损失与结构相似性指数度量(SSIM)和交集(IoU)损失的组合来监督训练。胡行涛等<sup>[12]</sup>将不确定性添加到检测系统中,可以有效提高检测精度,但该方法对特征遮挡和快速移动的目标存在漏检情况。Cai 等<sup>[13]</sup>提出一种基于毫米波雷达的异物检测系统,在雷达数据的基础上对不同维度的数据进行处理,以获得目标信息。He 等<sup>[14]</sup>结合多种多尺度增强方法,提高了小目标物体的检测能力,但是实时性较弱。Wang 等<sup>[15]</sup>提出一种基于深度生成网络和自回归模型的无监督入侵检测方法,通过重建和原始输入图像的差值定位异常区域。管岭等<sup>[16]</sup>提出一种改进的 YOLOv4-tiny 网络模型用于轨道异物检测,融合跨阶段结构和引入多头注意力机制,从而提高网络性能和效率。

随着网络结构的深入和更多新架构的出现<sup>[17]</sup>,现有的铁轨异物检测算法存在难以平衡检测精度和速度的问题,不能满足铁路入侵检测的实时性和准确性的需求。不同光照环境、天气、遮挡等复杂环境也影响了入侵异物的可视化特征,对铁轨异物目标的检测造成干扰。铁轨附近可能出现的异物尺寸复

杂多样,异物目标也会因与摄像头的距离差异而导致成像尺寸差异较大。这些因素都对目标检测算法的实用性和鲁棒性提出了更高的要求。现有的目标检测算法并不能保证在复杂环境下检测多尺度异物,对于姿态各异的工人以及远处的物体,在提取特征不足的情况下,可能难以检测出来。

针对上述问题,本文以 YOLOv7 为基础网络,提出了一种基于多尺度特征融合的 RMF-YOLO (rail multi-scale fusion YOLO) 铁轨异物入侵检测算法,旨在实现更小的参数量和更高的检测精度。通过引入改进的卷积注意力模块(ICBAM)、多尺度特征融合网络以及改进损失函数等策略,对 YOLOv7 网络进行改进,以提高异物检测的准确性,特别是针对不同尺度异物的特征提取能力。本文主要贡献如下。

(1)在特征提取网络中设计一种混合注意力机制,即 ICBAM。该注意力机制通过提取全局上下文信息,可以有效提高局部表达能力,更好地拟合入侵异物特征信息,提高异物入侵图像中对不同尺度异物的检测能力。

(2)为了降低网络的计算复杂度,在所有高效层聚集网络(ELAN)模块中引入 GhostConv 模块,代替普通卷积,实现模型压缩,以平衡检测精度和速度。

(3)在特征融合网络中,提出一种改进的特征金字塔网络 N-BiFPN 作为特征融合模块,代替原有的路径聚合网络(PANet),从而实现高效的双向跨尺度连接和加权特征融合,以获得更多的特征信息,避免了异物特征的丢失,显著提高了算法的检测能力。

(4)在损失函数的计算中引入 WIoUv3 (wise-intersection over union) 损失函数,采用梯度增益分配策略,增强模型对低质量锚框的聚集,提高模型在铁轨异物检测中的整体性能。

## 1 RMF-YOLO 算法的网络模型

### 1.1 YOLOv7 算法

YOLOv7<sup>[18]</sup> 的结构框架包括输入端、特征提取网络、特征融合网络以及头部架构。特征提取网络主要由一系列卷积层组成,专门用于从输入数据中提取特征。首先,将固定分辨率为  $640 \times 640$  像素的图像分批输入到骨干网络,进入主干层后产生双下采样特征图。之后,由 Conv 层和 ELAN 层卷积组

成网络层,产生 4 倍下采样图像。然后,进行 3 次重复的最大池化(核尺寸  $n=1$ )下采样,结合 ELAN 层,得到 32 倍下的采样特征图。在特征融合网络中,空间金字塔池化跨阶段模块(SPPCSPC)<sup>[19]</sup> 首先通过将特征通道的数量减半来汇集来自上部输出的特征。然后,采用路径聚合特征金字塔网络(PANet)模块进行多尺度的特征融合,采用自上而下的结构传递深层的强语义特征,以增强整个金字塔。同时,采用自下而上的结构,将浅层的图像结构、颜色、边缘等特征位置信息传递上去,从而实现不同层次特征高效融合。最后,头部架构采用 REP 结构对 PANet 结构进行通道数调整,将特征送入尺寸为  $1 \times 1$  的卷积模块,用来预测图像的置信度、类别、锚框信息,并最终生成检测结果。YOLOv7 算法结构图如图 1 所示。

### 1.2 RMF-YOLO 算法的网络结构

虽然 YOLOv7 网络具有良好的检测性能,但是对于铁路环境中异物特征不明显、遮挡、尺度变化较大的情况仍存在检测精度不高的问题,故而需要进行改进和优化。本文在 YOLOv7 网络结构上进行了改进,提出了一种基于多尺度特征融合的 RMF-YOLO 网络模型,使其可以适用于检测铁轨异物,准确识别复杂环境中大小不一的异物目标。RMF-YOLO 算法的整体结构如图 2 所示。

## 2 基于 RMF-YOLO 算法的异物目标

RMF-YOLO 算法集成了 ICBAM 混合注意力模块、GhostConv 模块、改进的 BiFPN 特征融合网络和 WIoUv3 损失函数,以提高对具有复杂特征的多个对象的识别能力。首先,将 ICBAM 注意力模块添加在特征提取网络中,帮助网络有效抑制无关对象的干扰,从而实现对外物的准确检测,提高特征的表达力。其次,用 GhostConv 代替普通卷积,可以让模块在几乎没有额外计算复杂度的情况下运行,提高模型性能,减少模型参数,有助于提高铁轨异物检测的速度。然后,在 PANet 的基础上进行改进,构建一种新的具有聚合路径的加权双向特征金字塔网络 N-BiFPN,从而有效地优化特征融合过程,可以更好地获取异物目标的多尺度特征,获得更高的检测精度。考虑对难以检测异物目标的关注,提出将 WIoUv3 损失作为边界框损伤的计算方法,有助于进一步提高模型的性能,具体见图 2 特征融合网络部分。

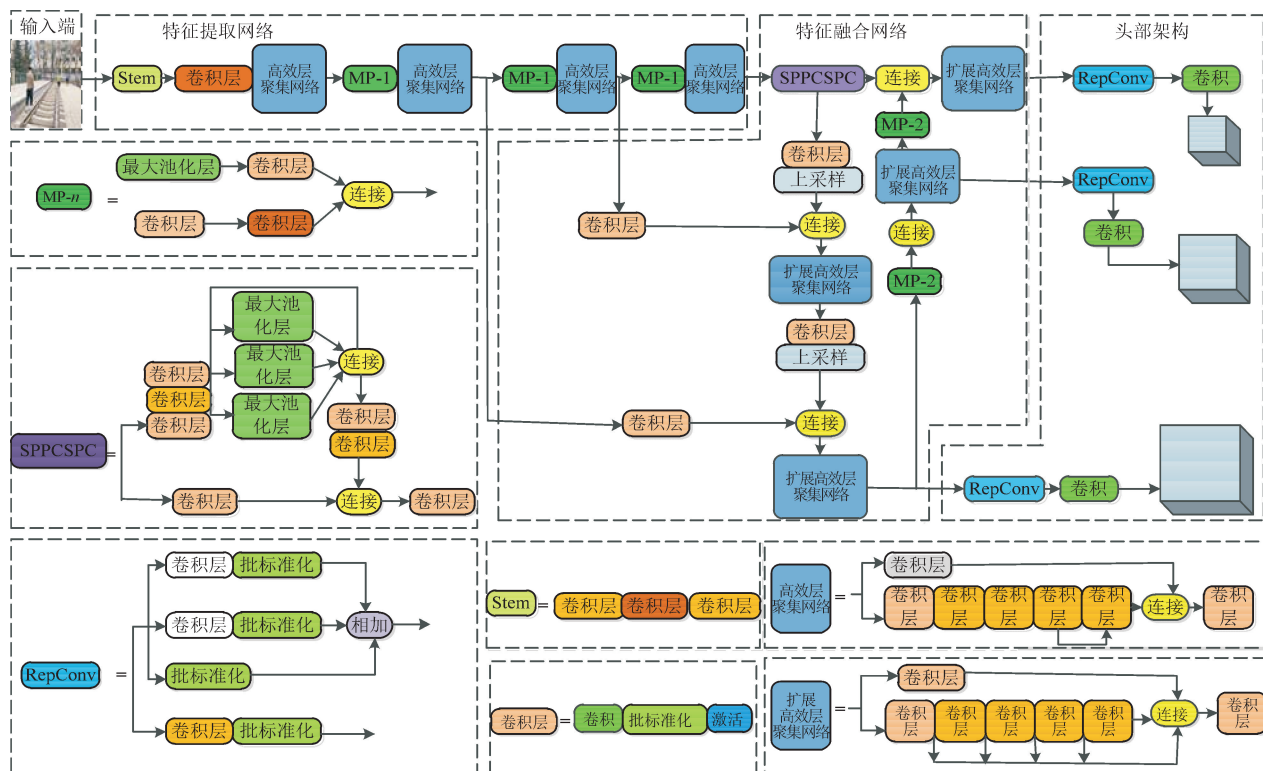


图 1 YOLOv7 算法结构图

Fig. 1 YOLOv7 network structure diagram

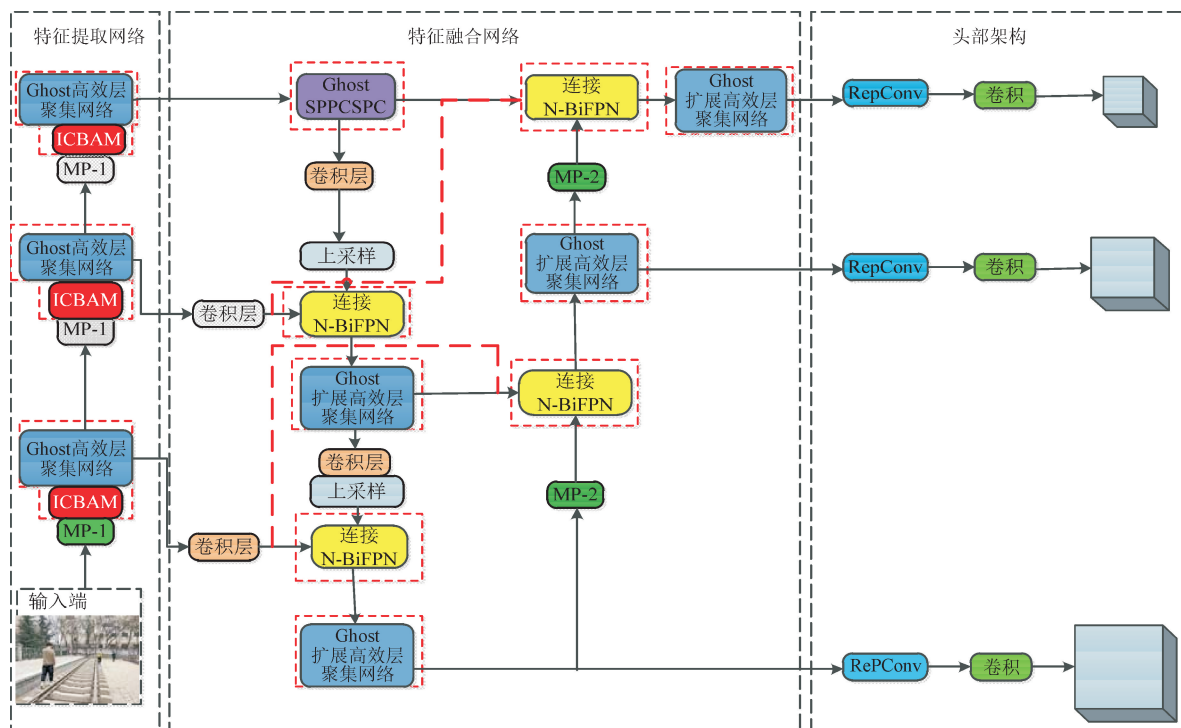


图 2 RMF-YOLO 算法结构图

Fig. 2 RMF-YOLO algorithm structure



## 2.1 混合注意力机制

在铁路轨道场景中,由于光照、遮挡、背景等因素的影响,可能产生对入侵异物的漏检和错检,尤其是对远端小物体的检测。为了进一步解决复杂环境中对铁轨入侵异物的检测,降低环境信息对异物检测的干扰,本文在卷积注意力模块(CBAM)<sup>[20]</sup>的基础上进行改进,得到ICBAM混合注意力机制,聚集特征图的通道和空间维度,从而进一步提高检测精度。ICBAM由两个相互连接的子模块组成,在输入和输出之间新增一个残差连接,以满足更短的有效梯度传播路径。通过使用具有sigmoid函数的全局最大池化来增强由全局平均池化生成的信道映射,对通道注意力模块进行改进,得到增强型通道注意力模块(ECAM),从而增强通道特征聚合并减少多层感知器(MLP)计算。通过ECAM和空间注意力模块(SAM),特征提取网络可以自适应地调整输入的特征图,并集中在其中重要的特征部分。

ICBAM混合注意力机制计算过程表示为

$$F_x = M_{eca}(F) \otimes F \quad (1)$$

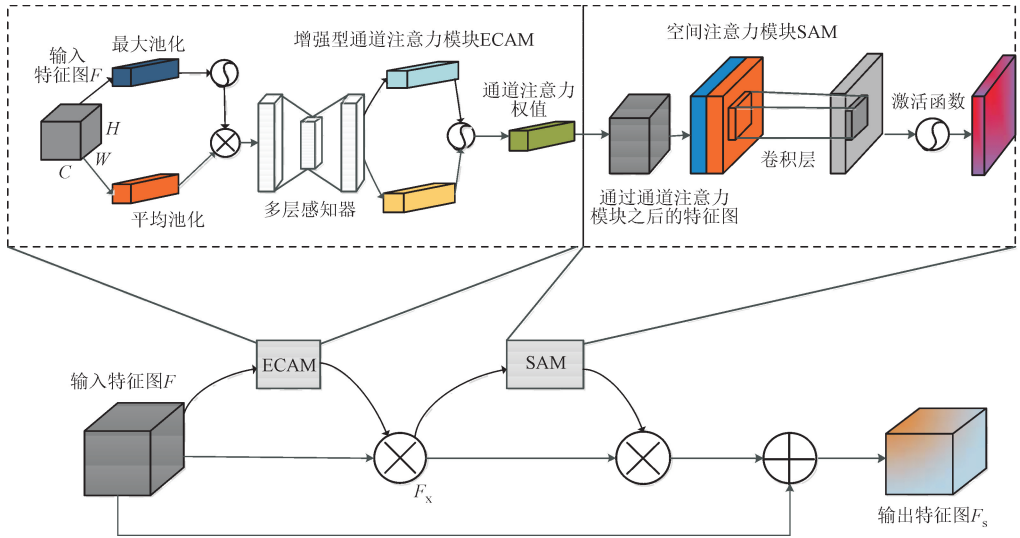


图3 ICBAM注意力机制结构图

Fig. 3 Structure of ICBAM attention mechanism

在混合注意力机制下,输入特征图 $F$ 经过ECAM,对特征图在空间维度上进行压缩。ECAM突出显示输入特征有意义的通道,为了有效地计算通道注意力特征,使用sigmoid函数将全局平均池化操作乘以全局最大池化操作,以增强通道注意力。然后,将得到的参数通过多层感知器MLP进行叠加,进一步压缩输入特征图的空间维度,以生成通道注意力特征 $M_{eca}(F)$ 。输出特征图计算公式为

$$M_{eca}(F) = \sigma(W_2(W_1(F_{avg}^{ec} \sigma(F_{max}^{ec})))) \quad (3)$$

式中: $\sigma$ 表示sigmoid函数; $F_{avg}^{ec}$ 和 $F_{max}^{ec}$ 分别为通道

$$F_s = (M_{sa}(F_x) \otimes F_x) \oplus F \quad (2)$$

式中:特征图 $F \in \mathbf{R}^{C \times H \times W}$ , $C$ 为特征图的通道数, $H \times W$ 为特征图大小; $M_{eca}(F)$ 为通道注意力特征; $M_{sa}(F_x)$ 为空间注意力特征; $F_x$ 表示通过ECAM得到的特征图。

在提取原网络特征之后,利用ICBAM注意力机制将通道注意力和空间注意力串联。在模型训练过程中对于异物目标分配更多的权值参数,对原有的入侵异物特征图进行信息提炼,突出前景异物目标的权重,使网络在学习过程中更关注于前景异物的特征,提高对目标对象的特征提取能力。将ICBAM注意力机制应用于铁路周界异物入侵检测,可以提升算法对异物目标的提取能力,改善由于复杂环境引起的异物目标特征信息减弱,致使检测的准确性降低等问题。应用该注意力机制可以保留更多的特征信息来提升网络对全局信息的学习,进而提升网络对异物目标的检测准确率。ICBAM混合注意力模块结构如图3所示。

注意力机制的全局平均池化运算和最大平均池化运算; $W_1$ 和 $W_2$ 分别为多层感知器第1层和第2层权重。

将给定特征图 $F_x$ 送入SAM,沿通道维度聚合空间信息,连接平均池化和最大池化特征图。然后,通过卷积和sigmoid函数激活,得到空间注意力特征 $M_{sa}(F)$ ,计算公式为

$$M_{sa}(F) = \sigma(f^{7 \times 7}(F_{avg}^s; F_{max}^s)) \quad (4)$$

式中: $f^{7 \times 7}$ 表示滤波器为 $7 \times 7$ 的卷积运算, $7 \times 7$ 表示卷积核的尺寸; $F_{avg}^s$ 和 $F_{max}^s$ 分别为空间注意力机

制的平均池化运算和最大平均池化运算。

## 2.2 Ghost 模块的构建

由于卷积神经网络的不断发展和对于嵌入式设备需求的增加,在有限的计算资源下部署更高效的神经网络已成为神经网络发展的迫切需求。Han 等<sup>[21]</sup>利用特征图的冗余特性进行一系列的线性变换,提出了一种更轻量级的卷积模块 GhostConv,使用更少的参数和计算实现更多的特征图。普通卷积和 GhostConv 卷积对比如图 4 所示。

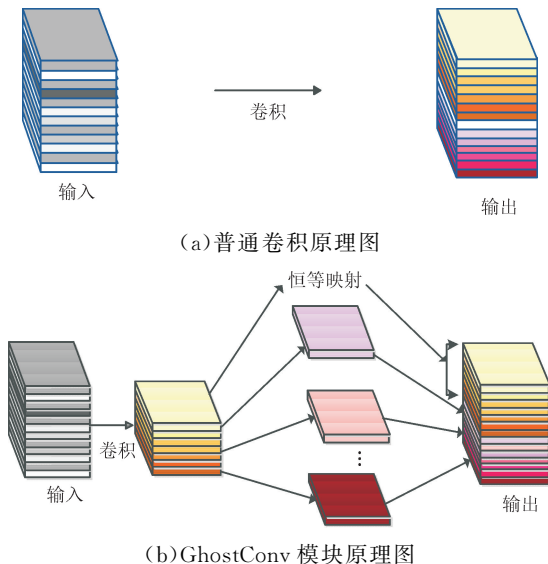


图 4 普通卷积和 GhostConv 卷积对比

Fig. 4 Comparison of normal convolution and GhostConv convolution

为了减少生成冗余特征图所需的滤波器的数量,本文利用这种轻量级的 GhostConv 模块代替 YOLOv7 模型中的普通卷积模块,可减少网络的计算量和参数量。使用 Ghost 卷积层替换标准卷积层,但保留了批处理归一化和 SiLU 激活函数,使得 CBS 模块变为 GBS 模块,也使得其他模块转变为 Ghost 对应模块。

GhostConv 卷积在常规卷积的基础上进行了线性运算,首先是通过提取特征信息进行少量卷积,获得一般特征图;然后利用普通卷积给出的固有特性映射,将一般特征图中的每一个通道的特征做线性运算,生成同等通道数的 Ghost 特征图,可以在几乎不影响检测精度的同时有效降低普通卷积的冗余性;最后将两次获得的特征图堆叠起来,形成新的输出特征图。GhostConv 卷积在保证模型性能的情况下,大幅度减少参数和计算量,从而实现更高效、更轻量级的网络部署。

假设输入特征图的高、宽和通道数分别为  $c$ 、 $h$ 、

$w$ ,经过一次卷积后输出特征的高度和宽度为  $H$ 、 $W$ ,卷积核数量为  $n$ ,卷积核大小为  $k$ ,线性变换卷积核大小为  $d$ ,经过  $s$  次变换,普通卷积操作的计算量与 GhostConv 卷积的计算量之比为

$$r_s = \frac{nHWckk}{(n/s)HWckk + (s-1)(n/s)HWdd} = \frac{ckk}{(1/s)ckk + ((s-1)/s)dd} \approx \frac{sc}{s+c-1} \approx s \quad (5)$$

式中:  $n/s$  为第 1 次变换时的输出通道数目;  $s-1$  是恒等映射不需要进行计算,但也算作第 2 次变换中的一部分。

计算卷积核的参数量为

$$r_c = \frac{nckk}{(n/s)ckk + (s-1)(n/s)dd} \approx \frac{sc}{s+c-1} \approx s \quad (6)$$

结合式(5)和式(6)可以看出,计算量和参数量之比受变换数量  $s$  的影响,即生成特征图越多,模型加速效果越好。在模型中引入 GhostConv 卷积可以有效减少模型的计算量和参数量,提高模型的运行速度和效率。

## 2.3 多尺度特征融合网络

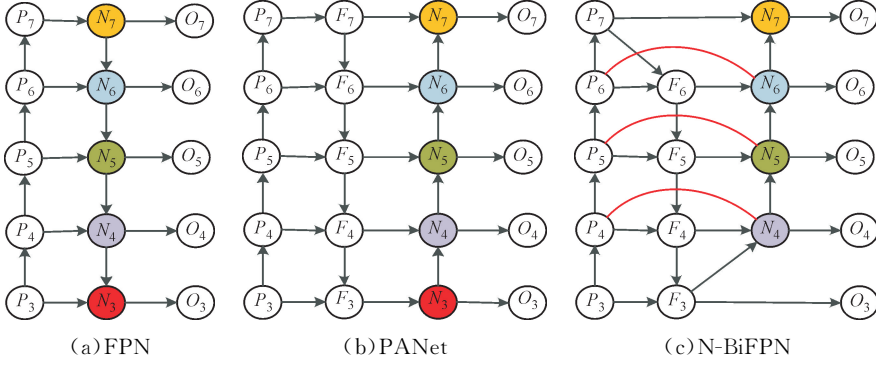
有效的特征融合可以增强目标检测的精度,不同特征金字塔如图 5 所示。特征金字塔网络(FPN)<sup>[22]</sup>提出了一种自顶向下的路径,具有横向连接,通过构造多尺度特征融合结构来传播语义上较强的路径。在 FPN 中建立一条自上而下的通路,进行特征融合,用融合后的具有更高语义信息的特征图进行预测,可以提高一定的精度,结构如图 5(a)所示。但是经过研究发现,这种自顶向下的 FPN 网络受到单向信息流约束的影响,精度不能得到有效提高。基于此,PANet 引入自底向上的路径,缩短底层与顶层特征之间的信息路径,通过双向融合网络深层特征图的语义信息和浅层特征图的定位信息提高了检测精度,让顶层网络在保持语义特征的同时增加了空间特征<sup>[23]</sup>。YOLOv7 在特征融合网络层采用 PANet 结构,该结构具有自顶向下和自底向上的特点,结构如图 5(b)所示。这种结构增加了参数量,降低了模型效率,只通过简单的求和运算来融合来自不同层的特征,导致特征融合不够充分。

为了解决这一问题,在以 PANet 为基础的框架上,借鉴加权双向特征金字塔网络(BiFPN)<sup>[24]</sup>结构的设计思想,本文提出了 N-BiFPN 作为多尺度特征融合模块。该模块旨在实现不同尺度特征图的融合,从而加强特征信息表达,实现对特征的多尺度融

合和上下文信息的丰富,结构如图 5(c)所示。

N-BiFPN 在 PANet 的基础上,移除 PANet 中的单输入节点,该节点所包含信息较少,没有进行特征融合,删除该节点可以简化多尺度融合结构,减少计算量。在原始输入节点和输出节点之间添加新的聚合路径,在节省资源消耗的同时融合更多的特征信息。N-BiFPN 通过构造双向通道实现跨尺度连接,将特征提取网络中的特征直接与自下而上路径

中的相对大小特征融合。自顶向下的路径可以提高网络的识别能力,向底层网络传递强语义信息,自底向上的路径增强网络特征层次的定位能力,保留了更浅层的细节特征。为改善多尺度特征融合效果,针对每个输入特征设置了一个可学习的特征融合因子,并采用通道级联的方式进行特征融合的操作。这种方式在保留更多特征信息的同时,可以简化融合计算,降低冗余操作。



$P_3 \sim P_7, N_3 \sim N_7, O_3 \sim O_7, F_3 \sim F_7$  均为网络节点特征

图 5 不同特征金字塔结构

Fig. 5 Structure of different characteristic pyramids

N-BiFPN 相当于一个具有权重的特征融合机制,权重信息在所有特征层之间共享,通过有效的路径优化实现路径增强,确保网络在训练过程中逐渐学习每个输入特征的重要信息。同时,多尺度特征融合可以提高模型的感受野,在不同尺度上提取和融合多尺度的特征信息,提高对多尺度目标的检测性能,使模型能够更好地理解所采集铁轨图像中的异物。这种多尺度的特征融合网络可以提高检测精度,使得模型能够在不同尺度物体上进行检测。

N-BiFPN 结构采用快速归一化融合方法进行权值融合,其加权特征融合公式为

$$O = \sum_i \frac{w_i}{\epsilon + \sum_j w_j} I_i \quad (7)$$

式中:  $O$  表示加权特征融合的输出特征;  $I_i$  表示输入的特征;  $w_i, w_j$  表示可学习的特征融合因子;  $\epsilon$  表示用于约束数值振荡的极小值学习率。

以节点  $P_4$  为例, N-BiFPN 特征融合过程如下

$$F_4 = \text{Conv} \left( \frac{w_3 P_4 + D_{BS}(P_5)}{w_3 + w_4 + \epsilon} \right) \quad (8)$$

$$N_4 = \text{Conv} \left( \frac{w_3 P_4 + w_4 F_4 + w_4 D_{BS}(N_3)}{w_3 + w_4 + w_5 + \epsilon} \right) \quad (9)$$

式中:  $P_i$  为第  $i$  层的输入特征;  $F_i$  为自上而下的中

间特征;  $N_i$  为自下而上的输出特征; Conv 表示卷积运算操作;  $w_i$  为各层对应的权重,描述特征融合过程中各个特征的重要程度;  $D_{BS}$  为输入的上采样或下采样运算。

## 2.4 损失函数的改进

损失函数用来衡量预测结果与实际标签之间的差异,合理设计的损失函数可以显著提高模型的检测性能。边界框损失函数作为目标检测损失函数的重要组成部分,良好的改进将为目标检测模型带来显著的性能提升。原 YOLOv7 采用的 CIoU (complete intersection over union) 损失函数<sup>[25]</sup>,考虑了预测框与真实框的长宽比、中心点距离、重叠面积,但是对于小标注框和低质量标注框的回归有一定局限性,且 CIoU 损失函数的计算十分耗时,最终导致较长的训练时间。CIoU 的相关公式如下

$$L_{\text{CIoU}} = 1 - I_{\text{ou}} + \frac{\rho^2(b, b^{\text{gt}})}{c^2} + \alpha v \quad (10)$$

$$\alpha = \frac{v}{(1 - I_{\text{ou}}) + v} \quad (11)$$

$$v = \frac{4}{\pi^2} \left( \arctan \left( \frac{w_{\text{gt}}}{h_{\text{gt}}} \right) - \arctan \left( \frac{w}{g} \right) \right)^2 \quad (12)$$

式中:  $I_{\text{ou}}$  为交并比,即预测框与真实框的交集除以二者的并集;  $b, b^{\text{gt}}$  分别为预测框和真实框的中心点;  $\rho$  为两个中心点之间的欧式距离;  $\alpha$  表示权重函

数;  $v$  为用来衡量预测框与真实框之间的宽高比的一致性;  $w_{gt}$ 、 $h_{gt}$  为真实框的宽和高。相关参数示意如图 6 所示, 其中  $c_h$  和  $c_w$  为预测框和真实框形成的最小封闭框的高度和宽度。

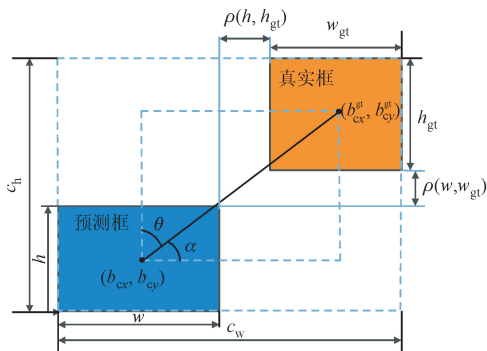


图 6 损失函数相关参数示意图

Fig. 6 Schematic of the parameters associated with the loss function

CIoU 中预测框的宽和高  $w$ 、 $h$  相对于  $v$  的梯度如下

$$\frac{\partial v}{\partial w} = \frac{8}{\pi^2} \left( \arctan \frac{w}{h} - \arctan \frac{w_{gt}}{h_{gt}} \right)^2 \frac{h}{h^2 + w^2} \quad (13)$$

$$\frac{\partial v}{\partial h} = \frac{8}{\pi^2} \left( \arctan \frac{w}{h} - \arctan \frac{w_{gt}}{h_{gt}} \right)^2 \frac{w}{h^2 + w^2} \quad (14)$$

由式(13)和式(14)可以看出, 以上两个梯度呈现相反态势, 即  $w$ 、 $h$  不能同时增大或者缩小。故而, 为降低低质量标注框对检测性能的干扰, 进一步提高检测精度, 本文引入  $L_{WIoUv3}$  损失函数<sup>[26]</sup>。WIoU 损失函数是基于动态非单调聚集机制设计的, 采用合理的梯度增益分配策略, 使模型更聚焦于低质量的锚框。WIoU 因为没有对纵横比进行计算, 相比于 CIoU 具有更快的速度, 耗时少。WIoU 相关公式如下

$$L_{WIoUv3} = R_{WIoU} L_{IoU} \quad (15)$$

$$R_{WIoU} = \exp \left( \frac{(b_{cr}^{gt} - b_{cr})^2 + (b_{cy}^{gt} - b_{cy})^2}{(c_w^2 + c_h^2)^*} \right) \quad (16)$$

$$I_{IoU} = 1 - I_oU \quad (17)$$

式中:  $R_{WIoU}$  为惩罚项;  $L$  为不同类型损失函数;  $(b_{cr}, b_{cy})$  和  $(b_{cr}^{gt}, b_{cy}^{gt})$  分别表示真实框和预测框的中心点坐标; 上标 \* 表示将  $c_h$  和  $c_w$  从计算图中分离, 防止  $R_{WIoU}$  产生阻碍收敛的梯度。

在式(15)的基础上, 构造单调聚集系数  $L_{IoU}^*$ , 得到

$$L_{WIoUv2} = L_{IoU}^* L_{WIoUv1}, \gamma > 0 \quad (18)$$

在模型训练过程中,  $L_{IoU}^*$  的梯度增益随着  $L_{IoU}$  的减小而减小, 故而引入  $L_{IoU}$  的均值作为归一化因子, 得到

$$L_{WIoUv2} = \left( \frac{L_{IoU}^*}{\bar{L}_{IoU}} \right)^\gamma L_{WIoUv1} \quad (19)$$

式中:  $\bar{L}_{IoU}$  为动量  $m$  的平均运行值; 动态更新归一化因子使梯度增益  $r = \left( \frac{L_{IoU}^*}{\bar{L}_{IoU}} \right)^\gamma$  保持在较高水平。

在式(19)的基础上, 引入动态非单调聚集系数, 得到

$$L_{WIoUv3} = r L_{WIoUv1} \quad (20)$$

$$r = \frac{\beta}{\delta \alpha^{\beta-\sigma}} \quad (21)$$

$$\beta = \frac{L_{IoU}^*}{L_{IoU}} \in [0, +\infty] \quad (22)$$

式中:  $\beta$  表示离群值;  $\alpha$  和  $\delta$  为控制梯度增益  $r$  的超参数。

WIoUv3 采用合理的梯度增益分配策略, 动态优化损失中高低质量锚框的权值, 使模型聚集于平均质量样本, 提高模型的整体性能。与 YOLOv7 模型中的原始损失计算相比, 在铁轨异物检测任务中, WIoUv3 损失函数可以动态优化不同尺度异物的损失权重, 提高模型的检测性能。

## 3 实验与结果分析

### 3.1 铁路轨道异物检测存在的问题

在实际铁路轨道交通场景中, 铁轨异物可能会因为列车在运行过程中, 导致摄像机采集的图像视角差异, 增加了异物目标误检的可能性。不同天气状况也可能导致误检率上升。为提高模型的泛化能力, 需要在训练过程中包含不同天气状况下的数据; 铁路轨道异物目标具有随机性, 容易发生异物的遮挡, 导致多个目标中可能出现漏检误检的问题。

现有的公开数据集并不适用于铁路轨道异物入侵检测, 需要构建专门的数据集, 并采取相应的方法以解决问题, 提高异物检测的准确率。

### 3.2 数据集

由于目前尚无公开的铁路入侵异物数据集, 根据实验的需要, 建立铁轨异物目标入侵的数据集, 模拟真实环境中的入侵。该数据集分为 2 部分, 一部分是从公开数据集(Pascal voc2007/2012)中选取具有与本文研究对象特征相近的数据; 另一部分从不同的拍摄角度和场景中采集所有可能妨碍铁路系统正常运行的物体, 以创建铁路异物检测数据集。在



实验中,选择行人、汽车、火车和动物等典型异物作为研究对象。

为增加模型在实际情况下的泛化能力,防止因数据不足而产生的过拟合现象,对所获得的原始数据采取图像增强的方法,增加训练数据集中图像的变化。通过随机调整对比度、增加噪声、旋转、调节亮度等操作,将训练样本的数量增加了 35%,进一步提高了模型的泛化能力,防止数据集中存在可能的偏差。

数据集分为 6 500 个训练集、1 000 个验证集和 1 000 个测试集,用于评估每种策略的性能实验。为了制备数据集,由于检测任务是有监督的机器学习问题,因此需要将地面真实标签返送到模型以进行训练。使用开源图像标注工具 LabelImg,对经过预处理的数据集进行人工标注,部分数据集标注结果如图 7 所示。



图 7 部分数据集标注结果

Fig. 7 Labeling results of some datasets

| 真实值 | 行人            | 牛             | 狗             | 火车            | 汽车            | 猫            |
|-----|---------------|---------------|---------------|---------------|---------------|--------------|
|     | 221<br>75.95% | 27<br>9.28%   | 18<br>6.19%   | 5<br>1.72%    | 7<br>2.41%    | 13<br>4.47%  |
|     | 6<br>4.05%    | 116<br>78.38% | 11<br>7.43%   | 3<br>2.03%    | 3<br>2.03%    | 9<br>6.08%   |
|     | 9<br>4.71%    | 10<br>5.24%   | 142<br>74.35% | 9<br>4.71%    | 7<br>3.66%    | 14<br>7.33%  |
|     | 4<br>2.82%    | 8<br>5.63%    | 8<br>5.63%    | 111<br>78.17% | 10<br>7.04%   | 1<br>0.70%   |
|     | 2<br>1.59%    | 4<br>3.17%    | 6<br>4.76%    | 11<br>8.73%   | 100<br>79.37% | 3<br>2.38%   |
|     | 2<br>1.96%    | 4<br>3.92%    | 8<br>7.84%    | 4<br>3.92%    | 7<br>6.86%    | 77<br>75.49% |
|     |               | 预测值           |               |               |               |              |

(a)YOLOv7

3.3 实验环境

本文的实验环境:基于 Python3. 9 版,CUDA 11. 2 版,编程软件为 PyCharm,实验模型的搭建、训练和测试均在深度学习 Pytorch 框架下完成,采用 Windows10 操作系统,搭载 NVIDIA RTX3060 显卡。

3.4 评价指标

本文采用以下指标来评估模型的性能:平均精确度(记作  $P_{mAP}$ )、精确度( $P$ )、召回率( $R$ )、检测速度(FPS)和计算量(GLOPs)。计算公式如下

$$P = \frac{T_P}{T_P + F_P}$$
(23)

$$R = \frac{T_P}{T_P + F_N}$$
(24)

$$P_{mAP} = \frac{1}{n} \sum_{i=1}^n P_A = \int_0^1 P(R) dR$$
(25)

式中:  $T_P$  为阳性样本被正确分类;  $T_N$  为阴性样本被正确分类;  $F_P$  为阴性样本被错误分类;  $F_N$  为阳性样本被错误分类;  $n$  为所有类别数量;  $P_A$  为目标类别的平均精确度。

3.5 铁轨异物入侵检测实验结果

3.5.1 混淆矩阵

混淆矩阵又称误差矩阵,是一种常见的用于评估分类器性能的方法,反映了样本数据的真实类别与分类器的预测类别之间的关系<sup>[27]</sup>。混淆矩阵中横轴表示预测值,纵轴表示真实值,图中对角线为正确识别的样本数,颜色越深表示数量越多。通过混淆矩阵对测试结果进行可视化,不仅能统计每个类别的正确识别数,而且可以了解错误识别的分布情况。在相同数据集中,由 YOLOv7 和 RMF-YOLO 检测到的混淆矩阵如图 8 所示。

| 真实值 | 行人            | 牛             | 狗             | 火车            | 汽车            | 猫            |
|-----|---------------|---------------|---------------|---------------|---------------|--------------|
|     | 238<br>81.79% | 18<br>6.19%   | 16<br>5.50%   | 0<br>0.00%    | 8<br>2.75%    | 11<br>3.78%  |
|     | 2<br>1.35%    | 129<br>87.16% | 5<br>3.38%    | 1<br>0.68%    | 2<br>1.35%    | 9<br>6.08%   |
|     | 5<br>2.65%    | 4<br>2.12%    | 158<br>83.60% | 6<br>3.17%    | 3<br>1.59%    | 13<br>6.88%  |
|     | 0<br>0.00%    | 3<br>2.10%    | 2<br>1.40%    | 126<br>88.11% | 8<br>5.59%    | 4<br>2.80%   |
|     | 0<br>0.00%    | 1<br>0.79%    | 3<br>2.38%    | 9<br>7.14%    | 111<br>88.10% | 2<br>1.59%   |
|     | 1<br>0.98%    | 8<br>7.84%    | 5<br>4.90%    | 1<br>0.98%    | 2<br>1.96%    | 85<br>83.33% |
|     |               | 预测值           |               |               |               |              |

(b)RMF-YOLO

图 8 混淆矩阵对比

Fig. 8 Confusion matrix comparison

从图 8 可以看出:本文提出的 RMF-YOLO 算法模型中混淆矩阵的对角线区域颜色比 YOLOv7 深,表明该算法正确预测目标类别的能力得到了增强。在原 YOLOv7 中,对于行人、狗、猫等小型物体

的漏检率较高,改进后的模型降低了这些类别的漏检率,检测正确率分别提高了 5.84%、8.78%、9.25%、7.84%。该算法各类别异物部分检测结果如图 9 所示,图中数值为目标类别的置信度。



图 9 部分检测结果图  
Fig. 9 Images of some detection results

3.5.2 对比实验

(1)不同注意力机制对比实验。为了验证本文所提出的 ICBAM 混合注意力模块的有效性,进行

实验验证,并将 ICBAM 混合注意力模块与其他常用的注意力模块进行比较,实验结果如表 1 所示, $P_{mAP@0.5}$  表示  $I_{ov}$  为 0.5 时的平均精度。

表 1 不同注意力机制模块处理效果的对比结果

Table 1 Comparative results of the processing effects of different attention mechanism modules

| 算法     | 输入图像尺寸/(像素×像素) | 精确度/% | 召回率/% | $P_{mAP@0.5}$ /% | 检测速度/(帧·s <sup>-1</sup> ) | 计算次数/10 <sup>9</sup> |
|--------|----------------|-------|-------|------------------|---------------------------|----------------------|
| YOLOv7 | 640×640        | 90.4  | 89.7  | 80.2             | 85                        | 104.5                |
| +CA    | 640×640        | 90.6  | 89.9  | 80.4             | 94                        | 104.8                |
| +ECA   | 640×640        | 90.7  | 90.2  | 80.4             | 78                        | 105.5                |
| +SENet | 640×640        | 91.3  | 90.1  | 80.5             | 102                       | 104.9                |
| +CBAM  | 640×640        | 91.2  | 90.7  | 80.8             | 85                        | 106.2                |
| +ICBAM | 640×640        | 91.8  | 91.2  | 81.4             | 90                        | 105.9                |

由表 1 可得,在基准模型中分别引进 5 种注意力机制后,本文所提出的 ICBAM 混合注意力机制检测精度优于其他 4 种注意力机制。与其他注意力机制相比,+ICBAM 的精确度、召回率和平均精确度  $P_{mAP}$  分别提高了 1.4%、1.5% 和 1.2%,图片处理速度提高到 90 帧/s。同时,ICBAM 混合注意力机制在保持良好检测精度的同时保证了检测的实时性,能够有效提升模型对不同尺度异物的检测能力,验证了该方法的可行性。在铁路轨道异物入侵检测过程中,利用深层次特征提取网

络虽然能增强入侵异物的特征感知,但是会影响模型的推理速度,提高网络的复杂度。使用注意力机制既能保证模型的速度,又能在一定程度上消除背景噪声干扰,提高特征提取网络对异物特征的感知能力。

(2)不同场景下多种检测算法检测结果的对比实验。使用训练好的检测模型分别针对遮挡、多尺度及沙尘天气等不同复杂场景,利用多种主流检测算法对不同的入侵异物进行实验对比,结果如图 10~图 12 所示。





图 10 异物遮挡场景下不同算法的检测结果

Fig. 10 Detection results of different algorithms in foreign object occlusion scene

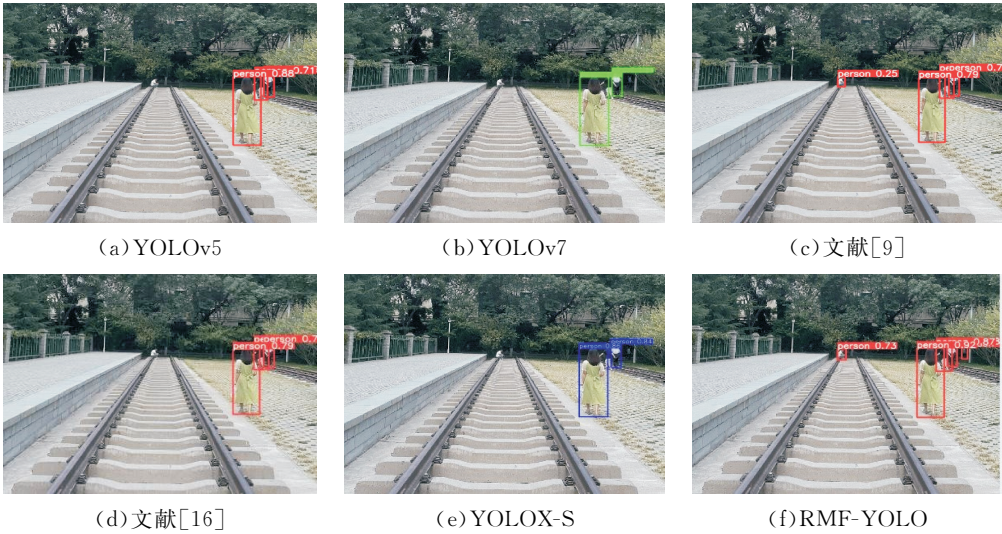


图 11 多尺度异物场景下不同算法的检测结果

Fig. 11 Detection results of different algorithms in multi-scale foreign object scenarios

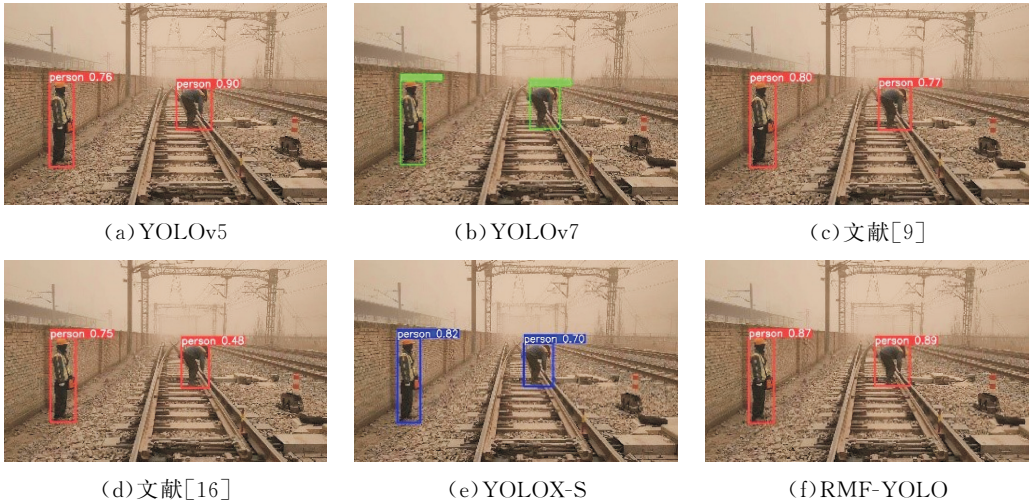


图 12 沙尘场景下不同算法的检测结果

Fig. 12 Detection results of different algorithms in sand and dust scenarios

对于异物具有遮挡的场景, YOLOv5 和 YOLOv7 算法均存在不同程度因遮挡发生少检、漏检情况;文献[9]算法对于近端的异物检测置信度较高,但对于远端的异物存在漏检,且置信度较低;文献[16]算法部分遮挡人物也出现了漏检的现象;YOLOX-S 算法<sup>[28]</sup>对于遮挡异物的检测有所提高,但是对于特殊姿势的检测率较低;本文算法 RMF-YOLO 可以检出全部异物,对于特殊姿势的异物检测更准确,且定位更加精确,置信度更高,证明本文的算法可以有效提高模型定位精度,提高模型鲁棒性。图 10 检测结果中工人手中的工具通常不会构成安全隐患,因此本文方法主要关注铁轨上的异物目标而非工人手中的工具,未单独检测到工具不影响我们方法的有效性和可靠性。

对于不同尺度的异物场景,尤其针对由于摄像头距离引起的远端异物的检测, YOLOv5、YOLOv7、YOLOX-S 和文献[16] 算法均存在像素占用小的漏检目标的情况,仅检测出大尺度的异物

目标。文献[9]算法虽然可以检测出较小尺寸的异物,但是置信度较低。本文算法 RMF-YOLO 可以有效检测出不同尺度的入侵异物,与 YOLOv7 相比,RMF-YOLO 在准确性和锚框回归方面表现出性能优势,证明本文方法提高了对不同尺度的铁轨入侵异物的感知能力。

对于沙尘复杂环境下的场景, YOLOv5、文献[9]和文献[16]算法受到环境的影响,较大尺度异物的置信度有所下降;文献[16]算法和 YOLOX-S 对于特殊姿势的异物检测精确度低;与 YOLOv7 相比,本文 RMF-YOLO 算法异物检测的置信度更高,对于不同尺度异物检测的精确度均有所提高,可以证明本文算法对于沙尘环境影响下的特征提取也有改善。

(3)主流目标检测算法对比实验。为了进一步验证本文算法的有效性和优越性,在同样的训练条件和数据的前提下,将本文提出的算法与其他算法进行对比,结果见表 2。

| 表 2 主流目标检测算法处理效果对比结果                                                               |                |                  |                           |                      |
|------------------------------------------------------------------------------------|----------------|------------------|---------------------------|----------------------|
| Table 2 Mainstream target detection algorithm processing effect comparison results |                |                  |                           |                      |
| 算法                                                                                 | 输入图像尺寸/(像素×像素) | $P_{mAP@0.5}/\%$ | 检测速度/(帧·s <sup>-1</sup> ) | 计算次数/10 <sup>9</sup> |
| SSD-MonileNet                                                                      | 640×640        | 77.4             | 59                        | 132.3                |
| Mask R-CNN                                                                         | 640×640        | 76.4             | 36                        | 56.5                 |
| YOLOv3-SPP <sup>[29]</sup>                                                         | 640×640        | 78.2             | 80                        | 48.6                 |
| YOLOv5                                                                             | 640×640        | 79.6             | 83                        | 109.0                |
| FCOS <sup>[30]</sup>                                                               | 640×640        | 78.3             | 41                        | 45.1                 |
| YOLOX-S                                                                            | 640×640        | 80.4             | 68                        | 164.5                |
| YOLOv7-tiny                                                                        | 640×640        | 78.6             | 118                       | 13.7                 |
| YOLOv7                                                                             | 640×640        | 80.2             | 85                        | 104.5                |
| RMF-YOLO                                                                           | 640×640        | 85.7             | 90                        | 91.7                 |

通过表 2 中不同算法的实验结果可以看出,本文算法 RMF-YOLO 的平均精确度  $P_{mAP}$  指标明显高于其他主流目标检测算法,具有较高的检测精度,比原 YOLOv7 的  $P_{mAP}$  提高了 5.5%,比 YOLOv5 提高了 6.1%,比 YOLOv3-SPP 提高了 7.5%。与 SSD-MonileNet、Mask R-CNN、YOLOX-S 和 YOLO7-tiny 算法相比,本文算法的检测准确率分别提高了 8.3%、9.3%、5.3%和 5.5%。相对于具有相近检测速度 YOLOv5 和 YOLOv3-SPP 算法来说,本文算法的检测准确率明显优势,虽然 YOLOv7-tiny 具有更高的检测速度,但其检测精度相对较低,无法满足对复杂的铁路环境下的异物检

测。本文算法的平均精确度  $P_{mAP}$  和检测速度都高于原 YOLOv7 目标检测算法,检测精度上有了显著的提高,每秒帧数提高了 5.88%,达到了速度和精度的平衡。综上所述,本文所提算法与其他主流算法相比更加符合复杂道路场景目标检测的实时性和准确性需求。

3.5.3 消融实验

消融实验是通过在完整模型中移除一个或多个模块,来检测对模型的性能影响。为了验证本文提出的优化措施的效果,依次对 RMF-YOLO 网络进行缩减,分别在测试集上进行消融实验,以评价各个优化模块对检测算法性能的影响,结果如表 3 所示。



表 3 不同消融方法处理结果的实验对比结果

Table 3 Experimental comparison results of the treatment results of different ablation methods

| 组别 | ICBAM | GhostConv | N-BiFPN | 损失函数 | 输入图像尺寸/(像素×像素) | $P_{\text{mAP}@0.5}/\%$ | 检测速度/(帧·s <sup>-1</sup> ) | 计算次数/10 <sup>9</sup> |
|----|-------|-----------|---------|------|----------------|-------------------------|---------------------------|----------------------|
| 1  |       |           |         |      | 640×640        | 80.2                    | 85                        | 104.5                |
| 2  |       | ✓         |         |      | 640×640        | 80.4                    | 93                        | 76.9                 |
| 3  | ✓     | ✓         |         |      | 640×640        | 81.3                    | 91                        | 82.1                 |
| 4  | ✓     |           | ✓       |      | 640×640        | 82.4                    | 73                        | 115.6                |
| 5  | ✓     |           |         | ✓    | 640×640        | 81.8                    | 78                        | 106.4                |
| 6  | ✓     | ✓         | ✓       |      | 640×640        | 82.6                    | 81                        | 97.3                 |
| 7  | ✓     | ✓         | ✓       | ✓    | 640×640        | 85.7                    | 90                        | 91.7                 |

注:实验以原始 YOLOv7 为基准,“✓”表示进行添加或改进。

从表 3 可以看出,第 2 组在原始模型基础上融合了 GhostConv 模块,与普通卷积模块相比,GhostConv 模块能够提高模型精度,同时降低模型复杂度,检测速度提高了 9.41%,浮点运算次数减少了 26.41%。第 3 组实验在第 2 组的基础上加入了 ICBAM 注意力机制模块,检测计算量大幅上升,识别速度有所降低,但是检测精度提高了 1.1%,表明该模块对原始模型的精度提升起到了关键作用。第 4 组实验则在原模型中引入了 ICBAM 和 N-BiFPN 模块,在异物检测的平均精度方面有了显著提高,较原始基准模型的  $P_{\text{mAP}}$  提高了 2.2%。第 5 组实验采用了 ICBAM 和改进的损失函数,精度提升了 1.6%,证明对识别目标的尺寸和位置进行有针对性的算法改进具有意义。第 6 组实验检测精度提高了 1.4%,而检测速度降低了 4.71%。第 7 组实验,在对损失函数进行替换后,模型性能达到最佳,与原基准实验相比, $P_{\text{mAP}}$  提高了 5.5%,帧数提高到 90,计算量减少了 12.25%,表明模型具有较好的泛化能力,进一步验证了本模型方案的有效性。

4 结束语

本文针对铁路轨道异物检测中的挑战,开展了基于多尺度特征融合 RMF-YOLO 算法的铁轨异物检测研究。以 YOLOv7 为骨干网络设计并引入 ICBAM 注意力机制,增强了网络对关键特征的关注和特征提取能力。采用 GhostConv 替代原模型中的普通卷积层,提高了检测速度。提出 N-BiFPN 结构进行跨层特征融合,充分结合深层语义特征和浅层位置特征,增强了网络的特征融合能力。最后,引入 WIoUv3 损失函数,采用动态样本分配策略,降低了模型对极端样本的关注,提升了模型整体效果。

实验结果表明,本文提出的 RMF-YOLO 算法具有更高的检测精度,对于遮挡和尺度不一的异物检测能力均有所提高,改善了漏检误检的情况,满足复杂环境背景下铁轨异物检测实时性的要求。相较于 YOLOv7 算法,检测精度提升了 5.5%,每秒帧数提高了 5.88%,计算量减少了 12.25%。与现有的主流算法相比,该方法在提高检测精度的同时降低了对计算能力的要求,易于部署在资源受限的设备上。

参考文献:

[1] RISTIC-DURRANT D, FRANKE M, MICHELS K. A review of vision-based on-board obstacle detection and distance estimation in railways [J]. Sensors, 2021, 21(10): 3452.

[2] LEI Yang, HE Xiuxiu, YAO Jincao, et al. Breast tumor segmentation in 3D automatic breast ultrasound using mask scoring R-CNN [J]. Medical Physics, 2021, 48(1): 204-214.

[3] LAW H, DENG Jia. CornerNet: detecting objects as paired keypoints [J]. International Journal of Computer Vision, 2020, 128(3): 642-656.

[4] XU Renjie, LIN Haifeng, LU Kangjie, et al. A forest fire detection system based on ensemble learning [J]. Forests, 2021, 12(2): 217.

[5] CAO Jinghao, ZHANG Junju, JIN Xin. A traffic-sign detection algorithm based on improved sparse R-CNN [J]. IEEE Access, 2021, 9: 122774-122788.

[6] WANG Dashuai, LI Zhuolin, DU Xiaoqiang, et al. Farmland obstacle detection from the perspective of UAVs based on non-local deformable DETR [J]. Agriculture, 2022, 12(12): 1983.

[7] GUO Baoqing, SHI Jiafeng, ZHU Liqiang, et al. High-speed railway clearance intrusion detection with

- improved SSD network [J]. *Applied Sciences*, 2019, 9 (15): 2981.
- [8] 汪世财, 谈东奎, 谢有浩, 等. 基于激光雷达点云密度特征的智能车障碍物检测与跟踪 [J]. *合肥工业大学学报(自然科学版)*, 2019, 42(10): 1311-1317.
- WANG Shicai, TAN Dongkui, XIE Youhao, et al. Obstacle detection and tracking for intelligent vehicle based on density characteristics of point cloud using 3D lidar [J]. *Journal of Hefei University of Technology (Natural Science)*, 2019, 42(10): 1311-1317.
- [9] 徐岩, 陶慧青, 虎丽丽. 基于 Faster R-CNN 网络模型的铁路异物入侵检测算法研究 [J]. *铁道学报*, 2020, 42(5): 91-98.
- XU Yan, TAO Huiqing, HU Lili. Railway foreign body intrusion detection based on Faster R-CNN network model [J]. *Journal of the China Railway Society*, 2020, 42(5): 91-98.
- [10] 张剑, 王等准, 莫光健, 等. 基于改进 YOLOv3 的高铁异物入侵检测算法 [J]. *计算机技术与发展*, 2022, 32(2): 69-74.
- ZHANG Jian, WANG Dengzhun, MO Guangjian, et al. High-speed rail foreign body intrusion detection algorithm based on improved YOLOv3 [J]. *Computer Technology and Development*, 2022, 32(2): 69-74.
- [11] HUSSAIN R, KARBHARI Y, IJAZ M F, et al. Reverse-net: exploiting reverse attention mechanism for salient object detection [J]. *Remote Sensing*, 2021, 13 (23): 4941.
- [12] 胡行涛, 刘大明, 虞发桐. 基于 FCN 不确定性特征的铁路入侵异物检测算法 [J]. *计算机应用与软件*, 2023, 40(4): 141-146.
- HU Xingtao, LIU Daming, YU Fatong. Foreign intrusion object detection algorithm for railway intrusion based on FCN uncertainty [J]. *Computer Applications and Software*, 2023, 40(4): 141-146.
- [13] CAI Huiling, LI Fei, GAO Dianzhu, et al. Foreign objects intrusion detection using millimeter wave radar on railway crossings [C]//2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC). Piscataway, NJ, USA: IEEE, 2020: 2776-2781.
- [14] HE Deqiang, QIU Yefeng, MIAO Jian, et al. Improved mask R-CNN for obstacle detection of rail transit [J]. *Measurement*, 2022, 190: 110728.
- [15] WANG Yao, YU Zujun, ZHU Liqiang. Intrusion detection for high-speed railways based on unsupervised anomaly detection models [J]. *Applied Intelligence*, 2023, 53(7): 8453-8466.
- [16] 管岭, 贾利民, 谢征宇. 融合注意力机制的轨道入侵异物检测轻量级模型研究 [J]. *铁道学报*, 2023, 45 (5): 72-81.
- GUAN Ling, JIA Limin, XIE Zhengyu. Research on lightweight model for railway intrusion detection integrating attention mechanism [J]. *Journal of the China Railway Society*, 2023, 45(5): 72-81.
- [17] HE Deqiang, YANG Yanjie, CHEN Yanjun, et al. An integrated optimization model of metro energy consumption based on regenerative energy and passenger transfer [J]. *Applied Energy*, 2020, 264: 114770.
- [18] 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 [C]//2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2023: 7464-7475.
- [19] HUANG Peile, WANG Shenghuai, CHEN Jianyu, et al. Lightweight model for pavement defect detection based on improved YOLOv7 [J]. *Sensors*, 2023, 23 (16): 7112.
- [20] WOO S, PARK J, LEE J Y, et al. CBAM: convolutional block attention module [C]//Computer Vision-ECCV 2018. Cham, Germany: Springer International Publishing, 2018: 3-19.
- [21] 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.
- [22] ZHAO Xiaotong, LI Wei, ZHANG Yifan, et al. Aggregated residual dilation-based feature pyramid network for object detection [J]. *IEEE Access*, 2019, 7: 134014-134027.
- [23] YU Haifeng, LI Xinbin, FENG Yankai, et al. Multiple attentional path aggregation network for marine object detection [J]. *Applied Intelligence*, 2023, 53(2): 2434-2451.
- [24] SONG Zhiwei, HUANG Xinbo, JI Chao, et al. Intelligent identification method of hydrophobic grade of composite insulator based on efficient GA-YOLO former network [J]. *IEEE Transactions on Electrical and Electronic Engineering*, 2023, 18(7): 1160-1175.
- [25] GAO Jinfeng, CHEN Yu, WEI Yongming, et al. Detection of specific building in remote sensing images using a novel YOLO-S-CIOU model. case: gas station identification [J]. *Sensors*, 2021, 21(4): 1375.
- [26] ZHAO Qing, WEI Honglei, ZHAI Xianyi. Improving

tire specification character recognition in the YOLOv5 network [J]. Applied Sciences, 2023, 13(12): 7310.

[27] 张开放, 苏华友, 窦勇. 一种基于混淆矩阵的多分类任务准确率评估新方法 [J]. 计算机工程与科学, 2021, 43(11): 1910-1919.

ZHANG Kaifang, SU Huayou, DOU Yong. A new multi-classification task accuracy evaluation method based on confusion matrix [J]. Computer Engineering & Science, 2021, 43(11): 1910-1919.

[28] GE Zheng, LIU Songtao, WANG Feng, et al. YOLOX: exceeding YOLO series in 2021 [EB/OL]. (2021-08-06) [2023-11-12]. <https://arxiv.org/abs/2107.08430>.

[29] SU Xiaodong, HU Jianxing, CHEN Linzhou, et al. Research on real-time dense small target detection algorithm of UAV based on YOLOv3-SPP [J]. Journal of the Brazilian Society of Mechanical Sciences and Engineering, 2023, 45(9): 488.

[30] TIAN Zhi, SHEN Chunhua, CHEN Hao, et al. FCOS: a simple and strong anchor-free object detector [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022, 44(4): 1922-1933.

(编辑 杜秀杰)