

# 基于改进 YOLOv8 的接触网异物检测算法<sup>\*</sup>

王卓萍<sup>1</sup> 张林煊<sup>1,2</sup> 李奕超<sup>3</sup> 陈燕楠<sup>1</sup>

(1. 新疆大学电气工程学院 乌鲁木齐 830017; 2. 清华大学国家计算机集成制造系统工程技术创新中心 北京 100084;  
3. 中国铁路乌鲁木齐局集团有限公司科学技术研究所 乌鲁木齐 830063)

**摘要:** 针对当前的铁路接触网异物检测算法存在检测精度不高和模型过大的问题,本研究提出一种基于改进 YOLOv8 的接触网异物检测算法(FRDW-YOLOv8)。首先,在主干网络中引入 FasterNet 模块,从而构建出 C2f-Faster 模块,降低模型复杂度,提升模型的计算效率;其次,在特征提取阶段引入感受野坐标注意力机制(RFCA),提高模型对于接触网异物区域的关注度,为其分配更多注意力;然后,在颈部网络中采用动态上采样器(Dysample),可以保留接触网异物更多的细节信息;最后,采用 WIoU v3 损失函数,通过动态调整权重因子,提高检测模型的整体性能。实验结果表明,改进后的算法 mAP 值达到 95.1%,较 YOLOv8 模型提升了 2.8%,模型的计算量和参数量仅为 7.3 G 和 2.7 M。改进后的算法进一步提高了模型的检测精度,且使模型具有轻量化的特性,充分说明改进后的算法检测性能优于目前主流的算法,能够更好地完成铁路接触网异物检测任务。

**关键词:** 接触网异物;目标检测;注意力机制;YOLOv8;损失函数

中图分类号: U226.8; TN919.8 文献标识码: A 国家标准学科分类代码: 510.4050

## Improved YOLOv8 for foreign object detection in catenary systems

Wang Zhuoping<sup>1</sup> Zhang Linxuan<sup>1,2</sup> Li Yichao<sup>3</sup> Chen Yannan<sup>1</sup>

(1. School of Electrical Engineering, Xinjiang University, Urumqi 830017, China;  
2. National Computer Integrated Manufacturing System Engineering Technology Research Center, Tsinghua University, Beijing 100084, China;  
3. Science and Technology Research Institute of China Railway Urumqi Group Corporation, Urumqi 830063, China)

**Abstract:** Aiming at the problems of low detection accuracy and large model size existing in the current foreign object detection algorithms for railway catenaries, this study proposes a foreign object detection algorithm for catenaries (FRDW-YOLOv8) based on the improved YOLOv8. Firstly, we propose the integration of a FasterNet module into the backbone network to construct the C2f-Faster module, which effectively reduces model complexity and enhances computational efficiency. Secondly, the Receptive-Field Coordinate Attention mechanism (RFCA) is introduced in the feature extraction stage to increase the model's attention to the foreign object areas of the catenary and allocate more attention to them. Then, a dynamic upsample (Dysample) is adopted in the neck network, which can retain more detailed information of the foreign objects on the catenary. Finally, the WIoU v3 loss function is used to improve the overall performance of the detection model by dynamically adjusting the weight factors. The experimental results show that the mAP value of the improved algorithm reaches 95.1%, which is 2.8% higher than that of the YOLOv8 model, and the floating-point operations and the number of parameters of the model are only 7.3 G and 2.7 M respectively. The improved algorithm further improves the detection accuracy of the model and makes the model lightweight. It fully demonstrates that the detection performance of the improved algorithm is superior to the current mainstream algorithms and can better complete the task of detecting foreign objects on railway catenaries.

**Keywords:** foreign objects on the catenary; target detection; attention mechanism; YOLOv8; loss function

## 0 引言

随着铁路交通运输业的不断发展,铁路作为关键性重

大基础设施,在经济社会发展中的作用越来越重要。铁路接触网作为特殊的输电线路,其状态与列车的安全、可靠运行密切相关。如果接触网上的异物无法及时被清除,可能

会造成接触网线路短路、绝缘子污闪等现象<sup>[1]</sup>。因此,对铁路接触网鸟巢、漂浮物等异物进行实时准确的检测是很有必要的。

目前,针对铁路接触网的目标检测方法,国内外学者进行了广泛研究。在传统图像检测方面,段旺旺等<sup>[2]</sup>首先提取鸟巢区域的 HOG 特征,然后通过训练好的 SVM 根据关键区域的 HOG 特征进行精确提取,自动识别出鸟巢区域,然而 HOG 仅能捕捉边缘、梯度方向等底层特征,缺乏对图像全局上下文的建模能力,检测精度较低。同时 SVM 难以处理复杂非线性边界,需依赖核函数映射到高维空间,计算成本高且效果有限,对于接触网鸟巢检测的实时性较差。祝振敏等<sup>[3]</sup>首先使用 Sobel 算子检测图像的边缘,接着通过霍夫变换对图像中的硬横梁进行检测,最后对图像二值化处理,根据白色区域面积的占比,对横梁上的鸟巢进行检测,然而 Sobel 算子对于接触网图像背景噪声敏感,边缘检测结果易受到光照等影响,可能导致边缘不连续或伪边缘,而且二值化阈值需手动设定,难以适应光照变化或复杂背景的干扰,仅通过白色区域占比规则过于简单,鸟巢形状、大小的变化易导致鸟巢的误检或漏检,检测精度较低。此外,Sobel、霍夫变换和二值化处理需分步计算,检测速度慢,无法满足实时性检测。Wu 等<sup>[4]</sup>利用 HOS 和 HLS 直方图表示检测到的树枝条纹的方向和长度分布,并利用 SVM 进行建模,对接触网上的鸟巢异物进行检测,然而直方图分箱数和 SVM 核函数的选择对检测结果影响较大,需要反复的人工调参,而且,直方图仅反映局部枝条属性,无法建模鸟巢与环境的语义关联,导致检测模型存在误检。此外,SVM 推理需高维特征计算,难以并行加速,检测实时性较差。因此,利用传统的图像检测方法依赖人工设计特征,而且传统的图像检测方法对接触网图像噪声、异物尺寸变化、形变敏感,缺乏语义理解和上下文建模能力,利用传统图像检测方法对接触网异物的误检率和漏检率较高,检测精度较低。虽然利用传统检测方法的模型参数量相对较少,但是存在计算效率低等缺点,无法适用于对于接触网异物检测实时性要求较高的场景。

近年来,基于深度学习的一系列算法在接触网异物检测中得到了大量的关注<sup>[5]</sup>,比如 SSD<sup>[6]</sup>、YOLO<sup>[7]</sup>等。吕嘉宜<sup>[8]</sup>通过 K-means 算法改进 SSD 网络的先验框设置,并引入视觉显著性修改网络结构,从而强化网络对接触网异物特征的提取能力,然而 K-means 对初始聚类中心敏感,可能导致生成的锚框尺寸分布不均衡,而且引入视觉显著性检测可能降低推理速度,影响实时性。此外,异物与背景相似时,显著性模块难以捕捉,可能导致误检和漏检。顾桂梅等<sup>[9]</sup>对 SSD 算法进行改进,引入 MobileNetV3 用于特征提取,设计了特征融合模块重构 SSD 的特征层,用于接触网上的小目标缺陷识别,然而 SSD 依赖多尺度特征图预测,MobileNetV3 的深层特征分辨率过低,可能导致小目标特征丢失,影响模型的检测精度。Zhao 等<sup>[10]</sup>提出了一种上下

文引导的从粗到细的检测模型,该模型充分利用了上下文信息来检测鸟巢,并采用流水线技术和多线程并行计算技术,以最大限度地利用计算资源,提高检测速度,然而该模型中级联了多个检测器,增加了模型的参数量,易导致模型过拟合,泛化能力下降,且不便于模型的移动端部署。郭翔羽等<sup>[11]</sup>提出了一种基于 YOLOv7 改进的网络模型对接触网异物进行检测,改进后的模型在检测精度和速度上均有所提升,然而该模型的参数量较大,不利于模型的实际工程部署。

尽管目前的研究在接触网异物检测领域取得了令人瞩目的进展,但是多数研究只针对接触网上的鸟巢进行检测和识别,而无法检测接触网上的漂浮物等其他异物。此外,当前的铁路接触网异物检测模型存在参数量和计算量较大或者检测精度较低的问题,无法较好地平衡检测模型的权重大小与检测精度。为了解决这些难题,本文提出一种基于改进 YOLOv8<sup>[12]</sup> 的接触网异物检测算法 FRDW-YOLOv8。FRDW-YOLOv8 算法将主干网络 C2f 模块中的 Bottleneck 替换为 FasterNet<sup>[13]</sup>,从而构建出 C2f-Faster 模块,降低模型参数量和计算量;其次,在特征提取阶段引入感受野坐标注意力机制 (receptive-field coordinate attention, RFCA)<sup>[14]</sup>,构建出感受野坐标注意力卷积模块 (RFCACConv),提高模型对于接触网异物区域的关注度,为其分配更多注意力;然后,在颈部网络中采用动态上采样器 (dynamic upsampler, Dysample)<sup>[15]</sup>,减少上采样过程中图像细节的丢失;最后,使用 Wise-IoU v3 (WIoU v3)<sup>[16]</sup> 损失函数优化模型,改善边界框回归的训练效果,提高模型的检测性能。

## 1 FRDW-YOLOv8 网络架构

在 YOLOv8 的众多版本中, YOLOv8n 相较于 YOLOv8s、YOLOv8m、YOLOv8l 以及 YOLOv8x 来说, YOLOv8n 所需的计算资源更少。其次, YOLOv8n 的结构相对简单,不仅训练效率更高,而且为后续改进提供了良好的起点。因此,综合考虑计算资源、训练效率、后续扩展等因素,本研究选用 YOLOv8n 作为基准模型。

FRDW-YOLOv8 网络架构如图 1 所示。整体结构由输入层 (Input)、主干网络 (Backbone)、颈部网络 (Neck)、检测头 (Head) 所构成。输入层负责接收待检测的接触网异物图像,并对其进行预处理;主干网络负责对图像的特征进行提取;颈部网络对不同尺寸的特征图进行融合;检测头将融合后的特征进行分类和回归,输出最终的检测结果。

## 2 YOLOv8 模型改进策略

### 2.1 构建 C2f-Faster 模块

由于 YOLOv8 中的 Bottleneck 依赖多次卷积和跨层连接,使得 YOLOv8 模型的参数量和计算量较大。为了便于网络模型的实际工程部署,使得模型更加轻量化,本研究

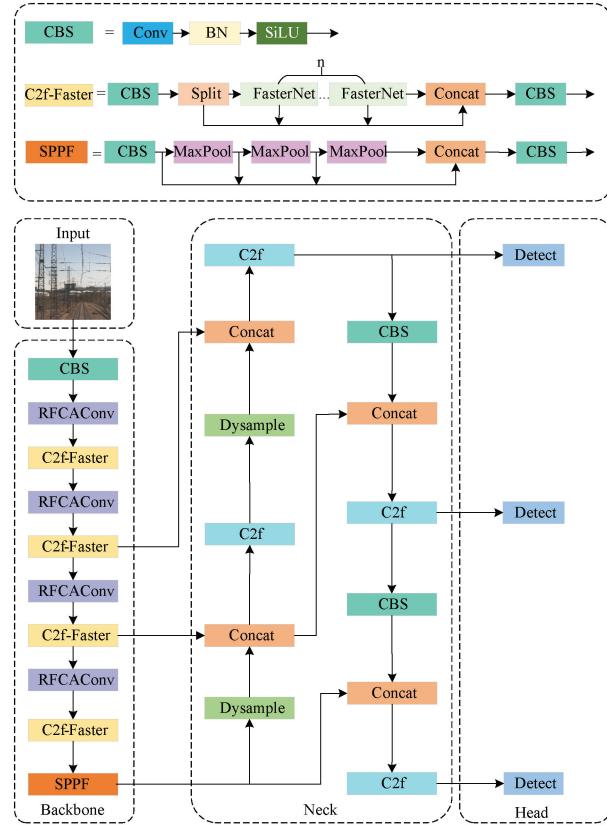


图 1 FRDW-YOLOv8 模型结构图

Fig. 1 Structure diagram of the FRDW-YOLOv8 model

将 YOLOv8 主干网络中 C2f 模块的 Bottleneck 替换为 FasterNet, 从而构建出 C2f-Faster 模块。FasterNet 通过设计部分卷积 (partial convolution, PConv)<sup>[17]</sup>, 减少冗余计算和内存访问, 通过优化通道分配和卷积操作顺序, 在保证特征表达能力的同时, 显著降低了模型的参数量以及计算量。

PConv 的工作原理如图 2(a) 所示, 假设输入和输出特征图的通道数相同,  $h$  和  $w$  分别表示特征图的高度和宽度,  $k$  表示内核大小,  $cp$  表示 PConv 的通道数。PConv 的 FLOPs 为  $h \times w \times k^2 \times cp^2$ , 当  $cp = c/4$  时, PConv 的 FLOPs 仅为标准卷积的  $1/16$ 。此外, PConv 的内存访问量为  $h \times w \times 2cp + k^2 \times cp^2 \approx h \times w \times 2cp$ , 当  $cp = c/4$  时, PConv 的内存访问量仅为标准卷积的  $1/4$ 。

图 2(b) 展示了 FasterNet 结构图, FasterNet 由一个 PConv 层, 以及两个  $1 \times 1$  卷积层所组成, 其中的残差连接用于重复使用输入特征。除此之外, 在中间的  $1 \times 1$  卷积层之后应用归一化层和激活层, 以保持特征多样性并实现更低的延迟。

## 2.2 引入 RFCA 注意力机制

由于 YOLOv8 主干网络的标准卷积操作每个滑块内相同位置的特征参数是固定的, 存在卷积核参数共享的问题。当接触网背景较为复杂时, YOLOv8 主干网络标准卷

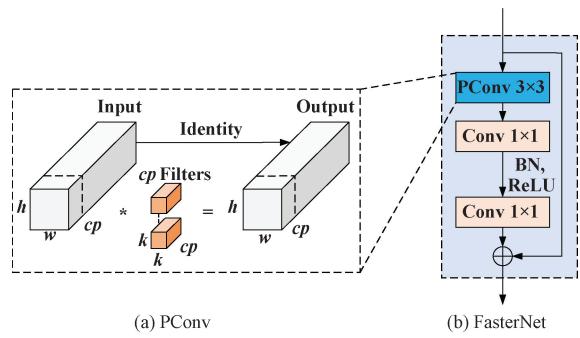


图 2 FasterNet 与 PConv 结构图

Fig. 2 Structure diagrams of FasterNet and PConv

积的卷积核在不同的感受野中使用相同的参数提取特征信息, 无法考虑到不同位置的特征差异, 使得检测模型出现漏检和误检的现象, 降低了对于接触网异物的检测精度。为了更加准确的对不同位置的特征进行有效提取, 在本研究中引入感受野坐标注意力 RFCA, 构建出感受野坐标注意力卷积 RFCACConv 模块, 用于替换主干网络中的标准卷积 CBS 模块, 彻底消除了卷积核参数共享的问题。

RFCAConv 的网络结构如图 3 所示。首先采用组卷积快速提取输入的感受野空间特征, 原始的特征被映射为新特征  $X$ , 调整后的特征图的宽和高均变为原来的  $K$  倍, 通过构建感受野空间特征, 避免了卷积核参数共享的问题, 从而能够捕获每个空间特征信息并减少计算冗余。然后利用全局平均池化压缩空间信息, 得到一对聚合特征图。接着将它们连接起来, 经过  $1 \times 1$  的卷积, 得到一个中间特征图。再沿空间维度将中间特征图分成两个独立的张量, 利用两个  $1 \times 1$  卷积操作, 得到与  $X$  具有相同通道数的注意力图。这样, 位置信息就能在生成的注意力图中得到保留, 然后通

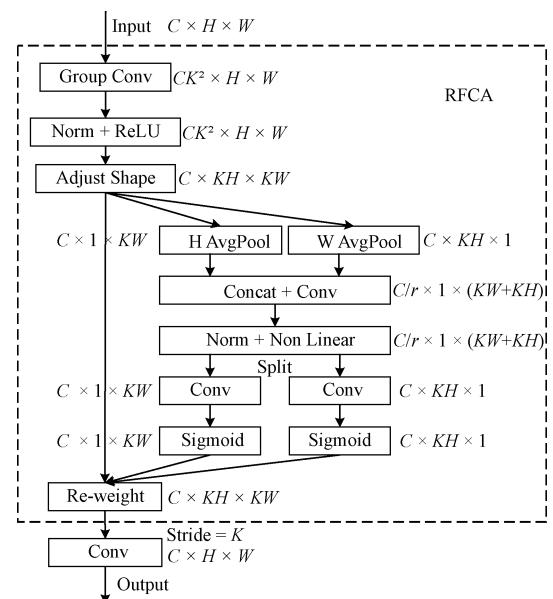


图 3 RFCAConv 模块结构图

Fig. 3 Structure diagram of the RFCAConv module

过乘法操作将这两个注意力图应用到感受野空间特征中,以强调感兴趣的表征。最终通过步长为  $K$  的  $K \times K$  卷积操作来提取特征信息。

RFCA 注意力模块将位置信息嵌入通道注意力中,同时考虑了通道与空间位置信息,使得网络模型在特征提取阶段更加关注图像中的异物区域,并为其分配更多注意力。

### 2.3 采用 Dysample 动态上采样

在 YOLOv8 检测模型中,使用最近邻插值法作为默认上采样方法,该方法对所有图像以相同的方式进行上采样,而不考虑图像的特征差异,一旦确定了上采样参数,就难以在不同的图像或任务中进行灵活调整。由于图像中的接触网异物尺寸和形态不一,可能会出现像素失真等问题,导致在上采样过程中图像细节的丢失,使得网络模型检测精度下降。为了解决这个问题,本文引入一种高度轻量级且有效的动态上采样器 Dysample,其核心在于从点采样角度重新构建上采样过程,通过考虑图像的局部结构和特征,使得网络模型在上采样过程中保留更多的细节信息,提高模型的检测精度。

Dysample 的网络结构图如图 4 所示。将上采样过程看作点采样,假设给定特征图  $X$  和采样集  $S$ ,输入特征图  $X$  经双线性插值成为连续特征图,然后生成内容感知的采样点对该连续特征图重新采样为  $X'$ ,即:

$$X' = \text{grid\_sample}(X, S) \quad (1)$$

点采样生成器如图 5 和 6 所示。以图 5 中的静态因子采样方法为例,若给定特征图  $X$  和静态范围因子为 0.25,特征图首先经过线性层生成偏移量  $O$ ,经像素洗牌重塑后,采样集  $S$  为偏移量  $O$  与原始采样网格  $G$  之和,即:

$$O = 0.25 \text{linear}(X) \quad (2)$$

$$S = G + O \quad (3)$$

最后,通过式(1)中的网格采样函数,可以生成上采样特征图  $X'$ ,从而实现动态上采样。通过这些步骤,Dysample 绕过了基于内核的模式,从点采样的角度制定上采样,并且可以通过 PyTorch 的标准内置函数轻松实现,能有效提高上采样的精度和性能。

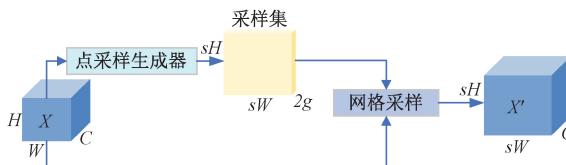


图 4 Dysample 结构图

Fig. 4 Structure diagram of the Dysample

### 2.4 优化损失函数

YOLOv8 模型中的损失函数采用了 CIoU,其损失值表示如下:

$$L_{\text{CIoU}} = 1 - R_{\text{IoU}} + \frac{(x - x_t)^2 + (y - y_t)^2}{(W^2 + H^2)} + \alpha v \quad (4)$$

其中,  $R_{\text{IoU}}$  表示交并比;  $W$  和  $H$  表示锚框与真实框所

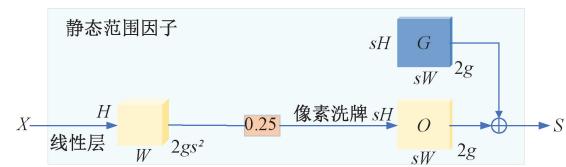


图 5 静态范围因子点采样生成器

Fig. 5 Static range factor point sampler

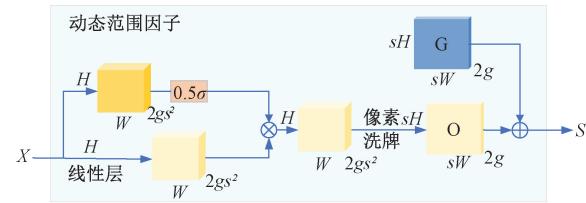


图 6 动态范围因子点采样生成器

Fig. 6 Dynamic range factor point sampler

围成的最小外接框的坐标值;  $x_t$  和  $y_t$  为真实框中心点的坐标值;  $v$  是修正因子,用来描述长宽比的一致性;  $\alpha$  为平衡参数。 $\alpha$  和  $v$  的具体计算方式分别为:

$$\alpha = \frac{v}{(1 - R_{\text{IoU}}) + v} \quad (5)$$

$$v = \frac{4}{\pi^2} (\arctan \frac{w_t}{h_t} - \arctan \frac{w}{h})^2 \quad (6)$$

其中,  $w$  和  $h$  分别表示预测框的宽和高;  $w_t$  和  $h_t$  分别为真实框的宽和高。

目前有关接触网异物的公开数据集较为稀少,本实验采用自建数据集,在训练数据中难免存在一些低质量样本。若在接触网异物检测过程中,采用 YOLOv8 默认的损失函数 CIoU,则边界框距离和长宽比等几何因素会加重对低质量样本的惩罚,从而降低模型的检测精度。因此,在本研究中采用 WIoU v3 损失函数,通过提出动态非单调的聚焦机制,使用“离群度”替代 IoU 对锚框进行质量评估,并提供了明智的梯度增益分配策略。损失值计算公式如下:

$$L_{\text{WIoUv3}} = r \cdot L_{\text{WIoUv1}} \quad (7)$$

$$L_{\text{WIoUv1}} = R_{\text{WIoU}} \cdot L_{\text{IoU}} \quad (8)$$

$$R_{\text{WIoU}} = \exp \left( \frac{(x - x_t)^2 + (y - y_t)^2}{(W^2 + H^2)^*} \right) \quad (9)$$

其中,  $R_{\text{WIoU}} \in [1, e]$ , 其将显著放大普通质量锚框的  $L_{\text{IoU}}$ ;  $r$  为非单调聚焦系数,其表示为:

$$r = \frac{\beta}{\delta \alpha^{\beta - \delta}} \quad (10)$$

其中,  $\delta$  和  $\alpha$  为超参数,当  $\beta = \delta$  时,  $r = 1$ 。这里的  $\beta$  为离群度,表示为:

$$\beta = \frac{L_{\text{IoU}}^*}{\overline{L}_{\text{IoU}}} \in [0, +\infty) \quad (11)$$

其中,  $\overline{L}_{\text{IoU}}$  为  $L_{\text{IoU}}$  的滑动平均值;  $L_{\text{IoU}}^*$  为梯度增益值。由于  $\overline{L}_{\text{IoU}}$  是动态的,锚框的质量划分标准也是动态的。通

过为不同离群度的锚框动态分配梯度增益,不仅降低了高质量锚框的竞争力,而且减少了低质量锚框产生的有害梯度<sup>[18]</sup>,最终使得 WIoU 可以聚焦于普通质量的锚框,从而提高网络模型的检测精度。

### 3 数据集扩充与数据增强

#### 3.1 数据集扩充

本文数据集为自建数据集,从某铁路局获取到含有接触网异物的图像 658 张,原始图像数据来源于列车的便携式接触网安全巡检装置(C2)。由于原始数据较少,直接用于训练后的网络模型可能出现过拟合现象。为了构建的数据集包含更多的语义信息,本文对原始图像经过平移、旋转、镜像、改变亮度以及添加噪声等方法进行图像扩充,扩充后的数据集共计 2 890 张。部分图像的扩充结果如图 7 所示。然后,使用 LabelImg 工具对接触网图像中的漂浮物(floater)和鸟巢(nest)异物进行标注,并按照 8:1:1 的比例划分训练集、验证集以及测试集。

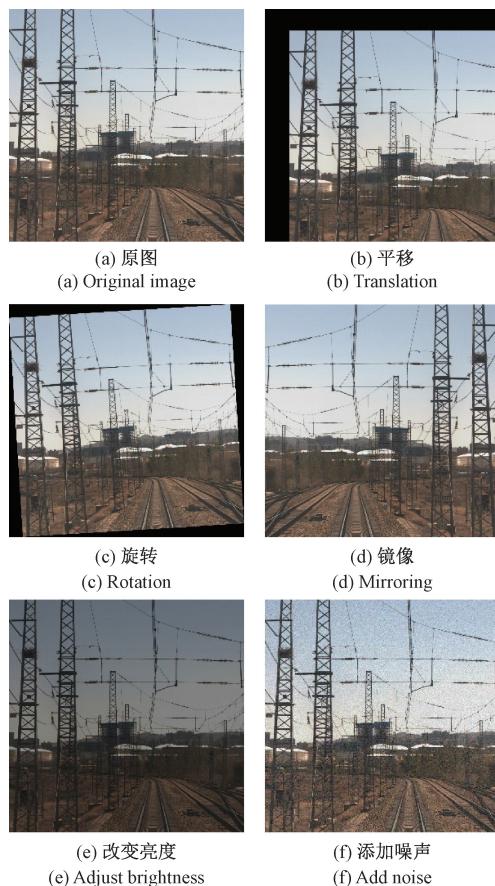


图 7 图像扩充样例图

Fig. 7 Augmented image sample diagram

#### 3.2 数据增强

为了增强数据的多样性以及网络模型的鲁棒性,提高网络模型对于小尺寸异物的检测能力,在 FRDW-YOLOv8

的 Input 层中引入 Mosaic 增强技术对图像进行预处理。Mosaic 数据增强效果如图 8 所示。



图 8 Mosaic 数据增强效果图

Fig. 8 Mosaic data augmentation renderings

通过随机选取 4 张图像进行拼接,使得组合后的图像样本更加多样化。拼接后的新图像中具有更加复杂丰富的背景,使网络模型检测过程中具有更好的抗干扰能力。此外,Mosaic 增强后的图像中拥有更多不同尺寸的目标,通过提高模型对小尺寸异物的检测能力。

### 4 实验结果与分析

#### 4.1 实验训练参数及环境配置

由于本文所用数据样本较少,且为自建数据集,因此为了使得模型训练时拥有更好的初始化性能和更快的收敛速度,在本文中采用官方提供的、基于 COCO 数据集预训练的 YOLOv8n.pt 权重文件作为模型初始权重。网络模型训练过程中,实验的训练参数以及环境配置分别如表 1 和 2 所示。

表 1 实验训练参数

Table 1 Experimental training parameters

参数	参数值
输入图像尺寸	640×640
训练周期	200
批量大小	16
初始学习率	0.01
权重衰减系数	0.000 5
动量因子	0.937

表 2 实验环境配置

Table 2 Experimental environment configuration

软件环境	硬件环境
操作系统:Ubuntu 20.04	CPU:AMD EPYC 9654 96-Core
编程语言版本:Python 3.8	GPU:RTX 4090(24 GB)×1
深度学习框架:PyTorch 1.11.0	
加速环境:CUDA 11.3	

#### 4.2 评价指标

本研究对于接触网异物图像识别采用了以下评价指标:

$$P = \frac{TP}{TP + FP} \times 100\% \quad (12)$$

$$R = \frac{TP}{TP + FN} \times 100\% \quad (13)$$

$$AP = \int_0^1 P(R) dR \times 100\% \quad (14)$$

$$mAP = \frac{1}{n} \sum_{k=1}^n AP_k \times 100\% \quad (15)$$

其中,  $TP$  表示预测正确的正样本个数,  $FN$  表示预测错误的负样本个数,  $FP$  表示预测错误的正样本个数。

$P$  称为精确率, 表示预测为正样本中实际为正样本所占的比例。 $R$  称为召回率, 表示所有正样本中, 被模型正确预测为正样本所占的比例。 $AP$  称为平均精度, 它反映了模型的精确率和召回率, 其值等于以精确率  $P$  和召回率  $R$  围成的  $P-R$  曲线与坐标轴围成的面积。在本文中, 通过平均精度均值  $mAP$  来衡量检测模型的总体性能,  $mAP$  为鸟巢和漂浮物的平均精度的平均值。由于  $AP$  值的计算过程中, 需要考虑交并比(IoU)阈值, 若 IoU 阈值较小, 则检测精确率  $P$  较小, 模型存在误检; 若 IoU 阈值较大, 则检测的召回率  $R$  较小, 模型存在漏检。为了能够反映模型在宽松定位要求下的综合性能, 在本研究中模型的 IoU 阈值取值为 0.5。

除了与检测精度相关的评价指标外, 本研究还包括参数量(Params)、浮点运算量(FLOPs)以及帧率(FPS)。浮点运算量表示模型在进行推理或训练时所需的浮点运算次数, 该指标反映了模型的复杂度, 而参数量表示模型中可学习的参数数量, 参数量大的模型需要更多的存储空间来存储模型参数, 并且在训练和推理过程中需要消耗更多的计算资源和时间。帧率表示模型每秒处理的图像帧数, 模型的帧率越高, 意味着模型的检测速度越快。

#### 4.3 注意力机制对比实验

为验证本研究中加入的 RFCA 注意力机制的优越性, 在 YOLOv8 模型中的相同位置分别添加几种不同的注意力机制进行对比实验, 其结果如表 3 所示。

表 3 不同注意力机制对比结果

Table 3 Comparison results of different attention mechanisms

网络模型	P/%	R/%	mAP/%
YOLOv8	93.2	90.5	92.3
YOLOv8+SimAM	93.4	88.6	91.6
YOLOv8+CBAM	92.8	91.8	92.4
YOLOv8+CA	91.2	92.3	92.7
YOLOv8+RFCA	<b>92.7</b>	<b>92.4</b>	<b>93.8</b>

从表 3 中可以看出, 相较于注意力机制 SimAM<sup>[19]</sup>、CBAM<sup>[20]</sup>、CA<sup>[21]</sup>来说, RFCA 注意力机制对于 YOLOv8 模型的检测精度提升更加明显, 模型的 mAP 值达到了 93.8%。SimAM 注意力机制在计算整张特征图的像素差平均值时加权可能会忽略小目标的重要性, 其感受野基于整个特征图计算, 可能无法很好地适应小目标尺度, 在接触网尺寸较小的异物检测任务中可能限制性能。CBAM 注意力机制利用卷积计算空间注意力, 然而卷积只能捕捉局

部关系, 无法捕捉视觉任务中至关重要的长距离依赖关系。CA 注意力机制只关注空间特征, 不能完全解决卷积核参数共享的问题。相比之下, RFCA 注意力机制能够使网络模型关注感受野空间特征, 提高网络模型对于异物区域的关注度, 为其分配更多注意力, 抑制复杂背景对于特征提取的干扰, 对于检测模型的性能提升更加明显。

为了对比添加 RFCA 前后的 YOLOv8 模型对不同区域的关注程度, 采用 Grad-CAM<sup>[22]</sup> 热力图进行可视化分析, 对比结果如图 9 所示。从图 9(a)可以看出, 在使用 YOLOv8 模型检测接触网异物时, 背景区域的噪声明显对于异物检测产生了干扰; 从图 9(b)可以看出, YOLOv8 模型添加 RFCA 后, 背景区域的噪声明显降低, 网络模型提高了对于接触网异物区域的关注度, 表明添加 RFCA 注意力机制的 YOLOv8 模型具有更好的检测效果和抗干扰能力。

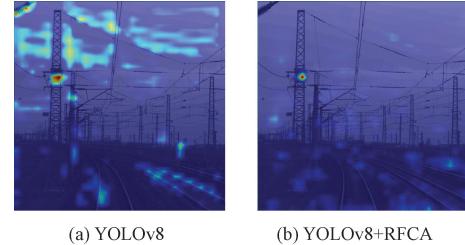


图 9 YOLOv8 改进前后热力图对比

Fig. 9 Comparison of heatmaps of YOLOv8 before and after improvement

#### 4.4 损失函数对比实验

为了探究不同损失函数对检测算法性能的影响, 在改进 YOLOv8 的模型中分别采用 CIoU、DIoU、SIoU 以及不同版本的 WIoU 损失函数进行对比实验, 其结果如表 4 所示。

表 4 不同损失函数对比结果

Table 4 Comparison results of different loss functions

损失函数	P	R	mAP	%
CIoU	92.9	92.1	93.5	
DIoU	92.7	90.4	92.7	
SIoU	93.7	91.5	93.8	
WIoU v1	93.2	92.2	94.3	
WIoU v2	95.1	89.4	93.4	
WIoU v3	<b>93.5</b>	<b>92.8</b>	<b>95.1</b>	

由表 4 的实验结果可知, 当 FRDW-YOLOv8 模型采用 WIoU v3 损失函数时, 网络模型的 mAP 最高。当选用 CIoU 损失函数时, 假设预测框与真实框的长宽比相同, 式(4)中的  $\alpha$  和  $v$  恒为 0,  $L_{CIoU}$  无法被稳定表达。此外, 距离和长宽比等几何因素会加重对低质量样本的惩罚, 从而

降低模型的泛化能力。当选用 DIoU 损失函数时,由于仅考虑了重叠面积和中心点距离,而没有考虑到预测框与真实框的长宽比。当选用 SIoU 损失函数时,需要计算多个损失函数并进行加权平均,计算过程相对复杂,可解释性较差。WIoU v3 相较于 WIoU v1、WIoU v2 以及其他损失函数来说,由于 WIoU v3 损失函数中采用了动态非单调聚

焦机制,为不同离群度的锚框动态分配梯度增益,可以更好地提高模型的整体性能。

#### 4.5 消融实验

为了探究不同损失函数为验证各个改进模块对接触网异物检测 FRDW-YOLOv8 模型的影响,本研究在自建数据集上进行消融实验,其结果如表 5 所示。

表 5 消融实验对比结果

Table 5 Comparison results of ablation experiment

模型	C2f-Faster	RFCA	Dysample	WIoU v3	mAP/%	Params/M	FLOPs/G	FPS
YOLOv8					92.3	3.0	8.1	97.6
YOLOv8-A	√				90.7	2.6	7.0	115.5
YOLOv8-B	√	√			92.8	2.7	7.3	106.4
YOLOv8-C	√	√	√		93.5	2.7	7.3	101.2
FRDW-YOLOv8	√	√	√	√	95.1	2.7	7.3	99.3

注:“√”表示添加该模块。

由表 5 可知,YOLOv8-A 模型相较于 YOLOv8 模型来说,通过构建 C2f-Faster 模块替代原模型中的 C2f 模块,使得模型的参数量和浮点运算量分别减少了 13.3% 以及 13.6%,而 mAP 只下降了 1.6%,且 FPS 明显高于原 YOLOv8 模型,说明此改进在不明显丢失检测精度的同时极大地减少了网络模型的参数量以及计算量,并提升了网络模型的推理速度。在 YOLOv8-A 模型的基础上,YOLOv8-B 模型中引入了 RFCA 注意力机制,mAP 值达到了 92.8%,相比于 YOLOv8-A 模型提高了 2.1%,表明在主干网络中添加 RFCA 注意力机制后,网络模型更加关注接触网上的异物区域。YOLOv8-C 模型相较于 YOLOv8-B 模型,采用了 Dysample 进行上采样,检测结果的 mAP 提高了 0.7%,证明 Dysample 能够在上采样过程中保留更多的细节信息,提高模型的检测精度。对 YOLOv8-C 模型的损失函数进行改进,将 CIoU 替换为 WIoU v3,最终得到本研究改进后的 FRDW-YOLOv8 模型,该模型相较于 YOLOv8-C 模型的 mAP 值提高了 1.6%,说明 WIoU v3 通过合理分配增益,有效提高了边框回归精度。最终,FRDW-YOLOv8 模型的 mAP 达到了 95.1%,mAP 相较于 YOLOv8 模型提升了 2.8%,参数量和计算量较 YOLOv8 模型分别减少 10.0% 和 9.9%,而且 FPS 提高了 1.7%。综合结果表明,本文的改进策略不仅使得网络模型更加轻量化,而且网络模型的检测精度和检测速度也得到一定程度的提升,能够更好地对接触网异物进行检测。

YOLOv8 模型与 FRDW-YOLOv8 模型训练过程中的 mAP 变化曲线如图 10 所示,从图像中可以看出,FRDW-YOLOv8 的 mAP 变化曲线整体更加平滑,而且其 mAP 变化曲线趋于稳定后,mAP 值也更大。

#### 4.6 不同检测模型的性能对比

为了进一步验证 FRDW-YOLOv8 模型的优越性,本

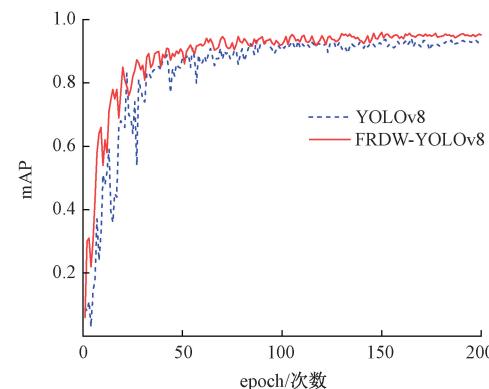


图 10 mAP 变化曲线

Fig. 10 mAP variation curve

研究选取 Faster-RCNN<sup>[23]</sup>、YOLOv5s、YOLOv7-tiny、YOLOv8n、RT-DETR<sup>[24]</sup>、YOLOv10n<sup>[25]</sup>、YOLOv11n<sup>[26]</sup>以及 YOLOv12n<sup>[27]</sup>与本文提出的 FRDW-YOLOv8 检测模型进行对比实验,实验结果如表 6 所示。

表 6 不同检测模型对比结果

Table 6 Comparison results of different detection models

模型	mAP/%	Params/M	FPS
Faster-RCNN	78.6	136.7	28.9
YOLOv5s	90.8	7.0	83.5
YOLOv7-tiny	90.1	6.2	90.8
YOLOv8n	92.3	3.0	97.6
RT-DETR	91.8	20.0	68.3
YOLOv10n	91.5	2.7	87.1
YOLOv11n	92.6	2.6	108.4
YOLOv12n	93.2	2.6	103.8
FRDW-YOLOv8	95.1	2.7	99.3

从表 6 可以看出, FRDW-YOLOv8 模型相较于 Faster-RCNN、YOLOv5s、YOLOv7-tiny、YOLOv8n 以及 RT-DETR 模型来说, FRDW-YOLOv8 模型的参数量相对较少, 检测精度更高, 而且检测速度也更快。FRDW-YOLOv8 与 YOLOv10n 网络模型的参数量相差不大, 但是该模型的检测精度和检测速度都优于 YOLOv10。FRDW-YOLOv8 与 YOLOv11n 和 YOLOv12n 网络模型相比, 虽然该模型的检测速度稍不及 YOLOv11n 和 YOLOv12n 网络模型, 但是 FRDW-YOLOv8 模型的检测精度明显更高, 且模型的参数量与 YOLOv11n 以及 YOLOv12n 模型比较接近, 说明 FRDW-YOLOv8 模型能够较好地进行边缘化部署。综上结果表明, FRDW-YOLOv8 模型能够更好地兼顾检测精度与模型的权重大小, 更加适合于接触网异物的检测。

#### 4.7 实验结果可视化

为了更加直观地展示 FRDW-YOLOv8 模型的检测效果, 将 YOLOv8 模型与 FRDW-YOLOv8 模型对接触网异物的检测结果进行对比, 结果如图 11 所示。



图 11 不同模型检测结果对比

Fig. 11 Comparison of detection results from different models

在图 11(a)中, YOLOv8 模型将接触网支柱顶端误检为了接触网异物;当图像中异物尺寸较小时, YOLOv8 网络模型存在漏检现象。而 FRDW-YOLOv8 模型检测结果如图 11(b)所示, FRDW-YOLOv8 模型能够有效降低误检率和漏检率, 对接触网上的异物进行准确识别。此外, 从图 11(a)和(b)中可以看出, 当接触网异物特征较为明显

时, YOLOv8 模型和 FRDW-YOLOv8 模型均能对接触网异物进行有效检测, 但 FRDW-YOLOv8 模型的检测置信度明显更高。

## 5 结 论

本文提出了一种基于 FRDW-YOLOv8 的接触网异物检测算法。创新点包含以下 4 个方面: 在原 YOLOv8 模型的主干部分, 本文引入了 FasterNet, 通过构建的 C2f-Faster 模块替换 C2f, 明显降低了网络模型的复杂度, 并提升了网络模型的推理速度; 其次, 在特征提取阶段引入 RFCA 注意力模块, 提高模型对于接触网异物区域的关注度, 为其分配更多注意力, 增强了主干网络的特征提取能力; 然后, 利用 Dysample 进行上采样, 可以保留接触网异物更多的细节信息, 提高了网络模型对于较小尺寸异物的检测能力; 最后, 采用 WIoU v3 损失函数优化模型, 通过为锚框动态分配梯度增益, 提高检测模型的整体性能。实验结果表明, FRDW-YOLOv8 模型在自建数据集上对接触网异物检测的 mAP 值达到 95.1%, 相较于原 YOLOv8 模型提升了 2.8%。此外, 参数量和计算量较 YOLOv8 模型分别减少 10.0% 和 9.9%, 且 FPS 提升了 1.7%。因此, YOLO-FRDW 模型不仅拥有较好的检测速度和更高的检测精度, 而且该模型具有轻量化特性, 有利于网络模型的实际部署。在未来的科研工作中, 将进一步扩充原始数据集, 收集更多类别的接触网异物用于网络模型的训练, 旨在增强模型的鲁棒性, 持续优化模型的检测能力。

## 参考文献

- [1] 高天姿. 基于深度卷积神经网络的接触网异物与缺陷检测算法研究[D]. 南昌: 华东交通大学, 2022.
- [2] GAO T Z. Research on foreign object and defect detection algorithm of overhead catenary system based on deep convolution neural network[D]. Nanchang: East China Jiaotong University, 2022.
- [3] 段旺旺, 唐鹏, 金伟东, 等. 基于关键区域 HOG 特征的铁路接触网鸟巢检测[J]. 中国铁路, 2015(8): 73-77.
- [4] DUAN W W, TANG P, JIN W D, et al. Bird nest detection in railway contact networks based on HOG features in critical areas[J]. China Railway, 2015(8): 73-77.
- [5] 祝振敏, 谢亮凯. 基于相对位置不变性的接触网鸟巢识别检测[J]. 铁道科学与工程学报, 2018, 15(4): 1043-1049.
- [6] ZHU ZH M, XIE L K. Detection of birds' nest in catenary based on relative position invariance [J]. Journal of Railway Science and Engineering, 2018, 15(4): 1043-1049.
- [7] WU X, YUAN P, PENG Q, et al. Detection of bird

- nests in overhead catenary system images for high-speed rail [J]. Pattern Recognition, 2016, 51: 242-254.
- [5] 张浩然, 吴松荣, 周懿, 等. 计及少样本的 YOLOv5s 轨枕掉块小目标缺陷检测方法研究[J]. 铁道标准设计, 2024, 68(5):52-59.
- ZHANG H R, WU S R, ZHOU Y, et al. Research on YOLOv5s sleeper falling block small target defect detection method considering few samples[J]. Railway Standard Design, 2024, 68(5):52-59.
- [6] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot multibox detector [C]. 2016 European Conference on Computer Vision (ECCV). Amsterdam, The Netherlands: ECVA, 2016:21-37.
- [7] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection [C]. 2016 IEEE Conference on Computer Vision and Pattern Recognition, 2016:779-788.
- [8] 吕嘉宜. 基于视觉的高速铁路接触网异物检测[D]. 杭州: 浙江大学, 2021.
- LYU J Y. Vision-based foreign objects detection for high speed railway catenary[D]. Hangzhou: Zhejiang University, 2021.
- [9] 顾桂梅, 王小亮. 深度学习的接触网小目标缺陷识别研究[J]. 电子测量与仪器学报, 2024, 38(4): 151-160.
- GU G M, WANG X L. Identification of the catenary small target defects in deep learning[J]. Journal of Electronic Measurement and Instrumentation, 2024, 38(4):151-160.
- [10] ZHAO H W, WU S Q, TIAN Z, et al. Context-guided coarse-to-fine detection model for bird nest detection on high-speed railway catenary [J]. Multimedia Systems, 2023, 29:2729-2746.
- [11] 郭翔羽, 石天怡, 陈燕楠, 等. 基于 YOLO-CDBW 模型的列车接触网异物检测研究[J]. 广西师范大学学报(自然科学版), 2025, 43(2): 56-69.
- GUO X Y, SHI T Y, CHEN Y N, et al. Research on foreign object detection in railway overhead contact system based on YOLO-CDBW model[J]. Journal of Guangxi Normal University(Natural Science Edition), 2025, 43(2): 56-69.
- [12] 冷睿轩. 基于 YOLOv8 的输电线路异物识别算法应用[D]. 哈尔滨: 东北农业大学, 2023.
- LENG R X. Application of foreign objects identification of transmission lines based on YOLOv8 algorithm [D]. Harbin: Northeast Agricultural University, 2023.
- [13] 张震, 周俊, 江自真, 等. 基于改进 YOLOv7 轻量化模型的自然果园环境下苹果识别方法[J]. 农业机械学报, 2024, 55(3):231-242, 262.
- ZHANG ZH, ZHOU J, JIANG Z ZH, et al. Lightweight apple recognition method in natural orchard environment based on improved YOLOv7 model[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(3):231-242, 262.
- [14] ZHANG X, LIU CH, SONG T T, et al. RFACConv: Innovating spatial attention and standard convolutional operation [J]. ArXiv preprint arXiv: 2304.03198, 2023.
- [15] TAN M X, PANG R M, LE Q V. Efficientdet: Scalable and efficient object detection[C]. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. IEEE, 2020:10778-10787.
- [16] TONG Z J, CHEN Y H, XU Z W, et al. Wise-IoU: Bounding box regression loss with dynamic focusing mechanism [J]. ArXiv preprint arXiv: 2301.10051, 2023.
- [17] 王婷婷, 王洪涛, 黄志贤, 等. 基于改进 YOLOv7 的露头区岩石裂缝检测识别[J]. 地质科技通报, 2025, 41(1): 1-14.
- WANG T T, WANG H T, HUANG ZH X, et al. Rock fracture detection and identification in outcrop area based on improved YOLOv7 [J]. Bulletin of Geological Science and Technology, 2025, 41 (1): 1-14.
- [18] 杨露露, 马萍, 王聪, 等. 结合特征重用与特征重建的 YOLO 绝缘子检测方法[J]. 计算机工程, 2024, 50(7): 303-313.
- YANG L L, MA P, WANG C, et al. YOLO insulator detection method combining feature reuse and feature rebuild [J]. Computer Engineering, 2024, 50(7): 303-313.
- [19] YANG L X, ZHANG R Y, LI L D, et al. SimAM: A simple, parameter-free attention module for convolutional neural networks[C]. 38th International Conference on Machine Learning. Virtual Event. IMLS, 2021:11863-11874.
- [20] WOO S, PARK J, LEE J Y, et al. CBAM: convolutional block attention module [C]. 15th European Conference on Computer Vision (ECCV), Munich, Germany: ECVA, 2018:3-19.
- [21] HOU Q B, ZHOU D Q, FENG J SH. Coordinate attention for efficient mobile network design [C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021: 13713-13722.

- [22] SELVARAJU R R, COGSWELL M, DAS A, et al. Grad-CAM: Visual explanations from deep networks via gradient-based localization[C]. IEEE International Conference on Computer Vision, 2017: 618-626.
- [23] REN SH 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, 2016, 39(6):1137-1149.
- [24] ZHAO Y, LYU W Y, XU SH L, DETRs beat YOLOs on real-time object detection [J]. ArXiv preprint arXiv:2304.08069, 2023.
- [25] WANG AO, CHEN H, LIU L H, et al. YOLOv10: Real-time end-to-end object detection [J]. ArXiv preprint arXiv:2405.14458, 2024.
- [26] KHANAM R, HUSSAIN M. YOLOv11: An overview of the key architectural enhancements[J]. ArXiv preprint arXiv:2410.17725, 2024.
- [27] TIAN Y J, YE Q X, DOERMANN D. YOLOv12: Attention-centric real-time object detectors[J]. ArXiv preprint arXiv:2502.12524, 2025.

### 作者简介

王卓萍,硕士研究生,主要研究方向为目標检测、深度学习。

E-mail: 1226954640@qq.com

张林鎧(通信作者),教授,博士生导师,主要研究方向为故障预测与设备健康管理。

E-mail: lxzhang@tsinghua.edu.cn