

# 基于改进 YOLO 算法地铁异物检测<sup>\*</sup>

何赛赛<sup>1,2</sup> 黄民<sup>1,2</sup> 王文胜<sup>1,2</sup>

(1. 北京信息科技大学机电学院 北京 100192; 2. 北京信息科技大学现代测控技术教育部重点实验室 北京 100192)

**摘要:** 对于地铁站台和列车间缝隙之间异物入侵造成的安全防护,以及现有深度学习算法低光照下检测精度低,空间位置精度和小目标检测准确性不足等问题,提出一种改进 YOLOv10 地铁缝隙异物检测的深度学习方法 SSS-YOLO,用于地铁缝隙中异物的监测,该方法首先通过 SCINet 构建了一个权重共享的光照学习过程,用于改善地铁缝隙暗环境下的图像质量问题,考虑到不同尺度部分特征权重,SSS-YOLO 模型中引入了无参注意力机制,同时强化了空间位置信息,减小了信息损失量,最后使用 Shape-IOU 损失函数,增强了地铁缝隙中小目标检测和回归预测框的准确性,提升了对缝隙中小目标的检测精度。实验结果表明,该方法准确率达到了 90.90%,平均检测精度提高了 3.62%。

**关键词:** 深度学习;异物检测;YOLOv10;低照度网络;注意力机制

**中图分类号:** TN29    **文献标识码:** A    **国家标准学科分类代码:** 520.6040

## Subway foreign body detection based on improved YOLO algorithm

He Saisai<sup>1,2</sup> Huang Min<sup>1,2</sup> Wang Wensheng<sup>1,2</sup>

(1. Faculty of Mechanical and Electrical Engineering, Beijing Information Science Technology University, Beijing 100192, China;

2. Key Laboratory of Modern Measurement and Control Technology of the Ministry of Education,

Beijing Information Science Technology University, Beijing 100192, China)

**Abstract:** In order to solve the problems of low detection accuracy under low light, insufficient spatial position accuracy and small target detection accuracy of existing deep learning algorithms, a deep learning method SSS-YOLO to improve the detection of foreign objects in subway cracks was proposed to improve YOLOv10 for the detection of foreign objects in subway cracks. In order to improve the image quality in the dark environment of subway gaps, considering the weights of some features at different scales, the parameter-free attention mechanism is introduced into the SSS-YOLO model, and the spatial position information is strengthened to reduce the amount of information loss, and finally the Shape-IOU loss function is used to enhance the accuracy of small target detection and regression prediction frame, and improve the detection accuracy of small and small targets in the gap. The experimental results show that the accuracy of the method reaches 90.90%, and the average detection accuracy is increased by 3.62%.

**Keywords:** deep learning; foreign body detection; YOLOv10; low-light networks; attention mechanisms

## 0 引言

在国内地铁轨道系统中,轨道与车门边缘站台并不是完全无缝对接的,这是因为考虑到轨道对齐的精度限制以及列车自身的轻微摆动留有阈值,但也以此造成了很大的安全隐患<sup>[1]</sup>。当屏蔽门打开,乘客穿行时,难免会有异物掉落其中,如小件行李、衣物等卡在列车门与站台屏蔽门之间的缝隙中,可能会导致列车延误,影响整体的运营时间,因此而造成事故也不在少数,甚至可能会使乘客卷入缝隙进而导致受伤<sup>[2]</sup>。此外,轨道上的异物如工具、掉落的导电

类的金属类小物体等可能导致列车脱轨或其他事故,除了提醒乘客上下车门要注意以外,及时检测地铁缝隙内入侵的异物,预防地铁事故的发生也是保障地铁运营安全的关键因素。

目前对于地铁缝隙异物的检测手段大致可以分为两类,一类是经典的传统检测,另一类是基于计算机图像处理的方法,其中传统的检测方法主要有人为巡查,还有一些基于传感器的检测方法,巡查这种方法消耗大量人力资源同时也给人员带来极大的工作强度,另一方面无法做到实时监督的作用。王双英<sup>[3]</sup>提出了一种多传感器异物检测,采

用视觉相机、激光雷达相互融合实现对异物的检测。徐东等<sup>[4]</sup>提出了红外和超声互补的检测方法用于检测异物入侵,能够实现在不同距离上的检测效果提升。

对于计算机处理异物检测的算法,主要以图像处理为主,特别是深度学习算法主要成为研究的对象。由费莉梅等<sup>[5]</sup>提出的基于改进 ViBe 的自适应运动目标检测算法。张鹤庆等<sup>[6]</sup>利用机器视觉进行地铁缝隙的背景建模判断特征变化程度进而检测是否存在异物。徐鑫等<sup>[7]</sup>利用深度学习改进 YOLOv5,引入多种数据增强。晏朋等<sup>[8]</sup>改进 YOLOv7 的高铁异物入侵检测算法,实现了对复杂环境下高铁轨道入侵异物的检测,但是此方法增大了模型参数量,检测速度有所下降。由于地铁缝隙中异物的目标大小普遍较小,赵宗扬等<sup>[9]</sup>改进多尺度 LGF-Net,提升了小目标轨道入侵异物检测性能,而洪炎等<sup>[10]</sup>从检测速度的角度出发,提出重建 YOLOv8n 模型中的 C2f 模块的 Bottleneck,使得模型大小减少 0.2 M,检测速度进一步提升。

以上方法并没有考虑到地铁缝隙中是一种低照度的环境,而且并没有考虑到模型的参数量问题,地铁异物检测是一个实时性要求很高的场景,综合考虑以上几点,本文提出了基于 YOLOv10 算法改进的 SSS-YOLO 模型,首先在提取图片特征前引入 SCINet 增强网络结构,解决在低光照条件下物体识别问题,其次,加入无参注意力机制 SimAm 用来消除不同特征层对预测结果的影响,加大了对主要特征的权重能够更好的实现对特征的提取。最后使用 Shape-IOU 损失函数优化边界框回归问题,进一步实现边对不同大小边界框的匹配精准性。最终实现对地铁缝隙中低照度环境下的精准预测。

## 1 经典 YOLOv10 算法

YOLOv10 (you only look once v10) 是最新提出的 YOLO 无 NMS 训练的一致双重分配,消除了 YOLO 系列非极大值训练依赖,同时延续了 YOLO 系列性能强大、消耗算力较少,高效检测的特点,其高效、准确的物体检测能力使其成为实时目标检测领域的理想选择<sup>[11]</sup>。YOLOv10 模型由骨干网络(Backbone)、颈部网络(Neck)和检测头(Head)3 部分组成,具体 YOLOv10 网络结构如图 1 所示。Backbone 主要是负责对图像的特征提取,由 Conv、C2f 和 SPPF 3 个模块组成,Neck 的作用 Backbone 是在不同层提取到的特征图进行拼接,使网络浅层特征与深层特征融合。Head 采用解耦头将分类、回归和置信度等不同的任务进行分开处理,最后汇聚输出。使用 CIoU 融入了中心点距离和纵横比,考虑到了预测框与真实框的重叠区域。最后使用一致的双重分配策略<sup>[12]</sup>,得到模型最终检测结果。

## 2 改进 YOLOv10 算法模型

本文以 YOLOv10 为基础,设计了 SSS-YOLO 算法实

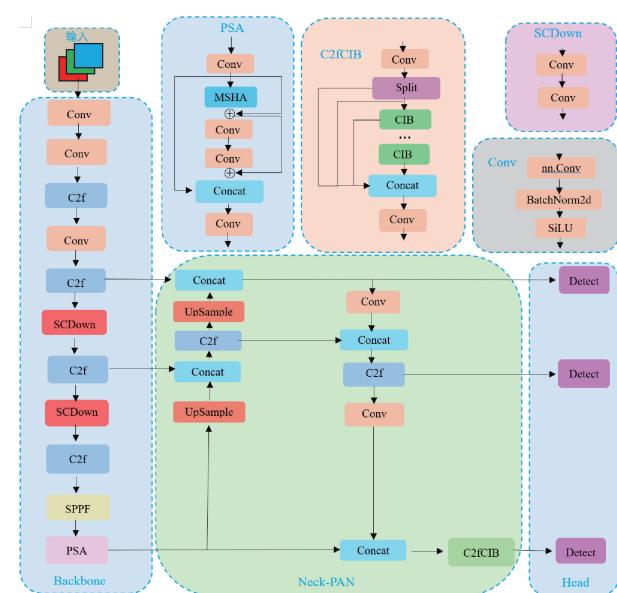


图 1 YOLOv10 模型结构

Fig. 1 Model structure of YOLOv10

现对异物检测。首先,利用低照度网络(SCINet)可以增强地铁缝隙环境中的亮度,引入照明估计模块<sup>[13]</sup>,可以显著提升低照度条件下图像中目标的识别效果,接着,无参注意力机制(SimAm),增强特征图的空间相关性<sup>[14]</sup>,更好的检测和感知到目标的所在位置,增强对异物的信息提取能力,最后,引入边界盒回归损失函数(Shape-IOU),解决现有损失函数局限性,增强回归结果准确性提升整体网络检测精度<sup>[15]</sup>。改进后的 SSS-YOLO 如图 2 所示。

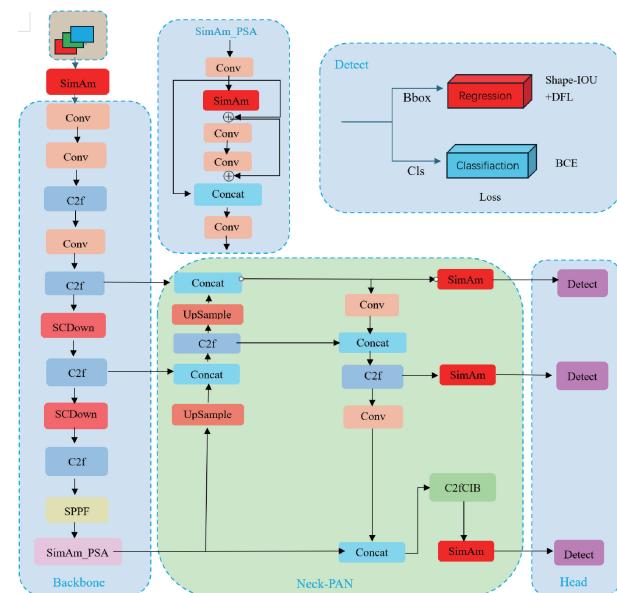


图 2 SSS-YOLO 网络模型结构

Fig. 2 SSS-YOLO network model structure

### 2.1 颜色暗小目标增强

YOLO 现有的图像技术普遍存在低密度光照下难以

处理视觉质量和计算效率,而且在未知的复杂场景中普遍无效,为了增强应对地铁缝隙暗环境下的算法检测图像质量退化问题,在 YOLOv10 基础上做了以下改进,首先引入 SCINet 低照度图像增强机制,SCINet 网络是一种无监督训练范式,该网络的结构参数量仅有 0.0003 M,属于一种实时很高的增强网络结构,它的核心在于构建了一个权重共享的光照学习过程,这一过程通过引入自校准模块来优化照明部分,从而增强图像。SCINet 网络可以分为两部分,一部分是自校正模块,一部分是照明估计模块,其结构如图 3 所示。

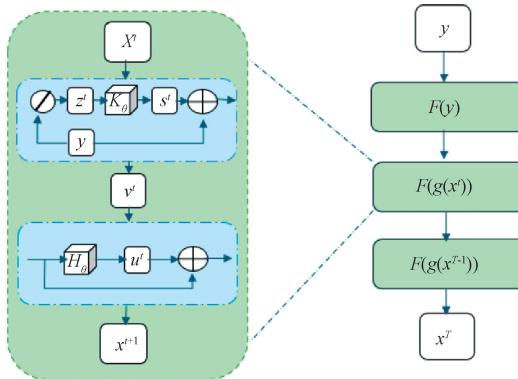


图 3 SCINet 网络结构原理图

Fig. 3 Schematic diagram of the SCINet network structure

在光照明的条件下,照明估计模块建立了一个低光照转化为高亮度优化的过程,通过连续多个处理阶段用于增强图片亮度,根据 Retinex 理论<sup>[16]</sup>,低光照观测图像等于清晰图像与光照的点乘,即  $y = z \otimes x$ ,基于该模型设计的方法中,光照的估计通常被视为主要的优化目标,得到精确的光照后,清晰图像能够上述关系直接得到,其原理如下:

$$F(x^t): \begin{cases} u^t = H_\theta(x^t), x^0 = y \\ x^{t+1} = x^t + u^t \end{cases} \quad (1)$$

其中,  $u^t$  和  $x^t$  分别表示的是第  $t$  阶段的残差与光照,  $H_\theta$  代表光照估计网络。在每一阶段光照估计网络均保持结构与参数共享状态。在渐进式优化与参数共享机制下,每个阶段均希望得到与目标接近的输出。

这是一个逐步优化的过程,每一次优化都是在前一步的基础上进一步提高图像的亮度和对比度,进而使得在暗环境下的物体也能变得清楚。为了在不损失细节的情况下提升整体图像质量,同时避免因过度调整导致的颜色失真,采用自校正模块用于校正当前阶段的图像输出,其校正公式为:

$$G(x^t): \begin{cases} z^t = y \odot x^t \\ s^t = K_\theta(z^t) \\ v^t = y + s^t \end{cases} \quad (2)$$

其中,  $v^t$  是校准后的用于下一阶段的输入,  $s^t$  表示自校正映射后的图像,  $z^t$  表示目标图像,  $K_\theta$  是参数操作

符,其中  $\theta$  参数是可学习变量,原本的光照学习过程中第二阶段及以后的输入变成了由上述公式得到的结果,也就是说光照优化过程被重新定义:

$$F(x^t) \rightarrow F(G(x^t)) \quad (3)$$

每一阶段的输入来间接地影响了每一阶段的输出,进而实现了阶段间的收敛。

## 2.2 增加无参注意力机制 SimAm

在 YOLOv10 中,骨干网络使用的是多尺度拼接进行特征融合,然而并没有考虑到不同尺度部分特征权重的重要性,这很有可能会造成不同尺度的特征对最后的预测结果造成较大的影响,而 SimAm 是一种轻量级、无参数的卷积神经网络注意力机制,它通过计算特征图的局部自相似性来生成注意力权重,用于调整不同尺度特征图,可以使网络更加专注于图像中的主要特征,而降低对次要特征的关注,使得网络的鲁棒性和抗干扰能力进一步增强。SimAm 注意力机制结构如图 4 所示,左半部分是 SimAm\_PSA,其内部的 SimAm 结构如图 4 右半部分所示。

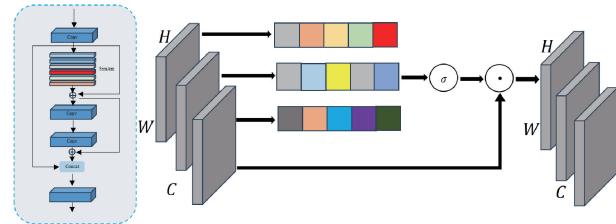


图 4 SimAm\_PSA 结构原理图

Fig. 4 SimAm\_PSA schematic diagram of the structure

SimAm 通过来计算当前神经元与其他神经元之间的线性可分性,因此,需要定义能量函数,使用二值标签并添加正则项的方式简化之后,每个神经元都有着不同级别的重要性,其中活跃的神经元可以抑制周围神经元的活跃性,所以当出现了明显抑制周围效应的神经元应该赋予更高级别的优先级,在 SimAm 中,每个神经元都定义了以下的能量函数:

$$e_t(w_t, b_t, y, x_i) = (y_t - \hat{t})^2 + \frac{1}{M-1} \sum_{i=1}^{M-1} (y_i - \hat{x}_i)^2 \quad (4)$$

其中,  $e_t$  为神经元  $t$  的能量函数,  $w_t$  和  $b_t$  表示线性变换的权值与偏置,  $y$  是这个通道上的输出,  $M$  表示在此通道上的神经元个数,  $y_i$  是其他神经元的输出,  $x_i$  是周围其他神经元,  $\hat{t}$  与  $\hat{x}_i$  是由  $t$  和  $x_i$  经过线性变化而来的。假设单个通道中的所有像素都遵循相同的分布,则每个独立神经元的重要程度可以通过以下最小能量函数来定义:

$$e_t^* = \frac{4(\hat{\sigma}^2 + \lambda)}{(t - \hat{\mu})^2 + 2\hat{\sigma}^2 + 2\lambda} \quad (5)$$

式中:  $\hat{\mu} = \frac{1}{M} \sum_{i=1}^M x_i$ ,  $\hat{\sigma}^2 = \frac{1}{M-1} \sum_{i=1}^{M-1} (x_i - \mu_t)^2$ ,  $e_t^*$  的值反映了重要程度,它的值越小说明与周围神经元的区别越大,

线性可分性也越强,因此 SimAm 将其作为权重系数,并使用归一化激活函数限制大小,得到新的特征图  $X'$ :

$$X' = \text{sigmoid}\left(\frac{1}{E}\right) \cdot X \quad (6)$$

### 2.3 Shape-IOU 损失函数优化

交并比(IoU)是一个用于衡量两个矩形框(预测框和真实框)重合程度的指标,决定预测的定位性能。YOLOv10 使用的是 CIoU 损失函数计算定位损失,但是这种损失函数没有考虑到所需真实框与预测框之间不匹配的方向,忽略了边界框自身的形状和尺寸也会对边界框回归产生影响的事实,同时它无法根据不同的检测器的检测目的进行自我调节,鲁棒性不强。本文引入了一种 Shape-IoU,一种基于边框回归损失函数的目标检测方法,能够通过聚焦边框自身形状与自身尺度来提高检测效果,对边界框的形状和尺度能更加的精准度量,如图 5 所示。

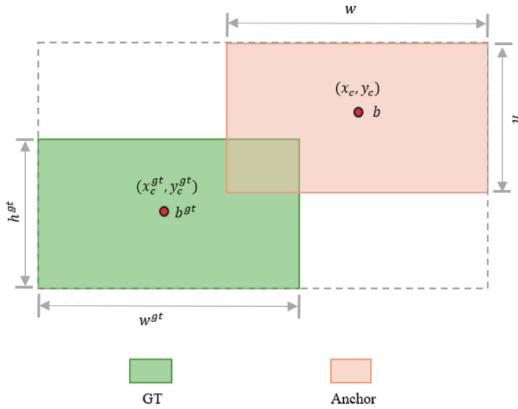


图 5 Shape-IoU 方法损失计算

Fig. 5 Shape-IoU method loss calculation

在 Shape-IoU 方法中,能够通过框自身形状和自身尺度来计算损失,进而使边界框的回归更加准确,其原理如下:

$$IoU = \frac{|B \cap B^{gt}|}{|B \cup B^{gt}|} \quad (7)$$

$$ww = \frac{2 \times (w^{gt})^{scale}}{(w^{gt})^{scale} + (h^{gt})^{scale}} \quad (8)$$

$$hh = \frac{2 \times (h^{gt})^{scale}}{(w^{gt})^{scale} + (h^{gt})^{scale}} \quad (9)$$

$$dis^{shape} = hh \times (x_c - x_c^{gt})^2 / c^2 + ww \times (y_c - y_c^{gt})^2 / c^2 \quad (10)$$

其中, scale 为尺度因子,与数据集中目标的尺度有关,  $ww$  与  $hh$  分别为水平方向与竖直方向的权重系数,其取值与 GT 框的形状有关。其对应的边框回归损失如下:

$$\Omega^{shape} = \sum_{t=w,h} (1 - e^{-\omega_t})^\theta, \theta = 4 \quad (11)$$

$$Loss_{\text{Shape-IOU}} = 1 - IoU + dis^{shape} + 0.5 \times \Omega^{shape} \quad (12)$$

## 3 实验结果与分析

### 3.1 实验环境与评价指标

由于关于地铁缝隙异物的公开数据集较少,实验数据是从实验室搭建的地铁缝隙平台自主拍摄的 800 张异物照片,同时扩充了 SpotGarbage 上的一些常见的物体图片 700 张,涵盖了不同大小,不同光线照明条件的图像。同时为了增强数据集的泛化能力,将这些数据集随机的进行裁剪,拼接,旋转,随机覆盖等操作,扩充到 5 000 张数据集,类别包括 box、knife、cup、bottle、stick、rope、racket、tape 和 bag 等 9 个类别,同时将其中 3 500 张作为训练集,1 000 张作为验证集,500 张作为测试集,数据集中的物体尺寸大小分布如图 6 所示,显示出大多数物体的长宽比例小于原图像尺寸的 0.1 倍左右,符合地体缝隙内小目标的相对尺寸定义,这一特征为小目标检测算法的研究提供了很好的实验基础。

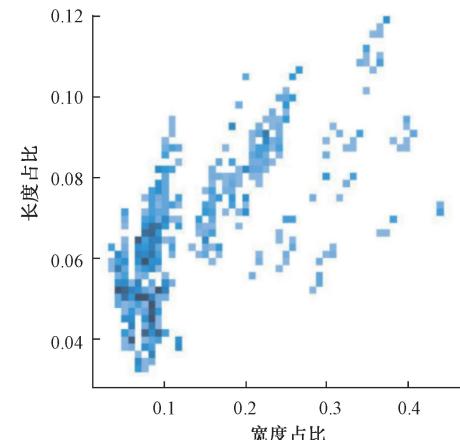


图 6 数据集物体大小分布

Fig. 6 Dataset object size distribution

实验软件使用了 OpenCV 和 PyCharm 开发工具,版本为 OpenCV4.8.1 和 PyCharm2023,实验使用的硬件平台参数如表 1 所示。

表 1 软硬件实验平台

Table 1 Software and hardware experimental platform

配置	参数
CPU	Intel(R) Core(TM) i5-8300H
GPU	NVIDIA GTX 1050 Ti
操作系统	Windows10
内存	16 G
模型	OpenCV4.8.1

模型的训练轮次为 100 个 Epoch,每批次训练数量尺度大小(size)为 16,初始学习率为 0.01,尺度因子设置

为 1.2。

### 3.2 模型性能评价指标

评估 YOLO 模型中,本文主要从精度(precision, P)、平均精度(mean average precision, mAP),参数量(Param),这 3 个评价指标来综合评估模型的性能表现,实验选用 mAP@0.5 和 mAP@0.5:0.95 作为性能参考的评价标准。具体公式如下:

$$P = \frac{T_p}{T_p + F_p} \times 100\% \quad (13)$$

$$R = \frac{T_p}{T_p + F_n} \times 100\% \quad (14)$$

其中,  $T_p$  表示正样本类别为正类且预测结果也为正类的样本数量;  $F_p$  表示样本是负样本类别但预测结果为正类的模型样本数量;  $F_n$  表示样本是正样本但预测结果为负类的模型样本数量。同时计算平均精度的公式如下:

$$AP = \int_0^1 P(R) dR \quad (15)$$

$$mAP = \frac{\sum AP}{n} \quad (16)$$

其中,  $AP$  表示的是数据集中标定的其中某一类的平均准确率,其大小反应的是模型预测的准确率,值越大说明模型预测的准确率越高。 $mAP$  反映表示的是平均准确率的平均,也就是所有类别预测准确率的平均值,  $n$  代表的是数据集中类别的个数。

在复杂度上,模型参数量(Params)是指在模型训练过程中所涉及的参数总量,也是衡量模型空间复杂度和规模的重要指标。

### 3.3 对比实验

#### 1) 不同注意力机制的对比试验

为了验证在不同注意力机制模块对 YOLOv10 网络性能造成的影响结果,在深度学习中常见的注意力机制模块有 GAM、SE 注意力机制和 ECA 注意力机制,几个注意力模块在其他条件保证一样的情况下,在模型的同一位置引入不同的注意力模块做对比实验。实验结果如表 2 所示。

表 2 注意力机制模块效果对比

Table 2 Comparison of the effect of the attention mechanism module

模块	P/%	mAP/%	Params/M
YOLOv10	87.99	87.82	2.3(+0.0)
v10+SE	88.56	88.21	3.7(+1.4)
v10+GAM	89.27	89.15	5.4(+2.1)
v10+ECA	87.12	86.53	3.8(+1.5)
v10+SimAm	<b>90.11</b>	<b>89.52</b>	<b>2.4(+0.1)</b>

可以看出 GAM 和 SE 注意力机制虽然在一定程度上提高模型的平均精度,但是引入了大量参数,使得模型运算

量大大增加,运行复杂度增高,从 ECA 注意力机制上也可以看出并不是所有的注意力机制都能优化模型,添加 ECA 注意力机制反而使得精度减少了 1.29%,SimAm 可以在几乎不引入参数的情况下,能够有效的提高原模型 1.7% 的平均精度,因此,SimAm 注意力机制相较于其他更有优势。由表 2 可知,整合注意力机制之后的改进模型在各个指标方面相对于其他改进此算法表现的更为突出,图 7 展示了不同注意力机制的热力图分布。

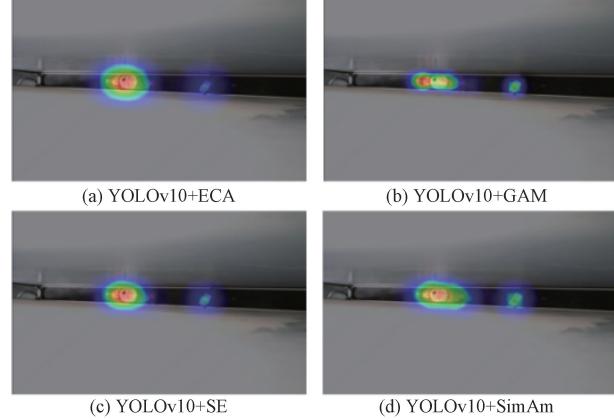


图 7 不同注意力机制下热力图对比

Fig. 7 Comparison of heat maps under different attention mechanisms

由图 7 可知,在 SimAm 注意力机制上目标注意力权重更贴合实际物体,可以有效地增强模型的检测性能。

#### 2) 消融对比试验

以 YOLOv10 模型为基础,通过消融实验以此引入不同模块以此验证不同模块的有效作用,实验结果表示:添加了低照明系统网络 SCINet,在不引入参数量的条件下可以看出,平均检测精度提高了 0.33%,添加了 SimAm 虽然在参数量上增加了 0.1 M,但是在平均检测精度上提高了 1.44%,由表 3 可知,在兼顾参数量和检测精度的平衡下,SimAm 网络是最优方案,在添加 SimAm 和 SCINet 的基础上又引入 Shape-IOU 可以进一步提高 0.17% 的平均精度,由此可见 Shape-IOU 损失函数(以下简写 S-IOU)是可以提高检测器的模型的性能。

表 3 消融实验对比

Table 3 Comparison of ablation experiments

模块	P/%	mAP/%	Params/M
YOLOv10	87.99	87.82	2.3
v10+SCINet	88.31	88.15	2.3
v10+S-IOU	89.42	89.26	2.3
v10+SimAm	90.11	89.52	2.4
v10+SimAm+S-IOU	90.62	89.89	2.4
v10+SimAm+S-IOU+SCINet	<b>90.90</b>	<b>90.06</b>	2.4

从消融实验可以看到在原有模型中逐级添加 SCINet

和 S-IOU 模块对于模型的参数量并没有发生变化,但是能够在很大程度上提高模型的检测精度,引入了 SimAm 后,虽然模型大小增大了 0.1 M,但是精度提高了 2.12%,表 3 展现了不同模块对网络模型的具体影响作用效果。

图 8 显示了 YOLOv10 改进前后 mAP @ 0.5 和 mAP@0.5: 0.95 对比曲线,可以明显的看出改进后的模型性能优于原模型,证明了改进的有效性。

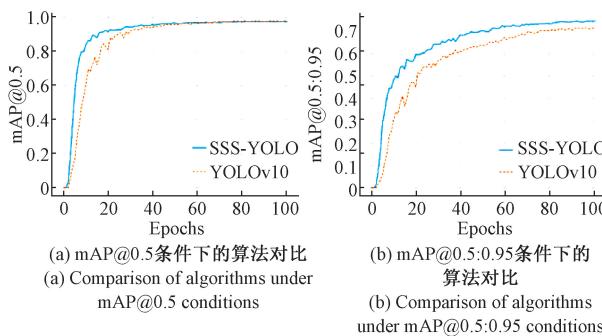


图 8 改进算法与原算法精度对比曲线图

Fig. 8 Comparison curve of the accuracy of the improved algorithm and the original algorithm

### 3) 不同模型检测效果对比

为了验证改进算法的性能,在数据集和实验环境包括实验参数设置相同的情况下,将本算法和目前较为主流 YOLO 系列变体以及先进的目标检测 DeeplabV3-CSPNet

等算法做对比,实验结果如表 4 所示。

表 4 不同系列算法检测对比

Table 4 Comparison of different algorithms

模块	P/%	mAP/%	Params/M
YOLOv5n <sup>[6]</sup>	85.00	83.15	2.51
YOLOv7-tiny <sup>[8]</sup>	82.59	81.27	6.02
YOLOv8n <sup>[10]</sup>	83.78	81.11	3.16
YOLOv10n <sup>[11]</sup>	85.12	81.51	2.30
DeeplabV3-CSPNet <sup>[17]</sup>	84.17	81.09	7.80
CNN-BiLSTM <sup>[18]</sup>	85.70	82.24	4.21
U-Net <sup>[19]</sup>	86.21	83.14	7.76
<b>SSS-YOLO</b>	<b>88.24</b>	<b>85.13</b>	2.38

通过表 4 的实验结果分析,SSS-YOLO 算法在平衡参数量和检测正确率的前提下,平均精度相对于原 YOLOv10 算法提高了 3.62%,相对于 YOLOv8n 提高了 4.2%,与 DeeplabV3-CSPNet 和 U-Net 相比,分别提高了 4.04% 和 1.99%,同时在参数量上也减少了数倍,与 CNN-BiLSTM 相比无论是在准确率还是在参数优化上都更胜一筹,这也表明此模型更具有实用性,使得算法模型在算力较低的设备上部署成为了可能。

为了进一步探究 SSS-YOLO 算法的改进效果,选取了数据集中不同条件下的异物图片,使用 YOLOv10n 和 SSS-YOLO 模型的检测效果做对比,结果如图 9 所示。

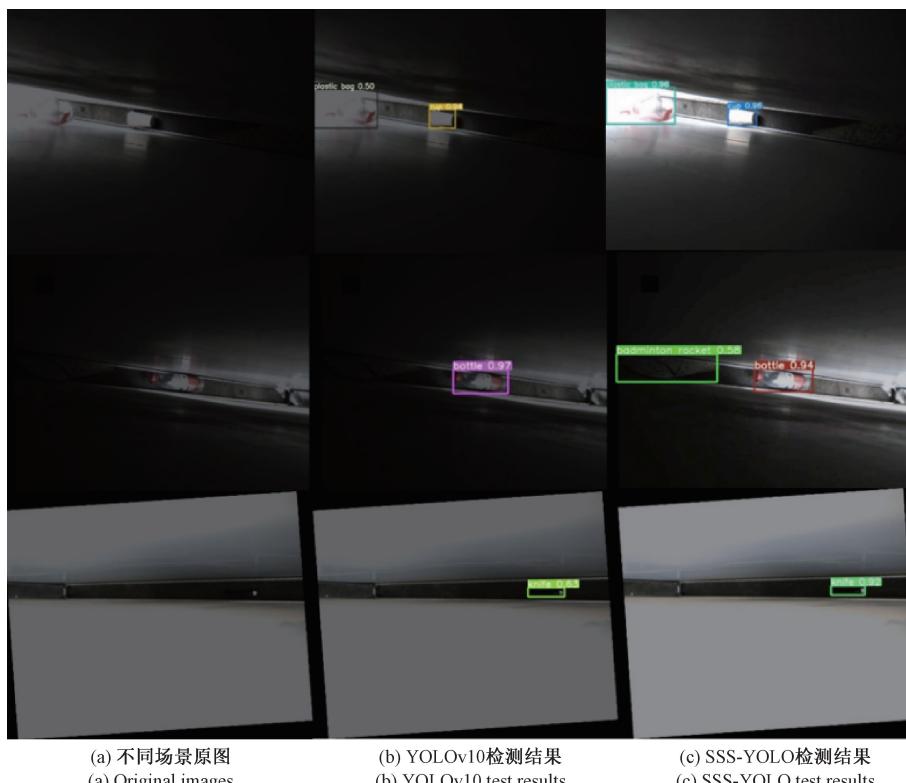


图 9 不同场景下改进算法检测效果对比

Fig. 9 Comparison of the detection effect of the improved algorithm in different scenarios

以看出,在多目标光线较好的条件下原 YOLOv10 相比较于改进 SSS-YOLO 算法,虽然也可以将目标检测出来,但是存在检测框不精准,与实际物体有部分出入,SSS-YOLO 检测框可以较好的拟合物体,且每种物体的置信度较高。在低光照多物体条件下,原 YOLOv10 算法检测精度较低,且会出现漏检的现象,而 SSS-YOLO 算法在引入了 SCINet 网络后,可以在不损害图像质量的条件下提升图像亮度,使得检测精度进一步提高。在小目标检测

中 SSS-YOLO 也能准确的检测出物体并给出较高的置信度。

#### 4) SSS-YOLO 泛化性分析

为了进一步验证算法的泛化性能和鲁棒性,实验的训练集使用的实验室平台采集的异物数据集,而泛化实验数据集使用的是从真实环境下采集的 500 张异物图片,包括不同物体不同位置的随机组合,图 10(a)为部分样张,实验结果如图 10(b)所示。



图 10 真实环境下泛化实验样张

Fig. 10 Samples of generalization experiments in real-world environments

在其他条件都相同的条件下,使用 YOLOv10 和改进后的 SSS-YOLO 算法模型进行泛化性对比,实验结果如表 5 所示。可见,提出的改进算法检测的泛化性明显高于原算法,准确率提升 3.38%,同时识别异物的种类可以给出更高的置信度。

表 5 前后模型效果对比

Table 5 Comparison of the effects of different models

模块	P	mAP	%
YOLOv10	86.16	85.63	
SSS-YOLO	89.54	88.76	

## 4 结论

本文提出了一种基于 YOLOv10 改进的地铁缝隙异物检测模型算法 SSS-YOLO,原有模型对地铁缝隙内低光照条件下识别效果不好,漏识别,以及对于预测边界框的回归问题,针对性的提出了改进。首先在主干网络首段添加低照度增强网络 SCINet,在进行特征提取前先进行图像增强,骨干网络中加入 SimAm 注意力机制,替换 Shape-IOU 损失函数使回归预测框更加精确,进一步解决模型的误检漏检问题。

参数量,改进的模型能够很大程度上提高对异物检测的识别精度,在原有模型的基础上,平均检测精度提高了 3.62%,且与其他主流的 YOLO 系列检测模型如 YOLOv5n、YOLOv7-tiny、YOLOv10 相比分别提高了 1.98%、3.86%、3.62% 的平均精度,跟目前比较先进的算法如 DeeplabV3-CSPNet 虽然精度上有小幅提升,但是在参数量上却小数倍,同时在自建数据集以及真实场景中可以有效的检测出异物,对地铁安全自动化检测具有应用价值和现实意义。

## 参考文献

- [1] 王先进,蔡昌俊,杨新征,等.中国城市轨道交通运营发展报告(2022—2023)[M].北京:社会科学文献出版社,2023.  
WANG X J, CAI CH J, YANG X ZH, et al. Report on the operation and development of urban rail transit in China (2022—2023) [M]. Beijing: Social Sciences Academic Press, 2023.
- [2] 侯妍君,李鑫磊,申樟虹.站台门与列车车门间异物探测装置的探讨[J].机电工程技术,2020,49(3):65-66,102.  
HOU Y J, LI X L, SHEN ZH H. Discussion on foreign object detection device between platform door and train door [J]. Mechatronics Engineering Technology, 2020, 49(3): 65-66, 102.
- [3] 王双英.基于多传感器融合的轨道异物检测系统[J].

以上的实验结果表明,引入的模块并没有增大模型的

- 今日制造与升级,2023(12):87-90.
- WANG SH Y. Orbital foreign body detection system based on multi-sensor fusion[J]. Today's Manufacturing and Upgrade,2023(12):87-90.
- [4] 徐东,李凯霖,王芃,等.红外和超声波互补式入侵检测方法[J].仪表技术与传感器,2022(5):93-97,104.
- XU D, LI K L, WANG P, et al. Infrared and ultrasonic complementary intrusion detection method [ J ]. Instrumentation Technology and Sensors, 2022 ( 5 ): 93-97,104.
- [5] 费莉梅,田翔,郑博仑.基于改进 ViBe 的自适应运动目标检测算法[J].计算机工程与设计,2024,45 ( 6 ): 1771-1779.
- FEI L M, TIAN X, ZHENG B L. Adaptive moving target detection algorithm based on improved ViBe [ J ]. Computer Engineering and Design, 2024, 45 ( 6 ): 1771-1779.
- [6] 张鹤庆,张金鑫.基于机器视觉的地铁站台门与列车门间异物检测方法研究[J].现代城市轨道交通,2020(7):51-55.
- ZHANG H Q, ZHANG J X. Research on foreign object detection method between subway platform door and train door based on machine vision[J]. Modern Urban Rail Transit,2020(7):51-55.
- [7] 徐鑫,潘杰,曹利安,等.基于深度学习的铁路异物侵限检测模型[J].铁路计算机应用,2023,32(10):7-12.
- XU X, PAN J, CAO L AN, et al. A deep learning based railway foreign object intrusion detection model [ J ]. Railway Computer Applications, 2023, 32(10): 7-12.
- [8] 晏朋,李遇鑫,李治林,等.基于改进 YOLOv7 的高铁异物入侵检测算法[J].无线电工程,2024,54(5):1099-1109.
- YAN P, LI Y X, LI ZH L, et al. Foreign body intrusion detection algorithm for high-speed railway based on improved YOLOv7[J]. Radio Engineering, 2024,54(5):1099-1109.
- [9] 赵宗扬,康杰虎,梁健,等.基于 LGF-Net 的全天候轨道入侵异物智能检测系统[J].仪器仪表学报,2023,44(9):287-301.
- ZHAO Z Y, KANG J H, LIANG J, et al. All-weather orbital foreign object intelligent detection system based on LGF-Net [ J ]. Chinese Journal of Scientific Instrument,2023,44(9):287-301.
- [10] 洪炎,汪磊,苏静明,等.基于改进 YOLOv8 的煤矿输送带异物检测[J].工矿自动化,2024,50(6):61-69.
- HONG Y, WANG L, SU J M, et al. Detection of foreign objects in coal mine conveyor belt based on improved YOLOv8[J]. Journal of Mine Automation, 2024,50(6):61-69.
- [11] KIM J Y. Comparison analysis of YOLOv10 and existing object detection model performance [ J ]. Journal of The Korea Society of Computer and
- Information, 2024, 29(8): 85-92.
- [12] 陈麒羽,邵洁,王超群,等.基于双重动态调整的改进非洲秃鹫优化算法[J].国外电子测量技术,2024,43(1):20-29.
- CHEN Q Y, SHAO J, WANG CH Q, et al. Improved African vulture optimization algorithm based on double dynamic adjustment[J]. Foreign Electronic Measurement Technology,2024,43(1):20-29.
- [13] LIU M, LI Y, HU J, et al. A new hybrid model based on SCINet and LSTM for short-term power load forecasting[J]. Energies, 2023, 17(1): 95.
- [14] MAHAADEVAN V C, NARAYANAMOORTHI R, GONO R, et al. Automatic identifier of socket for electrical vehicles using SWIN-transformer and SimAM attention mechanism-based EVS YOLO[J]. IEEE Access, 2023,11(3):111238-111254.
- [15] 刘龙哲,刘刚,徐红鹏,等.面向单阶段目标检测的损失函数优化设计[J].电光与控制,2024,31(3):86-93.
- LIU L ZH, LIU G, XU H P, et al. Optimal design of loss function for single-stage target detection [ J ]. Electro-Optics and Control,2024,31(3):86-93.
- [16] YEONAN-KIM J, BERTALMIO M. Analysis of retinal and cortical components of Retinex algorithms [ J ]. Journal of Electronic Imaging, 2017, 26 ( 3 ): 031208-031215.
- [17] 熊锐,程亮,胡涛,等.水面无人艇可行域及障碍物快速分割算法研究[J].电子测量与仪器学报,2023,37(2):11-20.
- XIONG R, CHENG L, HU T, et al. Research on feasible domain and obstacle fast segmentation algorithm for unmanned surface vehicle[J]. Journal of Electronic Measurement and Instrumentation, 2023, 37(2):11-20.
- [18] 梁欣怡,行鸿彦,侯天浩.基于自监督特征增强的 CNN-BiLSTM 网络入侵检测方法[J].电子测量与仪器学报,2022,36(10):65-73.
- LIANG X Y, XING H Y, HOU T H. CNN-BiLSTM network intrusion detection method based on self-supervised feature enhancement [ J ]. Journal of Electronic Measurement and Instrumentation, 2022, 36(10):65-73.
- [19] YADAVENDRA, CHAND S. Semantic segmentation of human cell nucleus using deep U-Net and other versions of U-Net models[J]. Network: Computation in Neural Systems, 2022, 33(3-4): 167-186.

## 作者简介

何赛赛(通信作者),硕士研究生,主要研究方向为计算机视觉图像处理。

E-mail:1960583048@qq.com

王文胜,博士,主要研究方向为目标检测与识别和计算机视觉。

E-mail:wangwensheng@bistu.edu.cn