

改进YOLOv10n的复杂场景设备检测算法

刘心怡^{1,2}, 卢 棚¹, 刘少清^{1*}

¹合肥综合性国家科学中心能源研究院(安徽省能源实验室), 安徽 合肥

²安徽理工大学计算机科学与工程学院, 安徽 淮南

收稿日期: 2025年12月16日; 录用日期: 2026年1月15日; 发布日期: 2026年1月23日

摘要

针对加速器中子源设备巡检场景中结构复杂、设备类型多样且易受遮挡干扰等问题, 现有目标检测方法在检测精度与稳定性方面仍存在不足。为此, 本文提出了一种基于改进YOLOv10的复杂场景设备检测算法: ECS-YOLO。首先, 在特征提取阶段引入EffectiveSE注意力机制, 对通道特征进行自适应重标定, 增强关键设备特征表达能力。其次, 设计C2f_RFACConv模块, 将RFA注意力机制嵌入C2f结构, 在不显著增加计算开销的前提下提升多尺度特征提取与感知能力。最后, 在损失函数中引入SIoU损失, 从距离、角度与形状匹配等多维度约束边界框回归过程, 提高目标定位精度。实验结果表明, 与基线模型相比, 改进的ECS-YOLO模型在mAP@50、P、R上分别提高了3.7%, 3.5%, 4.8%。

关键词

YOLOv10, 复杂场景, 设备检测, EffectiveSE, SIoU

Improved YOLOv10n-Based Device Detection Algorithm for Complex Scenes

Xinyi Liu^{1,2}, Peng Lu¹, Shaoqing Liu^{1*}

¹Institute of Energy, Hefei Comprehensive National Science Center (Anhui Energy Laboratory), Hefei Anhui

²School of Computer Science and Engineering, Anhui University of Science and Technology, Huainan Anhui

Received: December 16, 2025; accepted: January 15, 2026; published: January 23, 2026

Abstract

In accelerator neutron source equipment inspection scenarios, the complex structural environment, diverse equipment types, and frequent occlusion pose significant challenges to existing object detection methods, leading to insufficient detection accuracy and stability. To address these issues,

*通讯作者。

this paper proposes an improved YOLOv10-based device detection algorithm for complex scenes, termed ECS-YOLO. First, the EffectiveSE attention mechanism is introduced in the feature extraction stage to adaptively recalibrate channel features, thereby enhancing the representation of critical device features. Second, a C2f_RFACconv module is designed by embedding the RFA attention mechanism into the C2f structure, which improves multi-scale feature extraction and perception capability without significantly increasing computational overhead. Finally, the SIoU loss function is incorporated to constrain bounding box regression from multiple aspects, including distance, angle, and shape matching, further improving localization accuracy. Experimental results demonstrate that, compared with the baseline model, ECS-YOLO achieves improvements of 3.7%, 3.5%, and 4.8% in mAP@50, Precision, and Recall, respectively.

Keywords

YOLOv10, Complex Scenes, Device Detection, EffectiveSE, SIoU

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

随着肿瘤放射治疗技术的不断发展，硼中子俘获治疗(Boron Neutron Capture Therapy, BNCT)作为一种新型精准放疗手段，因其对肿瘤细胞的选择性杀伤能力而受到广泛关注。该技术通过向患者体内引入含硼药物，并利用热中子诱发核反应，从而实现对肿瘤组织的定向破坏。其中，加速器中子源作为BNCT系统的关键组成部分，其运行状态直接关系到治疗过程的安全性与稳定性。在实际应用中，加速器中子源设备通常工作于结构复杂、设备密集的环境中，对其进行长期、稳定的运行监测与巡检具有重要意义。然而，传统依赖人工的巡检方式在效率、准确性及一致性方面均存在明显不足，尤其在复杂场景下易出现漏检与误判问题。为此，有必要引入一种面向复杂环境的智能化设备检测方法，以辅助巡检人员实现对设备状态的高效、精准感知。

在复杂环境应用中，曾龙军[1]等通过改进 YOLOv8n 网络结构，结合轻量化骨干与多尺度特征融合策略，有效提升了自然场景下病害目标的检测精度；熊诗雨[2]等在 YOLOv11 框架基础上针对远距复杂场景小目标检测进行了结构优化，显著增强了模型对极小目标的感知能力。针对工业复杂场景中的遮挡与光照干扰问题，徐中懿[3]等通过引入协同注意与动态卷积机制，提高了模型在复杂工业环境下的检测鲁棒性。此外，郭莉[4]等基于 YOLOv10 对人体跌倒检测任务进行了改进，验证了该系列模型在复杂动态场景中的适用性；薛光辉[5]等提出的 PCSED-YOLO 在跨尺度多目标检测任务中取得了较好效果。

综上所述，尽管现有研究在复杂场景目标检测方面取得了一定进展，但在高精度检测方面仍存在不足。针对复杂设备巡检场景，本文基于 YOLOv10n 对模型结构进行进一步优化，以提升模型在复杂环境下的检测精度与实用性。

2. YOLOv10n 改进算法

YOLO 系列算法以较高的检测精度和实时性能著称，已在目标检测任务中得到广泛应用[6]。由清华大学团队开源的 YOLOv10 [7]是当前具有代表性的单阶段目标检测模型之一。该模型在 YOLOv8 [8]基础结构上进行了多项改进，引入了 PSA、SCDown 以及 C2fCIB 等模块，以增强特征表达能力和网络效率。针对传统后处理阶段存在的计算冗余问题，YOLOv10 采用一致性的双样本分配策略，取消了非极大值抑

制过程, 从而降低了推理开销。同时, 通过引入轻量化分类头并结合空间通道解耦的下采样方式, 进一步提升了模型的整体推理速度。

针对加速器设备巡检任务, 本文提出了一种基于改进 YOLOv10 的加速器设备检测算法, 其整体网络结构如图 1 所示。以 YOLOv10 为基准模型, 首先在特征提取阶段引入 EffectiveSE [9] 注意力机制, 对通道特征进行自适应重标定, 以增强网络对关键设备特征的表达能力并抑制冗余信息。其次, 设计 C2f_RFACConv [10] 模块, 将 RFA 注意力机制嵌入 C2f 结构中, 在保持模型轻量化特性的同时提升特征提取能力和多尺度目标感知能力。最后, 在损失函数设计中引入 SIoU [11] 损失函数, 从距离、角度及形状匹配等多个维度对边界框回归进行约束, 从而进一步提升模型在复杂场景下的定位精度。

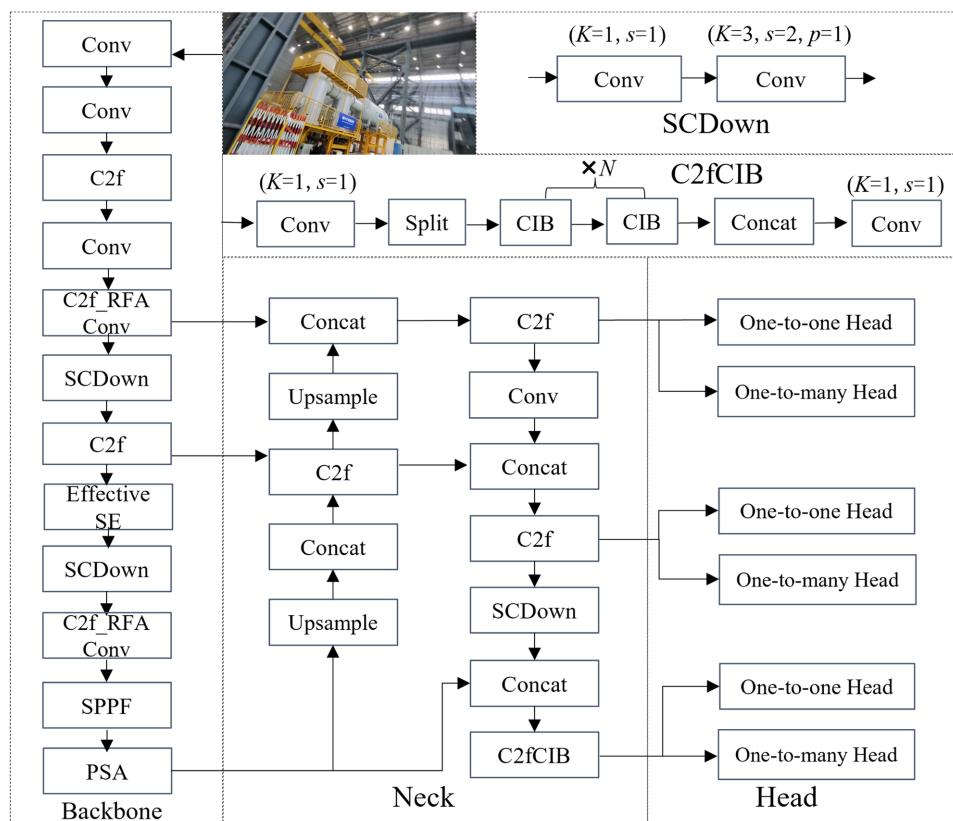


Figure 1. Architecture of the ECS-YOLO network

图 1. ECS-YOLO 网络结构图

2.1. EffectiveSE 注意力机制

EffectiveSE 是在 Squeeze-and-Excitation (SE) [12] 注意力机制基础上改进得到的一种轻量级通道注意力模块, 其核心目标是在降低计算开销的同时提升特征表达能力。传统 SE 模块通过对通道维度进行自适应建模, 动态调节不同通道在特征表示中的重要性, 从而增强网络对关键信息的感知能力。经典 SE 模块主要由 Squeeze 和 Excitation 两个阶段构成。其中, Squeeze 阶段采用全局平均池化对每个通道的空间信息进行压缩, 生成反映通道整体特征的描述向量; 随后, 在 Excitation 阶段中, 通过两层全连接层及非线性激活函数建模通道间的依赖关系, 并利用 Sigmoid 函数输出通道权重, 用于对输入特征图进行重新加权。然而, 该结构依赖全局平均池化和全连接运算, 不仅增加了模型的计算复杂度, 同时在压缩过程中可能丢失部分具有判别力的空间信息, 从而影响特征表示效果。针对上述不足, EffectiveSE 对传统 SE 结

构进行了针对性优化。该模块首先对输入特征图依次采用多个 3×3 卷积进行特征提取，并将得到的特征进行拼接融合，随后通过 1×1 卷积生成尺寸为 $C \times W \times H$ 的特征映射。基于该特征映射，EffectiveSE 通过加权方式直接计算各通道的注意力权重，并将其作用于原始特征图，实现通道维度的自适应增强。

与传统 SE 模块相比，EffectiveSE 在结构设计上更加简洁，有效减少了全连接层带来的计算负担，在保持甚至提升模型性能的同时，更好地保留了特征图中的空间信息，有助于增强网络对细粒度特征的表达能力。因此，本文将 EffectiveSE 模块引入 ECS-YOLO 网络，以强化特征表示能力，从而提升加速器设备的检测精度。

2.2. C2f_RFACConv 模块

在对目标进行检测时，传统的卷积存在着一定的局限：首先，感受野大小是固定的。这说明每个神经元只能关注所给图像特定范围内的像素，但是在实际的检测任务中，不同位置像素的重要性是有差异的；其次，传统的卷积获取特征信息是通过共享参数的滑动窗口来实现的。因此，在计算过程中难以获得不同位置的信息差异，导致小目标的特征无法有效地被 C2f 模块捕获。针对上述问题，将感受野注意力卷积(Receptive Field Attention Convolution, RFACConv)替换主干网络里 C2f 模块中 Bottleneck 的 Conv。该模块能够把空间特征信息通过卷积的方式进行重新排列，较好地解决了传统卷积核参数共享的难题。使模型在获取每个空间特征信息的同时减少其计算复杂度。感受野注意力空间注意力机制通过构建精准的注意力映射关系，对每个特征点进行强化。把通道的数量设置为 1，能够得出 RFA 的运算公式为：

$$\begin{cases} R_1 = X_1 \times L_1 \\ R_2 = X_2 \times L_2 \\ \vdots \\ R_n = X_n \times L_n \end{cases} \quad (1)$$

其中， R 是各个滑动窗口经过运算特征值后得出的结果。 X 是特征图宽度、高度和通道数的乘积，即输入的特征图。 L 是特征图中不同位置的信息。 n 是空间滑块的总数量。

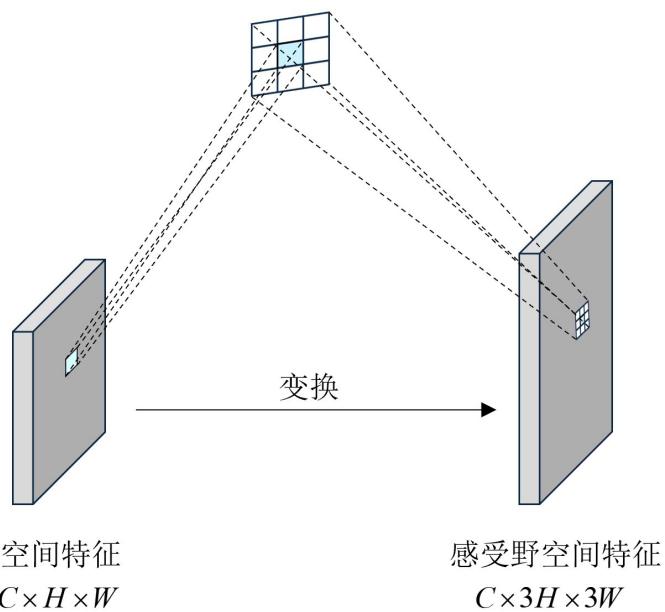


Figure 2. Spatial feature transformation
图 2. 空间特征变换

基于此,空间注意力机制 RFA 能够有效地解决卷积神经网络运算过程中出现使用相同参数的问题。将 RFA 与卷积操作相结合得到 RFACConv 模块,该模块能够将空间特征变换为感受野空间特征,如图 2 所示。图中 3×3 大小窗口是感受野滑块,感受野滑块内有着不同重要性的特征信息,根据重要性将感受野空间特征进行排序,这种方法有效地解决了传统卷积参数共享的问题。

RFACConv 模块的卷积结构图如图 3 所示。

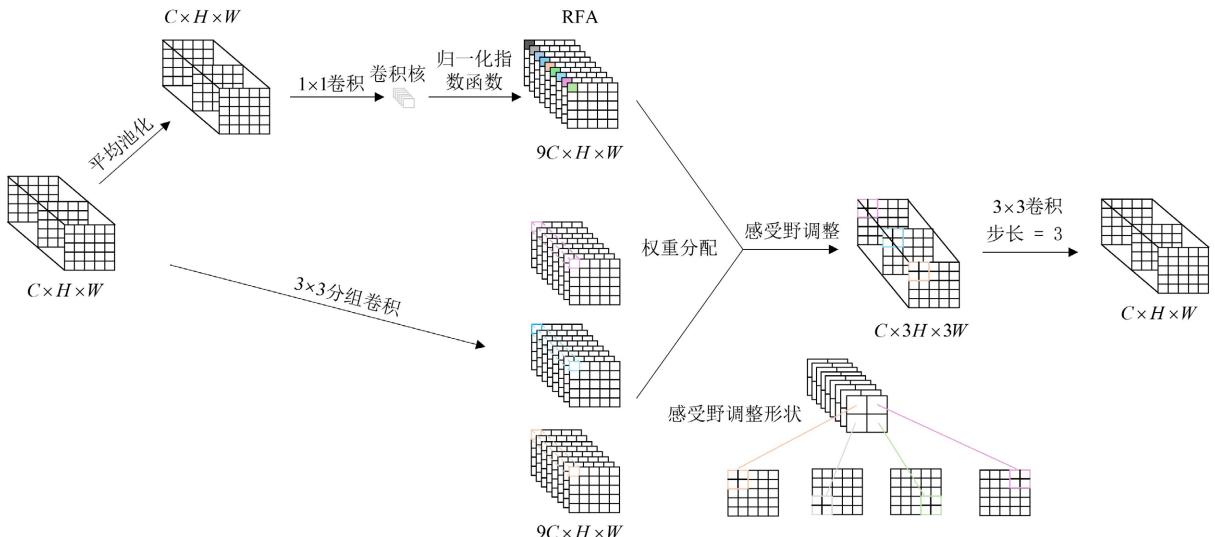


Figure 3. Architecture of the RFACConv convolution module

图 3. RFACConv 卷积结构图

其具体的运行步骤为:首先,通过平均池化将各个感受野特征的信息进行聚合。然后,采用 1×1 组卷积进行信息交互操作。最后,通过归一化指数函数的方法对感受野特征中各个特征参数的重要性进行强调。与此同时,通过 3×3 组卷积将感受野空间特征中的信息进行展开,此时特征信息的通道数增长到 9 倍,能够避免相邻特征信息出现相同的情况。

生成注意力分支的计算公式为:

$$A_{RF} = \text{Softmax} \left\{ g^{1 \times 1} \left[\text{AvgPool}(X) \right] \right\} \quad (2)$$

感受野空间特征的计算公式为:

$$F_{RF} = \text{ReLU} \left[\text{Norm} \left[g^{k \times k} (X) \right] \right] \quad (3)$$

因此 RFACConv 的计算公式为:

$$F = A_{RF} \times F_{RF} \quad (4)$$

通过将 RFACConv 与主干网络中的 C2f 结合,设计出 C2f_RFACConv 模块。该模块通过对感受野中各个空间特征重要性进行感应,从而在空间方面增强模型特征提取的能力并减少传统卷积因共享参数导致局部特征弱化的情况。在减少参数量和计算量的同时提升了模型的检测精度。

2.3. SIoU 损失函数

选用 SIoU 损失函数对原模型结构进行创新和优化。SIoU 引入向量角度,并将损失细分为:角度损失、距离损失、形状损失以及 IoU 损失。作为一种基于面积的损失函数,SIoU 的取值仅与目标框的面积

紧密相关。这一特性有效规避了边界框动态变化所带来的干扰，提供了更高的稳定性。SIoU 摒弃了复杂繁琐的距离度量方式与除法运算，大幅简化了计算流程。

除此之外，SIoU 能够迅速引导预测框向最近的轴靠近，极大的减少了模型训练时间，更高效地促使预测框朝着真实框的位置精准收敛，从而提升目标检测模型的检测精度。

SIoU 损失函数在计算过程中涉及到的各项参数，如图 4 所示：

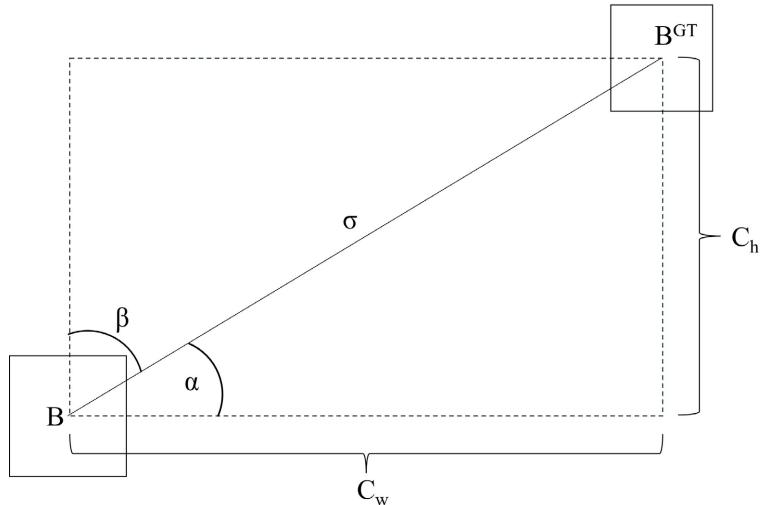


Figure 4. Schematic diagram of SIoU loss computation

图 4. SIoU 损失函数计算示意图

角度损失表示为预测框 B 与真实框 B^{GT} 中心点角度的偏差，其定义如公式(5)所示：

$$A = 1 - 2 \times \sin^2 \left(\arcsin x - \frac{\pi}{4} \right) \quad (5)$$

其中：

$$\begin{cases} x = \frac{c_h}{\sigma} \\ \sigma = \sqrt{(b_{c_x}^{gt} - b_{c_x})^2 + (b_{c_y}^{gt} - b_{c_y})^2} \\ c_h = \max(b_{c_y}^{gt}, b_{c_y}) - \min(b_{c_y}^{gt}, b_{c_y}) \end{cases} \quad (6)$$

C_h 表示真实框中心点和预测框中心点之间的高度差， σ 代表真实框中心点和预测框中心点之间的距离，坐标 $(b_{c_x}^{gt}, b_{c_y}^{gt})$ 指的是真实框的中心，坐标 (b_{c_x}, b_{c_y}) 指的是预测框的中心。

距离损失指的是预测框 B 与真实框 B^{GT} 最小外接距离的差异，距离损失的定义如公式(7)所示：

$$\begin{cases} \Delta = \sum_{i \in x, y} (1 - e^{-(2-\Delta)\rho_i}) \\ \rho_x = \left(\frac{b_{c_x}^{gt} - b_{c_x}}{c_w} \right)^2 \\ \rho_y = \left(\frac{b_{c_y}^{gt} - b_{c_y}}{c_h} \right)^2 \end{cases} \quad (7)$$

C_w 与 C_h 代表真实框与预测框在最小外接矩形上的宽和高。 ρ_x 和 ρ_y 表示 B^{GT} 和 B 中心点的位移差除以最小外接矩阵的宽(高)的平方。

形状损失表示为预测框 B 与真实框 B^{GT} 长宽比的偏差。其定义如公式(8)所示:

$$\Omega = \left(1 - e^{-W_w}\right)^\theta + \left(1 - e^{-W_h}\right)^\theta \quad (8)$$

其中:

$$\begin{cases} \omega_w = \frac{|w - w^{gt}|}{\max(w, w^{gt})} \\ \omega_h = \frac{|h - h^{gt}|}{\max(h, h^{gt})} \end{cases} \quad (9)$$

w 和 h 分别代表预测框的宽和高, w^{gt} 和 h^{gt} 分别表示真实框的宽和高。 θ 为形状损失关注度参数, ω_w 和 ω_h 是预测框相对于真实框在宽度和高度方向上的拉伸程度。

IoU 损失表示为预测框 B 与真实框 B^{GT} 的并集与其交集相除。其定义如公式(10)所示:

$$\text{IoU} = \frac{|b \cap b^{gt}|}{|b \cup b^{gt}|} \quad (10)$$

因此得出 SIoU 损失函数的公式如下:

$$\text{Loss}_{\text{SIoU}} = 1 - \text{IoU} + \frac{\Delta + \Omega}{2} \quad (11)$$

3. 实验结果与分析

3.1. 数据集

本研究以加速器中子源设备作为目标检测对象, 通过在实际运行现场对相关设备进行图像采集, 构建了专用设备检测数据集。为增强数据多样性并提高模型对复杂环境的适应能力, 采用了图像翻转、亮度调节、遮挡模拟以及噪声扰动等数据增强策略, 最终获得共 4328 张加速器设备图像样本。在数据集划分方面, 按照 8:1:1 的比例将样本划分为训练集、验证集和测试集, 分别包含 3462 张、433 张和 433 张图像。结合加速器中子源设备的结构组成及功能特征, 将检测目标划分为 11 类, 为后续目标检测模型的训练与评估提供数据支撑。

3.2. 实验环境

本文所使用的实验环境和配置如表 1 所示:

Table 1. Experimental environment and configuration
表 1. 实验环境及配置

配置	参数
GPU	NVIDIA GeForce RTX 4090
显存	24GiB
操作系统	Ubuntu22.04
开发环境	Python3.9.20 PyTorch2.0.1
训练环境	CUDA 12.6

本实验所用的超参数为：迭代次数设置为 100 轮，输入图片的分辨率设置为 640×640 ，批量大小设置为 8。

3.3. 评价指标

本文采用多个指标评估所提模型对加速器中子源设备的检测性能，具体通过准确度(precision, P)、召回率(recall, R)、平均精度均值(mean average precision, mAP)、参数量(params)、浮点计算量(floating-point operations per second, FLOPs)以及模型大小(Model Size)进行评估，其计算公式如下：

$$\begin{cases} P = \frac{T_p}{T_p + F_p} \times 100\% \\ AP = \int_0^1 P(R) dR \\ mAP = \frac{1}{n} \sum_{i=1}^n \int_0^1 P(R) dR \end{cases} \quad (12)$$

其中， T_p 表示被模型正确识别的样本数， F_p 表示误判样本数。 AP (Average Precision) 用于衡量模型在不同置信度阈值条件下的整体检测性能。

3.4. 消融实验

为验证各改进模块对模型性能的贡献，本文在 YOLOv10n 模型上依次引入 EffectiveSE 注意力机制、C2f_RFACConv 模块以及 SIoU 损失函数，并开展消融实验，结果如表 2 所示。

Table 2. Ablation experiments

表 2. 消融实验

模型	mAP@50/%	P/%	R/%	Params/M	FLOPs/G
YOLOv10n	89.5	85.8	81.4	2.7	8.4
YOLOv10n + EffectiveSE	90.1	83.5	83.4	3.9	14.4
YOLOv10n + EffectiveSE + C2f_RFACConv	90.1	82.6	85.2	3.7	14
YOLOv10n + EffectiveSE + C2f_RFACConv + SIoU (Ours)	93.2	89.3	86.2	3.7	14

在基线模型中引入 EffectiveSE 后，模型的 mAP@50 提升了 0.6%，召回率提高了 2.0%，表明通道注意力机制有助于增强特征表达能力，但同时带来了一定的参数量和计算量增长。在此基础上进一步融合 C2f_RFACConv 模块，模型召回率进一步提升了 1.8%，有效缓解了目标漏检问题，同时模型复杂度相较于仅引入 EffectiveSE 时有所降低，体现了较好的性能与效率权衡。最后，引入 SIoU 损失函数后，模型检测性能获得显著提升，其中 mAP@50 提升了 3.1%，召回率提升了 1.0%，准确度提升了 6.7%。在参数量和计算量基本保持不变的情况下，整体检测精度得到进一步增强，验证了所提出改进策略的有效性。综上所述，各改进模块在不同阶段均对模型性能产生了积极影响，且相互之间具有良好的互补性，使模型在保证计算效率的同时显著提升了检测性能。

3.5. 可视化分析

为直观展示 ECS-YOLO 与 YOLOv10n 在检测性能上的差异，随机选取数据集中的样本图片进行了可视化分析。如图 5 所示，第一列为真实标签，第二列为 YOLOv10n 的检测结果，第三列为改进后的 ECS-YOLO 检测效果。

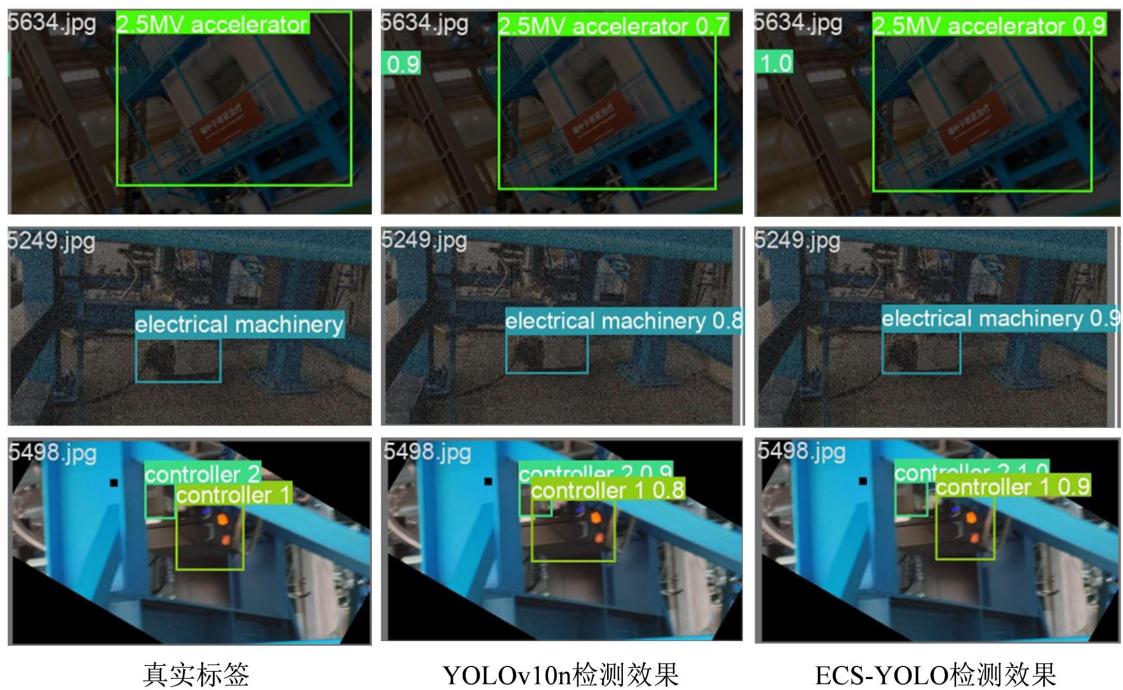


Figure 5. Visualization of detection results
图5. 检测结果可视化

4. 结论

本文针对加速器中子源设备巡检场景中结构复杂、类型多样及遮挡严重等问题，提出了一种改进的YOLOv10目标检测算法：ECS-YOLO。通过引入EffectiveSE注意力机制、设计C2f_RFACConv模块以及采用SIOU损失函数，在不显著增加计算开销的前提下有效提升了特征表达能力和目标定位精度。实验结果表明，与YOLOv10n基线模型相比，ECS-YOLO在mAP@50、精确率和召回率上分别提升了3.7%、3.5%和4.8%，在复杂巡检场景下表现出更优的检测性能与稳定性。该方法为加速器中子源设备的智能巡检与工程应用提供了可靠的技术支持。尽管该方法取得了较好的检测效果，但仍存在一定局限性，如模型训练依赖于特定场景的数据集，跨场景泛化能力和实际部署中的实时性仍有待进一步验证。未来研究将重点围绕模型轻量化与边缘部署展开，并探索引入红外、深度等多模态信息，以进一步提升复杂环境下的检测鲁棒性与工程应用价值。

基金项目

合肥综合性国家科学中心能源研究院(安徽省能源实验室)项目(21KZS202)。

参考文献

- [1] 曾龙军, 卢承方, 陈勇, 等. 基于YOLO-FBA的复杂环境下豆叶病害检测研究[J]. 江苏农业科学, 2025, 53(20): 315-324.
- [2] 熊诗雨, 狄永正, 纪雯, 等. 基于YOLOv11的远距复杂场景小目标检测[J]. 计算机系统应用, 2026, 35(1): 152-163.
- [3] 徐中懿, 张海军, 赵新元. 基于SDO-YOLO的复杂工业场景作业人员行为检测算法[J/OL]. 计算机应用与软件: 1-11. <https://link.cnki.net/urlid/31.1260.tp.20251105.1612.006>, 2026-01-19.
- [4] 郭莉, 张雪松, 李萌萌, 等. 改进YOLOv10的复杂场景人体跌倒检测方法[J/OL]. 电子测量技术: 1-11. <https://link.cnki.net/urlid/11.2175.tn.20251022.1028.012>, 2026-01-19.

-
- [5] 薛光辉, 闫朝阳, 吴冕. PCSED-YOLO: 复杂环境下跨尺度多目标穿戴检测算法研究[J/OL]. 计算机工程与应用: 1-20. <https://link.cnki.net/urlid/11.2127.TP.20251016.1507.016>, 2026-01-19.
 - [6] 周志耀, 马常霞, 杨丽莎, 等. 基于轻量级改进的 YOLOv8 水下目标检测模型[J]. 电子测量技术, 2024, 47(19): 181-189.
 - [7] Wang, A., Chen, H., Liu, L., *et al.* (2024) YOLOv10: Real Time End-To-End Object Detection. arXiv: 2405.14458.
 - [8] 杨长春, 贺轩轩, 王睿, 等. 基于改进 YOLOv8 光伏板缺陷检测算法[J]. 电子测量技术, 2024, 47(23): 181-192.
 - [9] Zhu, W., Han, X., Zhang, K., Lin, S. and Jin, J. (2025) Application of YOLO11 Model with Spatial Pyramid Dilation Convolution (SPD-Conv) and Effective Squeeze-Excitation (EffectiveSE) Fusion in Rail Track Defect Detection. *Sensors*, **25**, Article 2371. <https://doi.org/10.3390/s25082371>
 - [10] Shi, J., Jia, Y., Zhou, G., Wang, J. and Jia, Z. (2025) Small Target Insect Detection Based on Improved YOLOv8n. *ICASSP 2025—2025 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Hyderabad, 6-11 April 2025, 1-5. <https://doi.org/10.1109/icassp49660.2025.10890801>
 - [11] Gu, Z., Zhu, K. and You, S. (2023) YOLO-SSFS: A Method Combining SPD-Conv/STDL/IM-FPN/SIoU for Outdoor Small Target Vehicle Detection. *Electronics*, **12**, Article 3744. <https://doi.org/10.3390/electronics12183744>
 - [12] Hu, J., Shen, L. and Sun, G. (2018) Squeeze-And-Excitation Networks. 2018 *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City, 18-23 June 2018, 7132-7141. <https://doi.org/10.1109/cvpr.2018.00745>