

DOI:10.19651/j.cnki.emt.2519136

SGAD-YOLO: 基于改进 YOLO11 的工作服规范 穿戴检测算法*

杜俊男 杨 雯 刘志龙 王 成 王天一

(贵州大学大数据与信息工程学院 贵阳 550025)

摘要: 针对现有目标检测网络在工业场景中对规范穿戴工作服检测存在精确度不足、模型复杂和泛化性较低等问题,提出了一种基于 YOLO11 的高精度、轻量化新模型 SGAD-YOLO。首先,结合 StripBlock 结构和 CGLU 机制改进 C3k2 模块,通过多层次特征处理和动态特征增强,提升模型对细长特征和复杂纹理的感知能力,同时降低模型的参数量和计算量;其次,引入 AFGCAttention 注意力机制,通过全局上下文信息与局部特征的动态融合,增强模型对关键区域的聚焦能力,有效抑制背景噪声干扰;最后,重设计 Detect-SEAM 检测头,以提高模型在复杂环境下对遮挡物体和小目标物体的检测精度。实验结果表明,改进算法在电网现场作业环境数据集和公开数据集 Roboflow 5 中的 mAP@0.5 指标分别达到 93.6% 和 94.6%,相较于基线模型分别提升了 1.5% 和 2.1%,并且其参数数量和计算量分别下降了 8.3% 和 7.4%,证明了 SGAD-YOLO 算法在工业场景中对规范穿戴工作服检测任务具备更好的检测性能。

关键词: 工业场景;目标检测;YOLO11;轻量化改进;注意力机制

中图分类号: TP391.41;TN911.73 **文献标识码:** A **国家标准学科分类代码:** 510.4050

SGAD-YOLO: A standardized work uniform detection algorithm based on improved YOLO11

Du Junnan Yang Wen Liu Zhilong Wang Cheng Wang Tianyi

(Institute of Big Data and Information Engineering, Guizhou University, Guiyang 550025, China)

Abstract: To address the issues of inadequate accuracy, complex model architecture, and poor generalization in detecting standardized work uniform within industrial scenarios using existing object detection networks, a novel high-precision lightweight model named SGAD-YOLO based on YOLO11 is proposed. First, the C3k2 module is improved by combining the StripBlock structure and CGLU mechanism. Through multi-level feature processing and dynamic feature enhancement, the model's perception of slender features and complex textures is improved, while the model's parameters and computational complexity are reduced. Second, the AFGCAttention mechanism is introduced to enhance the model's focus on key regions and effectively suppress background noise interference through the dynamic fusion of global context information and local features. Finally, the Detect-SEAM detection head is redesigned to improve the model's detection accuracy for occluded and small objects in complex environments. Experimental results demonstrate that the improved algorithm achieves mAP@0.5 values of 93.6% and 94.6% on the power grid field operation dataset and the public Roboflow 5 dataset, respectively—representing improvements of 1.5% and 2.1% over the baseline model. Moreover, its parameters and computational complexity are reduced by 8.3% and 7.4%, respectively. This proves that the SGAD-YOLO algorithm has better detection performance for standardized work uniform detection tasks in industrial scenarios.

Keywords: industrial scenario; object detection; YOLO11; lightweight improvement; attention mechanism

0 引 言

在工业作业场景,尤其是变电站、化工厂这类高危场景

中,规范的穿戴工作服能够确保作业人员的安全,同时能有效地降低事故风险。传统检测规范穿戴工作服主要依靠人工巡检,以广东电网公司为例,为了给户外操作的作业

收稿日期:2025-06-18

* 基金项目:国家自然科学基金(62361009)、贵州省科技计划项目(黔科合基础-ZK[2021]304)资助

人员提供一份安全屏障,规定作业人员每天需要全身穿着工作服,并由监护人员监督。然而,在大型工业场地中受人力资源限制,这种检测方式导致工作人员很难全面、及时地覆盖所有区域进行检测,致使一些未规范穿戴的情况无法被及时察觉,给作业人员带来安全隐患。

随着计算机视觉技术的快速发展,基于深度学习的目标检测算法为规范穿戴工作服检测提供了新的解决方案。目前,在计算机视觉领域,相关研究方向丰富多样。一种是基于机器学习的方法对服装衣物识别分类,依赖人工设计特征结合传统分类器。例如,Hu等^[1]提出先将工人着装图像按安全帽、上身和下身进行结构划分,再分别提取各部分的方向梯度直方图(histogram of oriented gradient, HOG)和颜色直方图(histogram of color, HOC)特征,结合支持向量机(support vector machine, SVM)分类器进行检测,在混合场景下的检测准确率达到86.9%,但由于该算法场景特定性强,迁移到不同工业场景时需手动调整特征提取逻辑,因此难以适应工业场景中作业人员多样的穿戴情况,导致模型泛化性差。另外一种是基于深度学习的目标检测算法,以区域神经网络(region-convolution neural network, R-CNN)^[2]为代表通过区域建议网络生成候选框后处理的两阶段目标检测算法和YOLO(you only look once, YOLO)^[3]为代表通过单次卷积网络直接预测边界框与类别的单阶段目标检测算法。例如,孙国栋等^[4]通过将Faster R-CNN^[5]网络融合自注意力机制,以捕获多尺度全局信息,增强高层语义特征与感受野,在区域建议网络的训练中通过锚框补选增强,用于提高对工作人员安全帽的检测精度,该算法检测的平均精度达到94.3%,但该网络存在参数数量过大的问题(超过137 million)、这导致其计算成本高,检测速度较慢,难以满足工业场景实时性需求。宋春宁等^[6]针对复杂环境下的安全帽检测问题,提出改进YOLOv5算法,通过设计CAKConv卷积模块及引入EMA注意力机制,为工业场景下的防护装备检测提供了轻量化解决方案,但该算法在复杂环境下会导致图像特征严重失真。Li等^[7]在YOLOv5s算法的基础上,使用轻量级网络ShuffleNet v2^[8]替换骨干网络,以减少模型参数数量,提高模型在低计算能力平台上的运行速度,但该算法的检测精度仅为87.7%,模型检测精度有待提高。邝先验等^[9]基于YOLOv7-tiny设计了CDA-YOLO算法,通过低跨度上下文解耦检测头和并行级联注意力模块提升特征提取能力,为轻量化模型在复杂场景的应用提供了思路。该算法主要针对遥感图像优化,其解耦头设计在工业场景中工作服等柔性目标的边缘细节捕捉上存在局限。张羽等^[10]提出改进的YOLOv8轻量化检测算法,通过重构颈部网络,并引入EMA坐标注意力机制改进C2f模块,模型对反光衣和安全帽的检测精度提升了0.1%、参数数量下降了48.3%、计算量下降了32.1%,提升了算力平台运行速度,但导致特征提取能力下降,在复杂环境下模型鲁棒性不足,且不同工

业场景的工作服样式、检测需求存在差异,模型泛化能力受限。

综上所述,三类计算机视觉检测方法均依托特定的特征提取框架开展检测工作。其中,相比场景特定性强的机器学习检测算法与模型复杂、计算成本高的两阶段深度学习检测算法,YOLO系列目标检测算法凭借其高效的处理效率与实时响应能力,在各类实时目标检测任务中得到广泛应用。然而,YOLO系列检测算法在实际处理工业场景中仍面临以下挑战:工作服中的裤腿、袖口等细长特征在图像中占比小且形状特殊,导致传统卷积核难以有效捕捉这类特征;工业现场环境复杂,光照变化、遮挡以及背景杂乱等问题严重影响了检测精度;工业监控需要满足高实时性,以能达到实时监督目的,而现有模型往往在精度与速度之间难以取得平衡等。这些问题限制了现有算法在工业场景中的实际应用效果,因此需要一种能够兼顾精度、速度和鲁棒性的解决方案,以提升工业生产安全性、推动行业技术进步。

针对YOLO模型在处理复杂工业环境中存在的不足,主要做出以下工作:1)设计C3k2-StripCGLU模块。通过在C3k2模块中嵌入长条状卷积核,强化细长特征捕捉能力的同时,显著降低了模型参数量与计算开销;同时,融合卷积门控线性单元(convolutional gated linear unit, CGLU)^[11]机制,进一步提升了模块对复杂纹理与细节的感知能力。2)在特征金字塔主干网络的浅层部分嵌入自适应细粒度通道注意力(adaptive feature global context attention, AFGCAAttention)^[12]模块。通过动态融合全局上下文信息与局部特征,增强了模型对关键区域的聚焦能力,有效抑制背景噪声干扰,从而提升整体检测精度。3)提出Detect-SEAM检测头,集成了空间增强注意力模块(spatial enhanced attention module, SEAM)^[13]。通过空间维度的注意力权重分配机制,该检测头强化关键区域的特征响应,克服了传统检测头在复杂环境下的性能瓶颈

1 YOLO11 算法

YOLO系列算法属于一阶段目标检测算法,该算法在输入图像上运行一次卷积神经网络就可以预测多个边界框和它们的类别概率。由于其算法的特性,使得YOLO系列算法能够满足如工业监控场景这类对实时性要求较高的应用场景。其中,YOLO11为最新版本的YOLO算法,它在不同版本YOLO的基础上进行了显著的架构和训练方法改进,旨在提升目标检测的精度、速度和效率。YOLO11继承了YOLOv8模型设计长处的同时,在工程实践维度着重优化。

YOLO11网络结构分为主干网络、颈部网络和检测层三部分。主干网络设计方面,YOLO11借鉴了跨阶段局部(cross stage partial, CPS)模块思想,将YOLOv8中的C2f模块替换成了C3k2模块,新模块通过优化特征提取路径,减少了计算成本并提升了特征表达能力。此外,YOLO11

沿用了 YOLOv8 中的空间金字塔池化 (spatial pyramid pooling fast, SPPF) 模块,并在主干网络的最后一层嵌入了 CPS 模块和金字塔挤压注意力 (pyramid squeeze attention, PSA) 模块,以增强模型对关键区域的空间感知能力,从而提升检测精度。颈部网络方面,使用路径聚合网络-特征金字塔网络 (path aggregation network-feature pyramid network, PAN-FPN)^[14-15] 结构来构建特征金字塔。检测层方面,采用 Decoupled-Head 解耦头结构,如同 YOLOv8 一样将分类和检测头分离,但 YOLO11 在分类检测头中增加了两个深度可分离卷积,进一步减少了参数量和计算量。在损失函数的选择上, YOLO11 模型中使用完整交并比 (complete intersection over union, CIoU)^[16] 与分布 focal 损失 (distribution focal loss, DFL)^[17] 作为边界框回归的损失函数, CIoU 在 IoU^[18] 的基础上增加了对边界框中心点距离

和长宽比等几何特性的考量,使得损失函数能够更全面地反映预测框与真实框之间的差异。DFL 通过优化边界框回归的分布差异,提高了模型对边界框位置不确定性的建模能力,增强了模型在复杂场景下的鲁棒性。YOLO11 基于缩放系数提供了 n/s/m/x 尺度不同大小的模型,用于满足不同场景需求。相较而言,模型尺度较小的 YOLO11 s 在保证一定检测性能的同时,更契合工业实时监控场景下对资源高效利用以及快速响应的实际需求。因此,本文选定 YOLO11 s 作为基础模型。

2 SGAD-YOLO 算法

针对 YOLO11 s 模型在处理复杂工业环境中存在的不足,基于 YOLO11 s 模型提出 SGAD-YOLO 检测算法。SGAD-YOLO 算法的整体架构如图 1 所示。

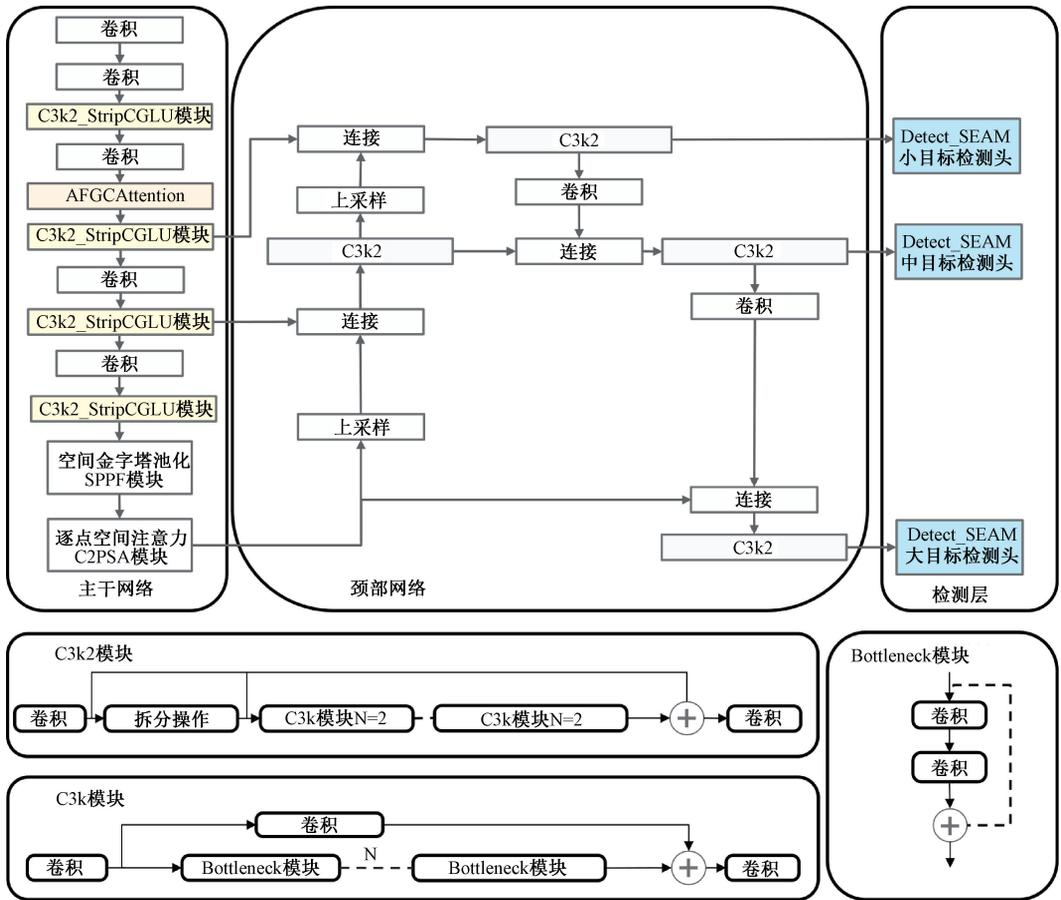


图 1 SGAD-YOLO 结构图
Fig. 1 The structure of the SGAD-YOLO

针对 SGAD-YOLO 算法做出的 3 个改进,下文将依次详细阐述 C3k2-StripCGLU 模块、AFGCAttention 注意力机制以及 Detect-SEAM 检测头的技术细节与实现原理。

2.1 C3k2-StripCGLU 模块

由于标准卷积在提取语义信息丰富的特征图中,会导致小目标特征发生显著的改变或直接丢失^[19],从而导致模

型的检测精度下降。如图 1 所示, C3k2 模块提取特征主要是通过标准卷积,导致其在处理工作服中丰富的细长特征时,难以有效捕捉其局部细节信息。为克服基础模型在下采样中存在的局限性,本文设计了 C3k2-StripCGLU 模块,模块结构如图 2 所示。

在 C3k2-StripCGLU 模块中, C3k-StripCGLU 模块是

通过结合 StripBlock^[20] 结构思想和卷积门控线性单元机制 所设计的新模块。C3k-StripCGLU 模块结构如图3所示。

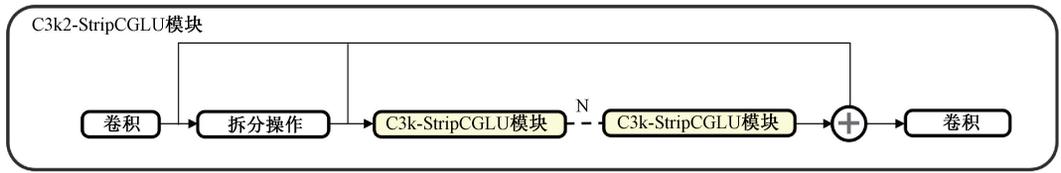


图2 C3k2-StripCGLU 模块
Fig. 2 The C3k2-StripCGLU module

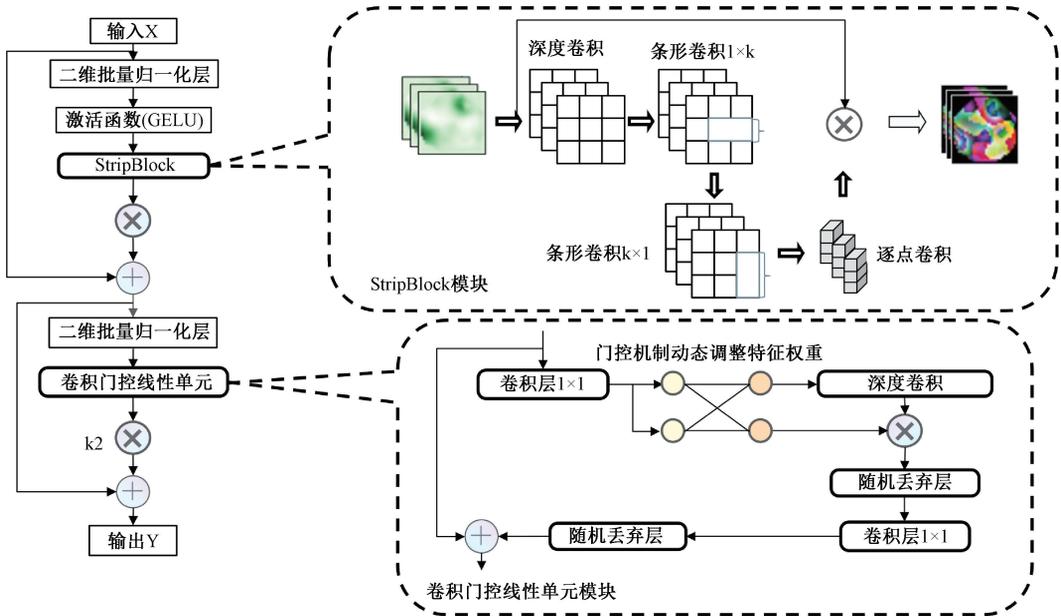


图3 C3k-StripCGLU 模块
Fig. 3 The C3k-StripCGLU module

由图3所示,模块在特征处理流程中,构建了一套层次化、多维度的优化机制:首先,输入特征图首先经过归一化处理,引入长条状卷积核打破传统卷积核的几何限制,针对性地捕捉细长形态特征,通过自适应注意力机制精准聚焦关键区域,在强化有效特征表达的同时,降低模型参数量与计算复杂度。随后,输出特征通过残差连接^[21]和动态权重调整,增强了对细长特征的捕捉能力。最后,经过细长条注意力模块处理后,进入卷积门控线性单元模块,对归一化后的特征进行非线性变换,通过门控机制动态调整特征权重,并引入随机丢弃层增加模型的鲁棒性,最终实现模型整体对复杂纹理感知能力的提升。

C3k-StripCGLU 模块通过窄带卷积核设计,有效捕获图像水平和垂直方向的长距离依赖关系,突破了传统方形卷积核和池化核的空间限制;并通过门控机制自动学习特征图各通道的重要性权重,有效缓解深层网络中的梯度消失问题,增强模型对复杂特征的建模能力。两种机制的有机结合,有效地解决原 C3k 模块中标准卷积对于丰富语义特征提取的局限性,兼具局部细节捕获和长距离特征关联能力,在轻量化的前提下提升特征表达能力。使整体算法

在工业场景下工作服缺陷检测、合规性判定等任务中,具备更快的推理速度与更强的环境适应性。

2.2 AFGCAAttention 注意力机制

Sun 等^[12]认为 SE 通道注意力^[22]机制广泛应用且有一定成效,但它仅依赖全连接层捕捉全局信息,缺乏与局部信息的有效交互,导致图像去雾时特征权重分配不准确。为实现全局和局部信息的有效整合与合理权重分配,提出 AFGCAAttention 模块。AFGCAAttention 模块结构如图4所示。

给定输入特征图为 $x \in R^{C \times H \times W}$,其中 C 是通道数、 H 是高度、 W 是宽度。模块对特征提取数学表达式如下:

$$U_n = GAP(F_n) = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W F_n(i, j) \quad (1)$$

$$\begin{cases} U_{lc} = \sum_{i=1}^k U \cdot b_i \\ U_{gc} = \sum_{i=1}^c U \cdot d_i \end{cases} \quad (2)$$

$$M = U_{gc} \cdot U_{lc}^T \quad (3)$$

$$W = \sigma(\sigma(\theta) \times \sigma(U_{gc}^w) + (1 - \sigma(\theta)) \times \sigma(U_{lc}^w)) \quad (4)$$

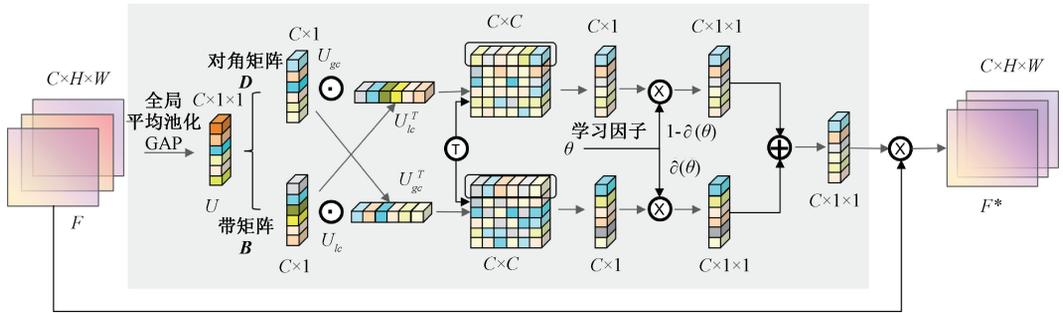


图 4 AFGCAttention 模块工作机制流程

Fig. 4 The AFGCAttention module workflow

该模块的结构设计基于对特征交互的深度优化:首先,模块利用带矩阵操作实现局部通道交互,深入挖掘特征图局部区域内通道之间的潜在联系。其次,结合对角矩阵运算实现全局信息整合,将全局层面的特征信息融入到特征处理过程中,使得模块能够动态地捕捉不同粒度特征之间的关联,针对性地强化浅层特征中的边缘和纹理通道。最后,通过交叉相关运算增强全局与局部信息的动态关联性,在复杂背景干扰下自适应抑制无关通道。模块融合全局平均池化和通道交叉相关操作,构建了全局语义引导下的局部细节增强机制。

YOLO11s 网络基于多尺度特征金字塔进行特征提取,特征融合方式通过简单的拼接操作融合不同尺度特征。这种传统融合策略在浅层网络中,由于语义信息匮乏,难以形成对目标的完整语义表达。同时,浅层网络缺乏动态注意力机制,无法自适应地判别各尺度特征的重要程度,这使得模型在处理小目标时,难以精准捕捉其细微

纹理、轮廓等关键特征。然而,传统通道注意力仅关注全局信息,同样难以捕捉小目标的局部纹理和边缘特征。针对以上问题,需一种适配浅层网络的注意力机制,在保留高分辨率空间细节的同时,弥补基础网络在小目标特征提取方面的缺陷。

有鉴于此,本文将 AFGCAttention 模块嵌入特征提取金字塔的浅层网络,旨在充分发挥其在细粒度特征聚焦与动态权重分配上的优势,有效弥补基础网络的短板。通过这一改进,YOLO11s 浅层网络在特征提取时能够保留高分辨率空间细节,克服基础网络对目标特征提取的不足和传统注意力机制的局限性,从而提高网络对目标的检测能力。

2.3 Detect-SEAM

SEAM 模块是一种针对遮挡和小目标优化的注意力机制,其核心是通过多尺度特征融合与空间增强,提升模型对关键区域的关注能力。其网络结构如图 5 所示。

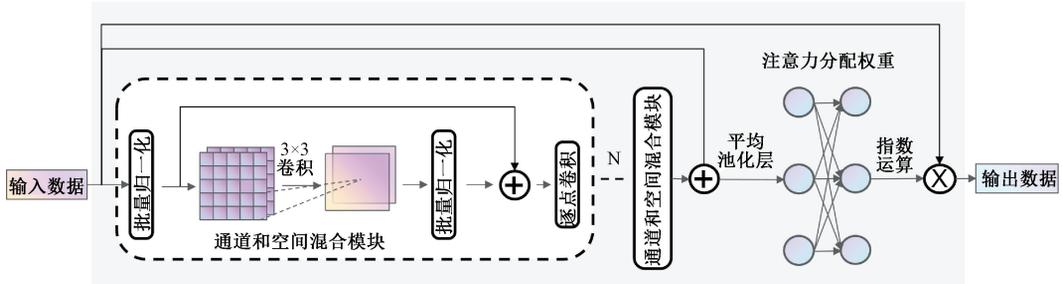


图 5 SEAM 模块结构

Fig. 5 The structure of the SEAM module

YOLO11s 原检测探头受限于固有设计机制,导致在复杂环境下检测时存在以下缺陷:常规卷积操作的固定感受野与静态特征融合策略,导致模型对遮挡区域的局部语义恢复能力不足;对关键细节,如被遮挡的衣领、挽裤腿等,因缺乏空间注意力引导而难以重建,加剧了不完整目标的漏检和错检风险。

而 SEAM 模块作为针对遮挡和小目标优化的注意力机制,通过多尺度特征融合与空间增强技术,可提升模型对关键区域的聚焦能力。该模块构建的协同互补特征提

取体系中,通道和空间混合模块先利用深度卷积强化局部空间特征提取,捕捉遮挡区域细节,再通过逐点卷积实现跨通道信息融合,精准保留反光条等小目标特征;多级深度可分离卷积路径的堆叠设计渐进式增强遮挡特征表达;全局平均池化结合瓶颈全连接结构动态生成通道注意力权重,有效抑制背景干扰;残差连接则在特征提取过程中保留原始特征上下文,辅助恢复遮挡区域语义完整性。基于 SEAM 模块的上述优势,本文提出 Detect-SEAM 检测头。模块结构如图 6 所示。

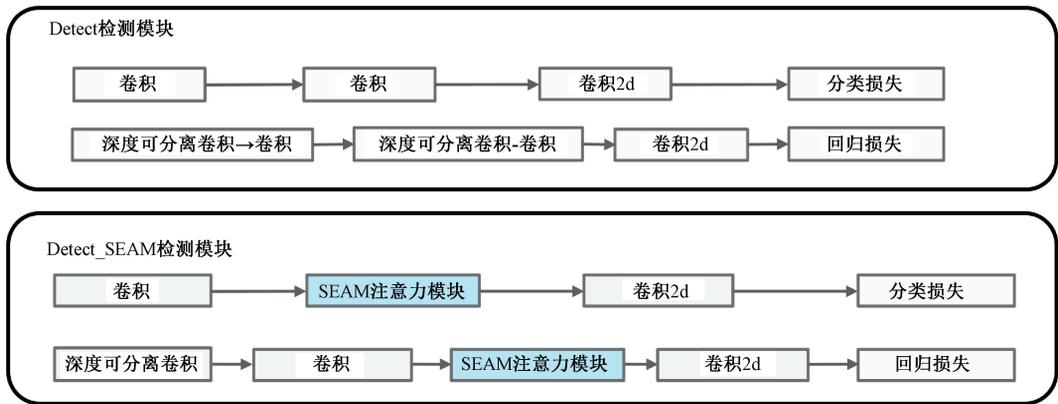


图6 Detect检测头和 Detect-SEAM检测头

Fig. 6 The Detect detection head and the Detect-SEAM detection head

Detect-SEAM 检测探头,通过通道与空间混合模块,捕捉遮挡区域底层特征并跨通道融合,重组碎片化特征以提升语义恢复完整性。利用多级深度可分离卷积渐进增强小目标特征响应,结合通道注意力抑制背景噪声,突出小目标细节。残差连接保留原始特征上下文,避免结构失真,增强对不完整目标的鲁棒性。

该检测头在维持模型计算量和参数量不变的前提下,能够有效克服原检测头在遮挡敏感性与小目标定位方面的核心难题。

3 实验结果与分析

3.1 实验环境和数据集

本文所有实验均在同一实验环境下进行,实验环境的操作系统为 Windows 22 621.525,CPU 型号为 13th Gen Intel Core i5-13400F,GPU 型号为 NVIDIA GeForce RTX 4060Ti,运行内存 32 GB,编程语言为 Python3.8.18,采用深度学习框架 PyTorch2.3.0 训练,CUDA 版本为 11.8。本文采用的数据集源自阿里巴巴天池挑战赛,针对广东电网智慧现场作业环境,专注于识别作业人员的规范穿戴工作服情况。采集时涵盖了多种天气条件,包括晴天、阴天及雾天,以模拟工业场景中可能遇到的各类环境光变化。为了构建高效且准确的实验环境,本文对原始数据集数据清洗,最终保留了 2 940 张高质量图像作为实验数据集。每张图片通过图像标注工具 labelimg 进行标注,内容覆盖了电力作业现场常见的四大类样本标签监护袖章(badge)、图中出现的所有在场人员(person)、合规穿戴工作服(clothes)、不合规穿戴工作服(wrongclothes),不合规穿戴工作服包括上衣开襟、挽裤腿、挽袖、不成套等。部分数据集如图 7 所示。

为了提供丰富的上下文信息和精准度,增强系统对复杂工作环境中各元素的识别能力,使得监控系统能够更精确地评估工作人员的安全状态、作业流程的合规性以及潜在的安全风险。本文采用随机划分的方式将该数据集按



图7 电力作业现场数据集

Fig. 7 The dataset of electric power operation site

照 7 : 2 : 1 的比例分成训练集、测试集和验证集。并且数据集训练参数配置表如表 1 所示。

表 1 训练参数配置表

Table 1 Training parameter configuration table

名称	参数
优化器(optimizer)	SGD
动量值(momentum)	0.937
迭代次数(epoch)	300
批量尺寸(batch_size)	8
学习率(learning_rate)	0.01
输入尺寸(img_size)	640×640

为了验证模型优化方法的有效性,设置了消融实验和对比实验,以探究提出的改进方法对 YOLO11 s 模型的性能影响。

3.2 评价指标

为了全面而客观地评估模型对该特定场景下的检测性能,本文采用了以下评价指标:平均精度均值(mean average precision, mAP)、精确度(precision, P)、召回率(recall, R)、模型参数量、浮点计算量(GFLOPS)与每秒帧数(FPS)。mAP@0.5 表示 IoU 阈值设为 0.5 时的平均精

准确度, mAP 越大则代表模型整体的精确度越高。精确率是指被模型预测为正类的样本中, 实际也为正类的比例。召回率是指所有实际为正类的样本中, 被模型正确预测为正类的比例, 每秒帧数是衡量模型推断速度的指标, 表示在 1 s 内处理的图像帧数。计算公式如下:

$$P = \frac{TP}{TP + FP} \quad (5)$$

$$R = \frac{TP}{TP + FN} \quad (6)$$

$$mAP = \frac{1}{N} \sum_{i=0}^n \int_0^1 iPRdr = \frac{1}{N} \sum_{i=0}^n AP_i \quad (7)$$

式中: TP 表示模型预测为正类且实际也为正类的样本数, FP 表示预测为正类但实际为负类的样本数, FN 表示预测为负类但实际为正类的样本数, N 表示类别的总数。

3.3 对比实验

本对比实验旨在系统性地比较几种先进的目标检测算法, 在相同硬件配置和软件环境下进行训练与测试, 以评估它们的性能差异。参与对比的算法包括 Faster-RCNN、YOLOv8s、YOLOv9s^[23] 和 YOLOv10s^[24] 算法。同时引入一些轻量化优化算法。实验结果如表 2 所示。

表 2 其他目标检测模型对比

Table 2 Comparison with other object detection models

算法	mAP@0.5/%	精确度 P/%	召回率 R/%	计算量/GFLOPS	参数量/m	FPS
Faster-RCNN	84.2	83.2	84.3	—	137.1	6.8
YOLOv8s	91.6	89.2	82.6	28.7	11.1	93.1
YOLOv9s	91.8	88.9	81.9	27.4	7.3	98.6
YOLOv10s	91.7	89.1	82.1	24.8	8.0	102.7
YOLO11s	92.1	90.9	86.5	21.6	9.5	108.5
YOLO11s-LADH	91.0	89.6	84.8	17.5	7.9	128.2
YOLO11s-EfficientHead	91.3	89.4	85.3	19.7	9.0	115.6
SGAD-YOLO	93.6	92.8	86.9	19.8	8.8	120.8

由实验结果可知, 在模型经过优化后, SGAD-YOLO 相较于基础模型 YOLO11 s 平均精度和精确度分别提升了 1.5% 和 1.9%, 其参数量和计算量下降了 8.3% 和 7.4%。YOLO11 s-LADH 和 YOLO11 s-EfficientHead 分别是通过 Lightweight Asymmetric Detection Head^[25] 和 EfficientDet^[26] 网络对 YOLO11 s 检测探头进行优化的轻量化算法。YOLO11 s-LADH 算法的计算量和参数量分别下降 19.0% 和 16.8%。YOLO11 s-EfficientHead 法的计算量和参数量分别下降 8.8% 和 5.3%, 但在 mAP@0.5 上分别下降了 1.1% 和 0.8%。综上所述, SGAD-YOLO

展现出更优的综合性能优势。实验证明, SGAD-YOLO 很好地平衡了模型轻量化和算法性能, 其中 FPS 指标显示出该模型满足实时安全监控的需求。

3.4 消融实验

为了探究 SGAD-YOLO 模型中所提出的改进方法对 YOLO11 s 模型的性能影响, 在相同的数据集和训练参数下设计了消融实验。消融实验结果如表 3 所示。

实验结果显示, SGAD-YOLO 在对各个模块进行优化后, mAP@0.5、精确度和召回率指标上均得到了提升。同时, 该模型的参数量与计算量分别得到下降, 有效平衡了

表 3 消融实验

Table 3 Ablation experiment

算法	mAP@0.5	精确度	召回率	计算量	参数量
	/%	P/%	R/%	/GFLOPS	/m
YOLO11s(基线模型)	92.1	90.9	86.5	21.5	9.4
YOLO11s+C3k2f-StripCGLU	92.9	91.6	87.6	20.3	8.9
YOLO11s+AFGCAttention	92.6	91.3	87.8	21.6	9.4
YOLO11s+Detect-SEAM	91.9	90.5	86.6	21.0	9.3
YOLO11s+C3k2f-StripCGLU+AFGCAttention	93.2	92.1	87.9	20.3	8.9
YOLO11s+C3k2f-StripCGLU+Detect-SEAM	92.7	91.0	86.4	19.8	8.8
YOLO11s+AFGCAttention+Detect-SEAM	93.3	91.5	87.8	20.8	9.3
YOLO11s+C3k2f-StripCGLU+AFGCAttention+Detect-SEAM	93.6	92.8	86.9	19.8	8.8

检测精度与计算效率。通过对各改进模块的独立验证与组合测试,结果表明每个优化模块均对模型性能提升产生影响,验证了本文所提出改进方法的有效性。

3.5 检测效果分析

为直观呈现出模型对图像中关键区域的注意力分布及预测结果的空间映射特征。在测试集中随机选取样本对其特征进行可视化处理,制作了模型优化前后的热力图对比分析,热力图分布如图 8 所示。

由热力图分布表明,对于工作服的细节特征,SGAD-YOLO 热力图在裤腿、袖口等关键部位形成显著峰值,而原模型仅能模糊覆盖服装整体区域,缺乏对纹理差异的敏

感度,反映出优化后模型对细粒度特征的捕捉能力更强。在背景抑制方面,SGAD-YOLO 对非目标区域呈现均匀的冷色调,原模型则在背景中存在零星高激活点,表明优化后模型通过增强目标背景的语义区分能力,有效过滤了无关干扰。总体而言,SGAD-YOLO 通过更合理的激活分布,直观反映出其在特征提取、目标定位及语义理解层面的综合提升。

为进一步验证优化方法在工业场景下对规范穿戴工作服检测性能的提升效果,直观的对比 SGAD-YOLO 算法与 YOLO11s 算法的检测效果,将两种算法用于电力人员工作的场景进行预测,预测部分结果如图 9 所示。

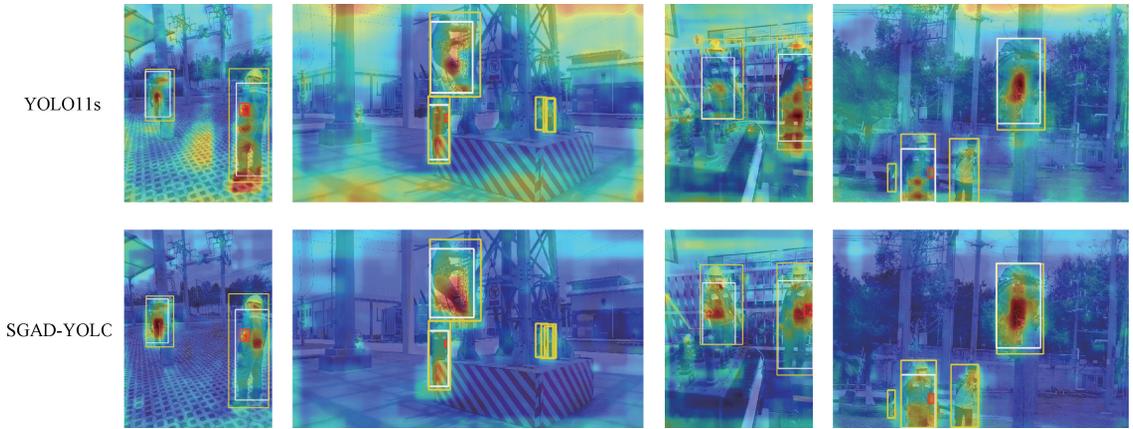
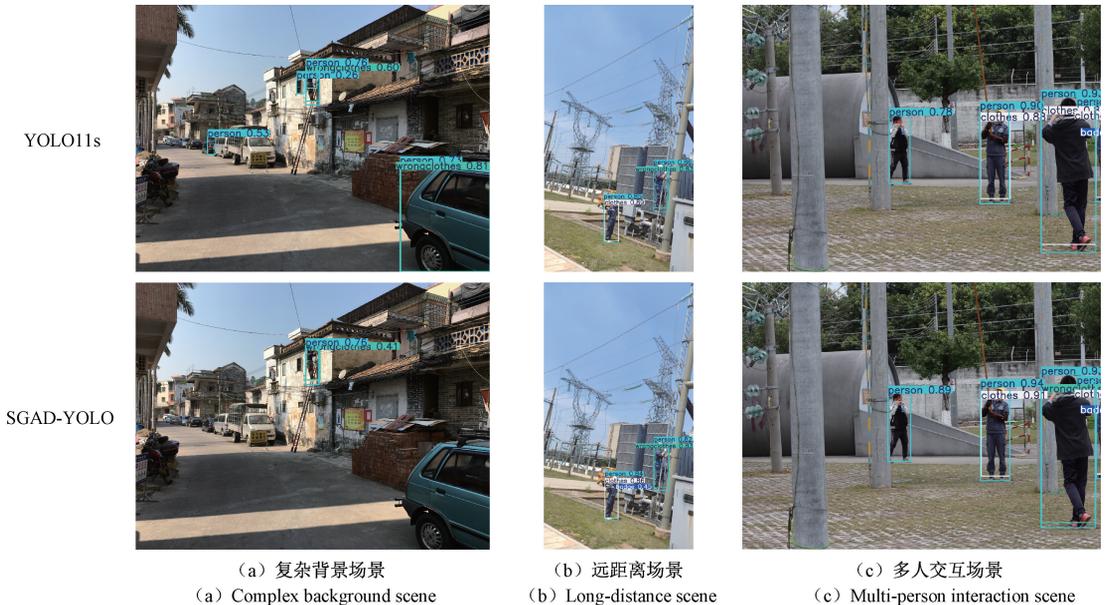


图 8 检测热力对比分析图

Fig. 8 Comparative analysis of detection heat maps



(a) 复杂背景场景

(a) Complex background scene

(b) 远距离场景

(b) Long-distance scene

(c) 多人交互场景

(c) Multi-person interaction scene

图 9 YOLO11s 和 SGAD-YOLO 在不同场景下的检测效果

Fig. 9 Detection performance of YOLO11s and SGAD-YOLO in diverse scenarios

如上图检测效果所示,图 9(a)在复杂背景的检测场景下,基础模型因环境干扰错误识别无关标签,而 SGAD-

YOLO 精准定位工作人员及其不合规穿戴状态,直观体现改进模型在背景抑制能力上的显著提升。图 9(b)在远距

离检测场景下,相较于基础模型出现的目标漏检问题,SGAD-YOLO 通过优化特征提取网络,实现了对远距离目标的完整捕捉,有效避免了因距离导致的检测盲区。图 9(c)在多人交互的复杂检测场景下,在多目标重叠、动作遮挡等干扰下,SGAD-YOLO 凭借增强的注意力机制与多尺度特征融合策略,对每个目标的定位与分类精度均超越原模型,验证了其在动态复杂工作环境中更强的适应性与检测鲁棒性。综上所述,SGAD-YOLO 在工业场景常见的复杂光照、多目标重叠等挑战性环境中,检测准确率与稳定性均展现出显著优势,为工作服穿戴规范的智能化监控提供了更可靠的技术方案。

为系统验证 SGAD-YOLO 在复杂的工业场景中相较于基础模型的检测性能优势,本文选择了 Roboflow 5 数据集进行实验验证,该数据集聚焦建筑作业场景,覆盖了常见的建筑作业现场常见的五大类样本标签,分别包括佩戴头盔、未佩戴头盔、未穿戴工作服、穿戴工作服、图中出现的在场人员,实验结果如表 4 所示。

表 4 Roboflow 5 数据集检测效果对比
Table 4 Comparison of detection performance on the Roboflow 5 dataset

算法	mAP@0.5 /%	精确度 P/%	召回率 R/%	计算量 /GFLOPS	参数量 /m
YOLO11s	92.5	92.6	88.5	21.6	9.5
SGAD-YOLO	94.6	93.8	88.6	19.8	8.8

实验结果表明,SGAD-YOLO 相较于基础模型,mAP@0.5 从 92.5%提高到了 94.6%,精确度从 92.6%提升到 93.8%。实验验证了 SGAD-YOLO 应对不同工业领域复杂环境的强大适应能力。为工业安全检测领域提供了更优的方案。

4 结 论

为提升工业生产安全性,本文聚焦于工业场景中规范穿戴工作服这一关键安全要点,针对 YOLO11s 算法在复杂工况下的检测不足,提出改进算法 SGAD-YOLO。传统标准卷积在语义特征提取中存在局限性,为此引入 C3k2 改进模块,将长条状卷积结构与卷积门控线性单元机制深度融合,增强模型对工作服复杂纹理和空间特征的解析能力,提高工业场景的适应性。浅层网络中,由于语义信息匮乏,难以形成对目标的完整语义表达,为此引入 AFGCAttention 注意力机制,使浅层特征在保留高分辨率空间细节的同时,通过局部通道交互挖掘潜在语义关联,有效弥补了浅层语义信息的缺失,提升模型对小目标的表征能力。结合 SEAM 模块对检测探头进行优化,增强模型捕捉遮挡区域细节信息的能力,进一步提升模型在复杂环境下检测遮挡物和小目标的能力。实验结果表明,无论是建筑施工环境,还是电网作业的特殊工况,SGAD-YOLO

均展现出稳定且高效的检测表现,在轻量化和算法性能平衡上具有优秀的表现。

未来将在该模型的基础上继续进行优化,为进一步拓展模型泛化能力,构建覆盖多行业的工业安全检测数据集,解决数据不平衡问题,并将算法应用延伸至石油化工、煤矿开采等高危领域。在性能优化上,结合边缘计算设备特性对模型进行轻量化改造,开发加速部署方案,探索多任务联合检测提升安全指标识别效率。

参考文献

- [1] HU J L, LIN X B, LI C H, et al. Protective clothing detection of substation workers using S-HOG + C operator[C]. 2018 International Conference on Power System Technology (POWERCON). Guangzhou: IEEE, 2018: 4181-4185.
- [2] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2014: 580-587.
- [3] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2016: 779-788.
- [4] 孙国栋,李超,张航.融合自注意力机制的安全帽佩戴检测方法[J].计算机工程与应用,2022,58(20):300-304.
SUN G D, LI CH, ZHANG H. Safety helmet wearing detection method fused with self-attention mechanism[J]. Computer Engineering and Applications, 2022, 58(20): 300-304.
- [5] 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.
- [6] 宋春宁,李寅中.面向复杂环境的改进 YOLOv5 安全帽检测算法[J].电子测量技术,2025,48(7):163-170.
SONG CH N, LI Y ZH. Improved YOLOv5 safety helmet detection algorithm for complex environments[J]. Electronic Measurement Technology, 2025, 48(7): 163-170.
- [7] LI W, WEI Y. A lightweight YOLOv5 model used for safety helmet and reflective clothing detection[C]. 2022 2nd International Conference on Algorithms, High Performance Computing and Artificial Intelligence (AHPICAI). Guangzhou: IEEE, 2022: 582-587.
- [8] MA N N, ZHANG X Y, ZHENG H T, et al. Shufflenet

- v2: Practical guidelines for efficient cnn architecture design[C]. Proceedings of the European conference on computer vision(ECCV), 2018: 116-131.
- [9] 邝先验,程福军,吴翠琴,等. 基于改进YOLOv7-tiny的高效轻量遥感图像目标检测方法[J]. 电子测量与仪器学报, 2024, 38(7): 22-33.
- KUANG X Y, CHENG F J, WU C Q, et al. Efficient and lightweight remote sensing image target detection method based on improved YOLOv7-tiny[J]. Journal of Electronic Measurement and Instrumentation, 2024, 38(7): 22-33.
- [10] 张羽,刘皓洋,贾升凯,等. 施工环境中安全帽和反光衣轻量化检测方法[J]. 电子测量技术, 2025, 48(8): 133-143.
- ZHANG Y, LIU H Y, JIA SH K, et al. Lightweight detection method for safety helmets and reflective clothing in construction environments[J]. Electronic Measurement Technology, 2025, 48(8): 133-143.
- [11] SHI D. Transnext: Robust foveal visual perception for vision transformers[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2024: 17773-17783.
- [12] SUN H, WEN Y, FENG H, et al. Unsupervised bidirectional contrastive reconstruction and adaptive fine-grained channel attention networks for image dehazing[J]. Neural Networks, 2024, 176: 106314.
- [13] YU Z P, HUANG H B, CHEN W J, et al. YOLO-face2: A scale and occlusion aware face detector[J]. Pattern Recognition, 2024, 155: 110714.
- [14] LIU SH, QI L, QIN H F, et al. Path aggregation network for instance segmentation[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2018: 8759-8768.
- [15] LIN T Y, DOLLAR P, GIRSHICK R, et al. Feature pyramid networks for object detection[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2017: 2117-2125.
- [16] ZHENG ZH H, WANG P, LIU W, et al. Distance-IoU loss: Faster and better learning for bounding box regression[C]. AAAI Conference on Artificial Intelligence, 2020, 34(7): 12993-13000.
- [17] LI X, WANG W, WU L, et al. Generalized focal loss: Learning qualified and distributed bounding boxes for dense object detection[J]. Advances in Neural Information Processing Systems, 2020, 33: 21002-21012.
- [18] YU J, JIANG Y, WANG Z, et al. Unitbox: An advanced object detection network[C]. 24th ACM International Conference on Multimedia, 2016: 516-520.
- [19] LU W, CHEN S B, TANG J, et al. A robust feature downsampling module for remote-sensing visual tasks[J]. IEEE Transactions on Geoscience and Remote Sensing, 2023, 61: 1-12.
- [20] YUAN X B, ZHENG ZH H, LI Y X, et al. Strip R-CNN: Large strip convolution for remote sensing object detection[J]. ArXiv preprint arXiv:2501.03775, 2025.
- [21] HE K M, ZHANG X Y, REN SH Q, et al. Deep residual learning for image recognition[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2016: 770-778.
- [22] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2018: 7132-7141.
- [23] WANG C Y, YE H I, MARK L H Y. YOLOv9: Learning what you want to learn using programmable gradient information[C]. European Conference on Computer Vision, 2025: 1-21.
- [24] WANG AO, CHEN H, LIU L H, et al. YOLOv10: Real-time end-to-end object detection[J]. Advances in Neural Information Processing Systems, 2024, 37: 107984-108011.
- [25] ZHANG J R, CHEN Z H, YAN G X, et al. Faster and lightweight: An improved YOLOv5 object detector for remote sensing images[J]. Remote Sensing, 2023, 15(20): 4974.
- [26] TAN M X, PANG R, LE Q V. Efficientdet: Scalable and efficient object detection[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 10781-10790.

作者简介

杜俊男, 硕士研究生, 主要研究方向为深度学习与图像处理。

E-mail: 1277052091@qq.com

杨雯, 硕士研究生, 主要研究方向为量子通信与深度学习。

E-mail: gs.wenyang23@gzu.edu.cn

刘志龙, 硕士研究生, 主要研究方向为深度学习与图像处理。

E-mail: 2485253878@qq.com

王成, 硕士研究生, 主要研究方向为深度学习与图像处理。

E-mail: 2231375292@qq.com

王天一(通信作者), 副教授, 硕士生导师, 主要研究方向为机器学习理论与应用。

E-mail: tywang@gzu.edu.cn