基于注意尺度融合跨空间学习的矿山井下传送带异物检测方法*

赵炳文[1],韦伟[2],杨光耀[3],郑振[4],潘博文[5]

(1.陕西陕煤榆北煤业有限公司, 陕西 榆林市 719000;
2.中国电信股份有限公司榆林分公司, 陕西 榆林市 719000;
3.陕西小保当矿业有限公司, 陕西 神木市 719302;
4.中国电信股份有限公司陕西分公司, 陕西 西安 710600;
5.西安科技大学 电气与控制工程学院, 陕西 西安 710054)

摘要:针对现有矿山井下传送带异物检测算法多以通道降维方式建模跨通道依赖,引起计算开销大且检测精度受限的问题,提出了一种基于注意尺度融合跨空间学习的异物检测方法(ES-YOLO)。首先,在主干网络中引入高效多尺度注意(EMA)模块,通过将通道重塑为批量维度,并将通道维度分组到多个子特征中,使空间语义特征在各特征组内均匀分布。随后,采用尺度序列特征融合(SSFF)模块提升网络的多尺度信息提取能力,该模块将不同尺度特征图水平堆叠,并利用3D卷积提取尺度序列特征,从而更好地融合深度特征图的高维信息与浅层特征图的细节信息。此外,维度感知选择性整合(DASI)模块自适应分配高维、低维特征图权重,增强模型对小目标的识别能力。试验结果表明,相比于基准模型YOLOv8n,该方法的平均精度均值提升了1.5个百分点,达89.7%;对大块矿料的类平均精度提升了4.1个百分点,参数量降低了18%;与主流模型相比,该模型在检测精度与轻量化程度方面均具有显著优势,更适合于矿山井下传送带异物实时检测场景。

关键词:异物检测; 跨空间学习; 矿山传送带; 特征融合

中图分类号:TD528;TP18 **文献标识码**:A **文章编号**:1005-2763(2026)06-0174-09

A Method for Detecting Foreign Objects in Underground Conveyor Belts in Mines Based on Attention-Scale Fusion and Cross-Spatial Learning

Zhao Bingwen[1], *Wei Wei*[2], *Yang Guangyao*[3], *Zheng Zhen*[4], *Pan Bowen*[5]

(1.SHCCIG Yubei Coal Industry Co., Ltd., Yulin, Shaanxi 719000, China;
2.Yulin Branch, China Telecom Corporation Limited, Yulin, Shaanxi 719000, China;
3.Shaanxi Xiaobaodang Mining Co., Ltd., Yulin, Shaanxi 719302, China;
4.Shaanxi Branch, China Telecom Corporation Limited, Xi'an, Shaanxi 710600, China;
5.College of Electrical and Control Engineering, Xi'an University of Science and Technology, Xi'an, Shaanxi 710054, China)

Abstract: To address the problems that existing foreign object detection algorithms for underground mine conveyor belts mainly model cross-channel dependencies through channel dimensionality reduction, resulting in high computational overhead and limited detection accuracy, a foreign object detection method based on attention-scale fusion and cross-spatial learning, termed ES-YOLO, was proposed. Firstly, an Efficient Multi-Scale Attention (EMA) module was introduced into the backbone network. By reshaping channels into the batch dimension and grouping channel dimensions into multiple sub-features, spatial semantic features were evenly distributed across different feature groups. Subsequently, a Scale Sequence Feature Fusion (SSFF) module was introduced to improve the network's capability for multi-scale feature extraction. By horizontally stacking feature maps at different scales and applying 3D convolution to capture scale-sequence features, the module effectively integrated high-dimensional semantic information from deep feature maps with detailed information from shallow feature maps. Furthermore, a Dimension-Aware Selective Integration (DASI) module was introduced to adaptively assign weights to high-dimensional and low-dimensional feature maps, thereby improving the model's ability to detect small objects. Experimental

* **收稿日期**:2025-06-25
基金项目:国家自然科学基金面上项目(52074213);陕西省重点研发计划项目(2021GY-131)
作者简介:赵炳文(1966—),男,陕西眉县人,高级工程师,主要从事煤矿生产智能化方面的研究。E-mail:Zhaobingwen1966@163.com

results demonstrate that, compared with the baseline YOLOv8n model, the proposed method increases the mean average precision by 1.5 percentage points to 89.7%. The class-wise mean average precision for large ore blocks improve by 4.1 percentage points, while the model parameters are reduced by 18%. Compared with mainstream models, the proposed method has significant advantages in detection accuracy and lightweight, making it more suitable for real-time foreign object detection in underground conveyor belts in mines.

Key words: Foreign object detection, Cross-spatial learning, Mine conveyor belt, Feature fusion

0 引言

矿山井下运输带对各类矿山的生产效率与安全至关重要,如何保障其稳定运行是亟待解决的关键问题。在开采和运输过程中,一旦有大块矿料或杂物等异物混入,轻则对运输带造成损坏导致生产中断,重则可能引发安全事故[1]。传统的人工检测方法不仅费时费力,而且工人在井下昏暗、高粉尘、强噪声等环境中长时间工作,极易因疲劳出现疏忽和误判,难以满足实时精准的检测需求。随着深度学习技术的发展,特别是兼具高精度与高实时性的目标检测算法不断涌现,基于机器视觉的传送带异物检测方法已日益成为研究热点,为提升矿山智能化水平和实现本质安全提供了有效技术途径[2]。

目标检测领域的方法主要分为两类,一类为两阶段(Two-stage)目标检测模型,典型算法以 Fast R-CNN、Faster R-CNN 等为代表;另一类为单阶段(One-stage)目标检测模型,代表算法包括 SSD(Single Shot MultiBox Detector)和 YOLO(You Only Look Once)系列。郝帅等[3]提出了在 YOLOv5 框架下引入卷积注意力机制(Convolutional Block Attention Module, CBAM)的方法,解决了传统算法由于光照、传送带运行速度及空气中粉尘干扰所导致的检测精度低等问题。然而,该方法由于需结合空间和通道注意力且更依赖于全局特征的计算,导致计算复杂度较高。Huang 等[4]以 YOLOv5 为基线,将多尺度视网膜增强算法(Multi-Scale Retinex,MSR)结合深度卷积,在一定程度上降低了计算量并简化了网络结构,但参数规模仍较大,推理速度未能满足实时检测要求。随着 YOLO 系列算法的发展[5],李宗霖等[6]将 GhostNetV2 和 HeadC2f_CA 模块引入了 YOLOv8 网络,在压缩模型大小的同时有效提升了检测精度;然而该模型特征融合能力仍存在不足,对部分遮挡及小目标的检测效果受限[7-8]。

针对上述问题,本文提出一种面向矿山井下传送带的基于注意尺度融合跨空间学习的异物检测方法(Enhanced Cross-Spatial Scale-Integration YOLO,ES-YOLO)。该方法首先在 YOLOv8 主干网络的两次卷积与 C2f 特征复用跨阶段部分结构中引入高效多尺度注意(Efficient Multi-Scale Attention,EMA)模块,以增强复杂背景下的特征提取能力。其次,在颈部引入尺度序列特征融合(Scale Sequence Feature Fusion, SSFF)模块,通过对特征图进行 3D 卷积提取尺度特征,使模型能更好地融合高低维信息。随后,提出维度感知选择性整合(Dimension-Aware Selective Integration, DASI)模块,该模块不仅能在不同尺寸特征图之间进行特征自适应融合,还可增强对图像底层特征的关注,从而提高模型在复杂背景下的鲁棒性。最后,通过试验验证了该模型算法在矿山井下传送带场景的高精度,能够在该场景下对传送带异物有较好的检测能力。

1 网络整体框架

针对现有方法中特征解耦不充分、多尺度上下文建模能力不足及高低维特征利用失衡等问题,提出一种面向矿山井下传送带的异物检测模型 ES-YOLO。该模型的整体结构分为三个部分:跨阶段融合高效多头注意力模块、基本卷积模块及由尺度序列特征融合模块和维度感知选择性整合模块组成的特征融合网络。其中,基本卷积层由 $k\times k$ 卷积层、归一化层和 SiLU 激活函数组成,k 代表卷积核尺寸。C2f 模块通过瓶颈分支增强了梯度传播与特征信息流动,ES-YOLO 算法结构如图 1 所示。

1.1 跨阶段融合高效多头注意力模块

由于矿山井下图像通常具有高噪声、低对比度等特点,传统 C2f 结构虽然可以提取到多尺度特征,但缺乏注意力机制对前景特征的关注,导致特征提取过程易受背景噪声干扰,从而影响检测精度。为

此,在 YOLOv8n 模型主干网络的 C2f 中添加 EMA 注意力机制,使网络更加关注与目标物体相关的特征,忽略无关背景信息,从而提升模型特征提取能力[9]。C2f-EMA 算法结构如图 2 所示。C2f-EMA 模块主要是在传统的 C2f 结构中的分割和瓶颈模块间添加了 EMA 注意力机制模块。

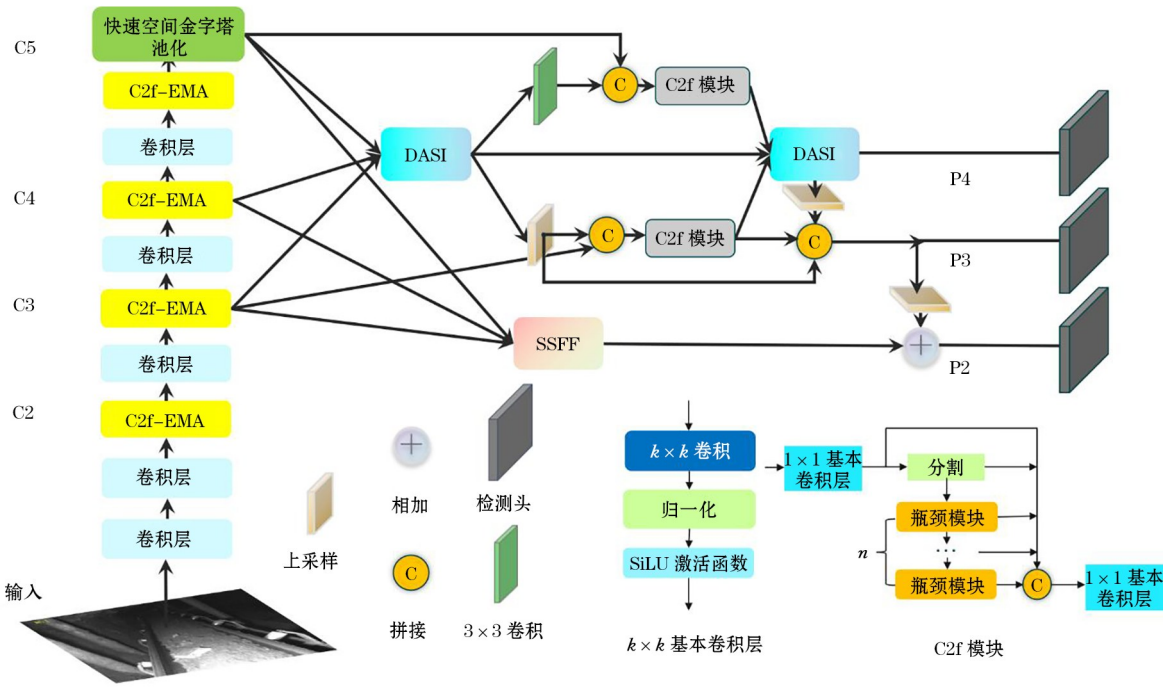


图 1 ES-YOLO 算法结构

Fig.1 ES-YOLO algorithm structure

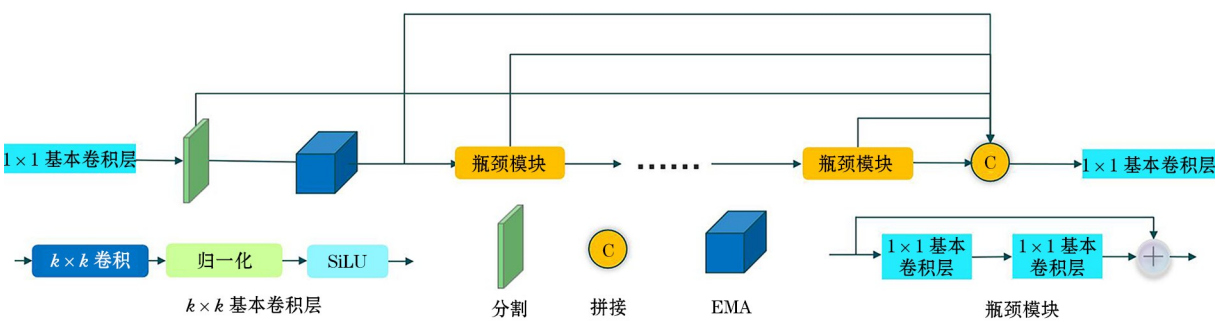


图 2 C2f-EMA 算法结构

Fig.2 C2f-EMA algorithm structure

首先,图像特征信息经过分割后被划分为 g 个子特征图。添加 EMA 注意力机制模块后,可对多个分支的特征图应用注意力机制,着重关注有效信息。在特征图传递至瓶颈模块之前,关键信息已经通过注意力机制得到过滤和强化,进而增强整个网络对局部特征的捕获能力[10]。EMA 注意力机制模块结构如图 3 所示,图中的 h、w、c 分别代表特征图的高度、宽度及通道数。

EMA 注意力机制可以有效建立特征间短期与长期的依赖关系,其由三个平行分支组成,其中两个分支分别对 X 和 Y 方向进行全局平均池化,第三个分支为 3×3 卷积分支。对 X 和 Y 方向的池化结果进行拼接后,再通过卷积实现特征融合;而 3×3 分支仅通过堆叠单个卷积核捕获多尺度特征。最后,特征经交叉空间学习模块和激活函数进一步加权,生成最终的注意力特征图。总体而言,为保留各通道信息并且减少计算开销,EMA 将通道维度分为多个分支,确保语义特征在各特征组内更均匀分布。通过并行的子结构,可避免更多的顺序处理和更大的深度。

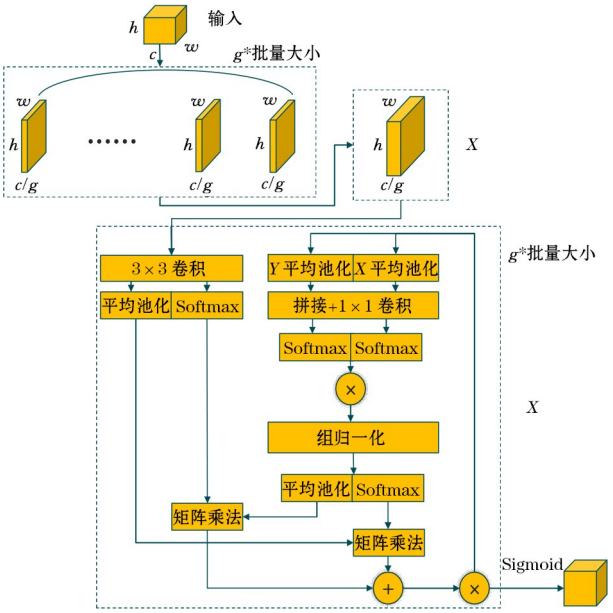


图 3 EMA 注意力机制模块结构

Fig.3 Structure of EMA attention mechanism module

1.2 尺度序列特征融合模块

针对井下大块矿料和锚杆图像的多尺度特征问题[11],现有文献中的特征金字塔结构多采用求和与拼接方式进行多尺度特征融合[12]。然而,传统特征金字塔结构难以有效利用不同维度特征图间的相互关系[13]。为此,在 ES-YOLO 模型中引入尺度序列特征融合模块,该模块旨在融合特征图中的全局信息与局部细节信息,尺度序列特征融合模块如图 4 所示。

从图 4 中可以看出,尺度序列特征融合模块首先使用 1×1 的卷积将 P4 和 P5 层特征的通道数统一为 256,并通过插值将其特征图调整为 P3 级别大小。随后,不同维度的特征图通过以下公式进行调整:

$$F_\sigma(w,h)=G_\sigma(w,h)\times f(w,h) \tag{1}$$

$$G_\sigma(w,h)=\frac{1}{2\pi\sigma^2}e^{-\frac{(w^2+h^2)}{2\sigma^2}} \tag{2}$$

式中:$f(w,h)$表示输入宽度 w 和高度 h 的二维图像;$G_\sigma(w,h)$表示一个二维的高斯滤波器,用于模糊输入图像;σ 为标准差;$F_\sigma(w,h)$为经高斯模糊后的图像。

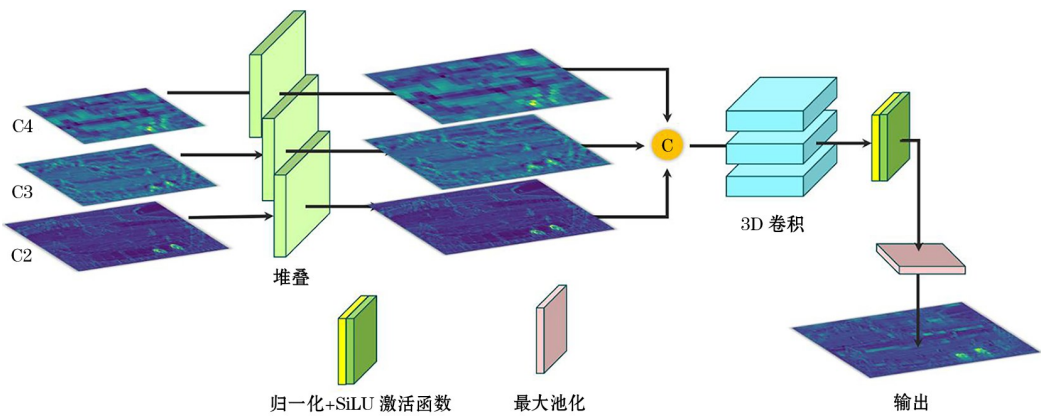


图 4 尺度序列特征融合模块结构

Fig.4 Structure of scale sequence feature fusion module

经过高斯模糊处理后,图像特征中的高频噪声被抑制,同时尺度对齐更平滑,有利于后续 3D 卷积操作。生成后的特征图空间分辨率相同但尺度不同。随后增加每个特征层的维度数量,使特征层(高度、宽度、通道)引入深度维度,将其扩展为四维张量(深度、高度、宽度、通道),并进一步拼接形成可用于 3D 卷积特征表示[14]。最后,利用 3D 卷积、归一化处理及 SiLU 激活函数完成尺度序列特征提取。

尺度序列特征融合模块通过融合不同层级和尺度的特征图,可以更全面地表征出传送带异物的形态与边缘特征,尤其适用于粉尘及低光照等复杂环境,从而提升模型的鲁棒性,降低环境干扰对识别精度的影响。

1.3 维度感知选择性整合模块

由于矿山井下传送带图像采集装置常处于高位,大块煤异物在整个图像中往往表现为小目标特征。传统金字塔网络无法有效检测小目标物体,一旦发生漏检,将对井下安全构成潜在威胁。基于此,在金字塔网络中引入维度感知选择性整合模块,以提升模型对小目标的检测能力。

在传统网络的降采样阶段,高维特征极易丢失图像的局部细节信息,而低维特征无法结合全局信息。维度感知选择性整合模块可以根据物体大小及背景信息,自适应选择高、中、低维度特征进行融合,维度感知选择性整合模块结构如图 5 所示。

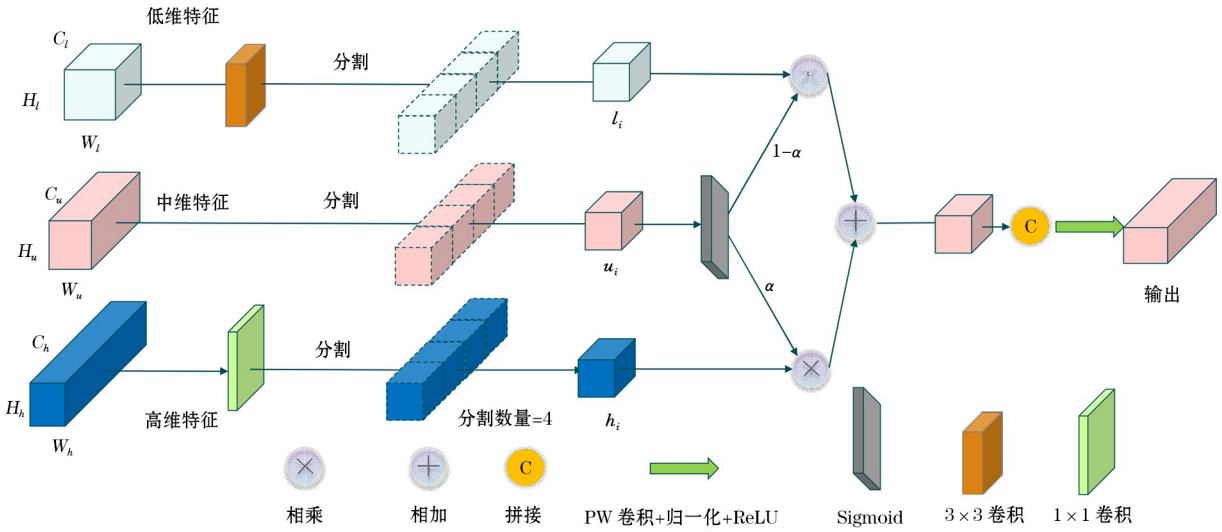


图 5 维度感知选择性整合模块结构

Fig.5 Structure of dimension-aware selective integration module

首先对高维特征 $F_h \in R^{H_h \times W_h \times C_h}$、中维度特征 $F_u \in R^{H_u \times W_u \times C_u}$ 和低维特征 $F_l \in R^{H_l \times W_l \times C_l}$ 通过卷积和插值等操作完成尺度对齐。随后执行分割操作,将其均分为通道数一致的四个部分,分割后的高中低三个维度的特征分别表示为:

$$(h_i)_{i=1}^{4} \in R^{H \times W \times \frac{C}{4}} \tag{3}$$

$$(u_i)_{i=1}^{4} \in R^{H \times W \times \frac{C}{4}} \tag{4}$$

$$(l_i)_{i=1}^{4} \in R^{H \times W \times \frac{C}{4}} \tag{5}$$

式中,l_i、u_i、h_i 表示低维、中维、高维的第 i 分区特征。随后,对中维度特征 u_i 采用 Sigmoid 激活函数处理,得到 α[14]。其中,$S(\cdot)$ 表示对特征施加 Sigmoid 激活函数。以上可以表示为:

$$\alpha = S(u_i) \tag{6}$$

$$u'_i = \alpha l_i + (1-\alpha) h_i \tag{7}$$

若 α 值大于 0.5,则说明模型特征缺少局部细节,需优先考虑细粒度特征;反之,则缺少全局信息,侧重上下文特征。最后,汇总各分区的选择性结果,在通道维度上进行拼接合并,u'_i 表示各分区的选择性聚合结果,得到输出特征 $u'_i \in R^{H \times W \times C}$。

2 试验结果与分析

2.1 试验数据集

试验所用数据集主要来自中国矿业大学公开数据集 CUMT-BelT[15]。该矿山井下图像数据集采集自矿井下皮带的运输环境,共分为三类:锚杆异物、大块矿料及正常煤流,并通过旋转、变色变换等方法进行数据增强,共获得 1 800 张图像。数据集按 7∶1∶2 的比例划分,其中训练集 1 260 张、验证集 180 张、测试集 360 张。试验中选取锚杆异物和大块矿料作为目标异物。部分数据集示例如图 6 所示。



图 6 部分数据集示例

Fig.6 Examples of some data sets

2.2 试验环境与参数配置

试验训练参数配置:输入图像尺寸均为 640×640,批量大小为 4,训练周期为 300。模型训练过程均开启混合精度训练,初始学习率为 0.001,模型采用 GPU 训练。模型训练的试验环境见表 1。

表 1 模型训练试验环境

Table 1 Model training test environment

参数	配置
CPU	AMD Ryzen 7 7800H
GPU	NVIDIA GeForce RTX 4060
内存	32 GB
系统	Windows 11
CUDA	12.1
Python	3.9.16
Torch	2.1.0

2.3 评价指标与预测结果

为验证改进算法的有效性与合理性,试验选取交并比为 0.5 条件下[16]的平均精度均值($mAP@50$)、类平均精度、参数量、计算量作为模型评价指

标,对改进网络的性能进行对比。其中,参数量用于衡量模型可训练参数总数,单位 M 表示百万级参数量;计算量用于衡量模型计算复杂程度,单位 G 指的是十亿次浮点运算次数。上述指标对评估模型规模与计算复杂度具有重要意义。

2.4 对比试验

2.4.1 注意力机制模块性能对比

为验证 C2f-EMA 模块的有效性,在主干部分的 C2f 中分别添加 CBAM、可变形注意力机制等主流算法,与 C2f-EMA 进行对比试验,试验结果见表 2。

表 2 添加注意力模块性能对比

Table 2 Performance comparison after adding attention module

模块	*mAP*@50/%	类平均精度/%		计算量/G	参数量/M
		大块矿料	锚杆		
Baseline	88.2	83.5	93.0	8.1	3.01
CBAM	88.6	83.3	94.0	8.8	3.12
DAttention	88.8	84.9	92.8	8.1	3.07
EMA	88.8	83.4	94.1	7.1	2.65

由表 2 可知,在 C2f 中添加 CBAM 注意力机制后,相较于 YOLOv8n 基线模型(Baseline),*mAP*@50 提升了 0.4 个百分点,但参数量和计算量均有所增加。将可变形注意力机制(DAttention)[17]引入 C2f 模块后,模型可根据输入特征动态调整注意力权重,从而提升细节捕捉能力,在 *mAP*@50 提升了 0.6 个百分点的同时,参数量与原始模型基本一致。相比之下,在 C2f 中添加 EMA 注意力机制,不仅使模型参数量减少 0.36 M、计算量降低 1.0 G,同时 *mAP*@50 提升了 0.6 个百分点。这主要得益于 EMA 通过将部分通道重塑为批量维度,使空间语义特征在每个特征组中均匀分布,有效增强模型对传送带异物的检测能力。因此,本文选用 C2f-EMA 作为主干网络的改进方案。

2.4.2 特征融合网络对比

为验证维度感知选择性整合模块与尺度序列特征融合组成的特征金字塔融合网络的有效性,分别与 BiFPN[18]、FDPN、ASF 等主流特征融合网络进行对比,试验结果见表 3。

表 3 特征融合网络性能对比

Table 3 Performance comparison of feature fusion networks

模块	*mAP*@50/%	类平均精度/%		计算量/G	参数量/M
		大块矿料	锚杆		
Baseline	88.2	83.5	93.0	8.1	3.01
BiFPN	88.0	81.8	94.3	7.1	1.99
FDPN	87.5	82.5	92.4	9.4	3.04
ASF	88.5	82.8	94.3	8.5	3.04
ES-YOLO	89.3	85.6	93.0	10.0	2.78

相比于 ASF 模型,本文提出的 ES-YOLO 网络在 *mAP*@50 上提升了 0.8 个百分点,同时参数量降低了 0.26 M。尽管计算量较基线模型有所增加,但精度的提升表明该网络具有更优的特征融合能力。相比之下,BiFPN 在参数量和计算量方面虽低于本文模型,但过度精简参数削弱了其特征提取能力。

2.5 消融试验

为验证所引入的各模块是有效的,开展了消融试验,结果见表 4。在 YOLOv8n 原模型基础上,将主干网络中原有的 C2f 模块替换为 C2f-EMA(记为模块Ⅰ),有效增强目标相关特征的提取能力,*mAP*@50 提升 0.6 个百分点的同时,模型计算量和参数均有所下降。引入维度感知选择性整合模块(记为模块Ⅱ),通过融合局部细节和全局语义信息,进一步增强对传送带异物的检测能力,*mAP*@50 相较于模块Ⅰ提升 0.1 个百分点。最后,在特征金字塔结构中加入尺度序列特征融合模块(记为模块Ⅲ),融合不同层级和不同尺寸的特征图,*mAP*@50 相较于模块Ⅰ+Ⅱ提升 0.4 个百分点。

由表 4 可知,ES-YOLO 不仅提升了检测精度,同时,可在一定程度上实现模型的轻量化,减少了模型的参数量和计算量。三项改进对矿山井下传送带异物检测效果呈递进提升,验证了改进方法的合理性和有效性。

表 4 消融试验性能对比

Table 4 Performance comparison of ablation tests

模块			*mAP*@50/%	类平均精度/%		计算量/G	参数量/M
Ⅰ	Ⅱ	Ⅲ		大块矿料	锚杆		
×	×	×	88.2	83.5	93.0	8.1	3.01
√	×	×	88.8	83.4	94.1	7.1	2.65
×	√	×	88.9	87.4	90.3	8.4	2.74
×	×	√	88.5	83.6	93.3	8.5	3.05
√	√	×	89.3	85.6	93.0	10.0	2.78
√	√	√	89.7	87.6	91.8	8.9	2.48

2.6 其他传送带异物检测模型对比

为进一步验证 ES-YOLO 的检测性能,在相同配置和数据集条件下,将其与三种矿山井下典型方法及主流目标检测模型 YOLOv5n、YOLOv8n、YOLOv8s、YOLOX 等方法开展对比试验,结果见表 5。

由表 5 可知,本文模型在检测精度方面表现最优,同时参数量进一步降低。与 YOLOv8n 和 YOLOv8s 基线模型相比,*mAP*@50 分别提升了1.5 和 1.3 个百分点;在模型复杂度方面,其参数量和计算量均显著低于 YOLOv8s;与 YOLOv8n 相比,本

文模型的参数量降低了 18%,大块矿料类平均精度提高了 4.1 个百分点。此外,与 YOLOv10n、RT-DETR 相比,本文模型在检测精度和参数规模上均具有一定优势。

此外,大多数由基线网络训练出的模型对大块矿料的检测能力较弱,原因在于异物颜色与背景难以分辨,而本文模型能有效提升其检测精度,维持较好的综合性能。

表 5 各网络模型性能对比

Table 5 Performance comparison of each network model

网络模型	*mAP*@50/%	类平均精度/%		计算量/G	参数量/M
		大块矿料	锚杆		
CBAM-YOLOv5s[3]	89.2	87.1	91.3	25.10	9.41
DWConv-YOLOv8[4]	88.7	83.4	93.3	8.00	3.03
GhostHGNetV2-YOLOv8[6]	88.9	86.5	91.3	6.80	2.31
YOLOv5	87.6	83.4	91.8	7.10	2.50
YOLOv7-tiny	87.0	83.1	91.0	13.10	6.01
YOLOv8n	88.2	83.5	93.0	8.10	3.01
YOLOv8s	88.4	86.0	90.8	28.40	11.12
YOLOX	85.1	79.1	89.9	26.70	8.94
YOLOv10n	87.8	83.0	92.7	8.22	2.71
RT-DETR	87.4	83.1	91.7	56.90	19.87
ES-YOLO	89.7	87.6	91.8	8.90	2.48

选取部分算法网络识别精度 *mAP*@50 和参数量的关系,如图 7 所示,可以看出 ES-YOLO 在保持较高精度的同时,参数量较低。

YOLOv8n、YOLOv10n、RT-DETR 与 ES-YOLO 模型的可视化检测结果如图 8 所示。从图 8 可以看出,在大块矿料检测任务中,ES-YOLO 得益于多尺度特征融合方法,能够更有效地融合局部细节与全局信息,从而提升对模糊边界目标的识别能力。与其他算法相比,该模型不仅能够检测到对比算法漏检的大块矿料,而且在已检测目标上的置信度更高。由此可见,ES-YOLO 在煤矿井下传送带异物检测场景下的目标检测能力优于当前主流算法。

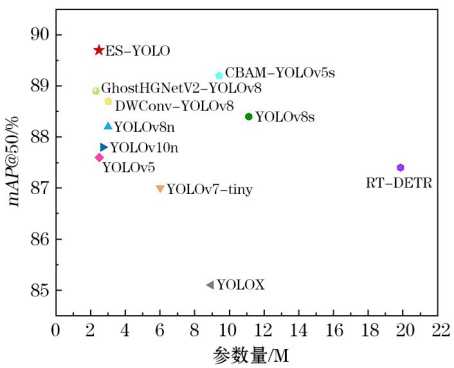
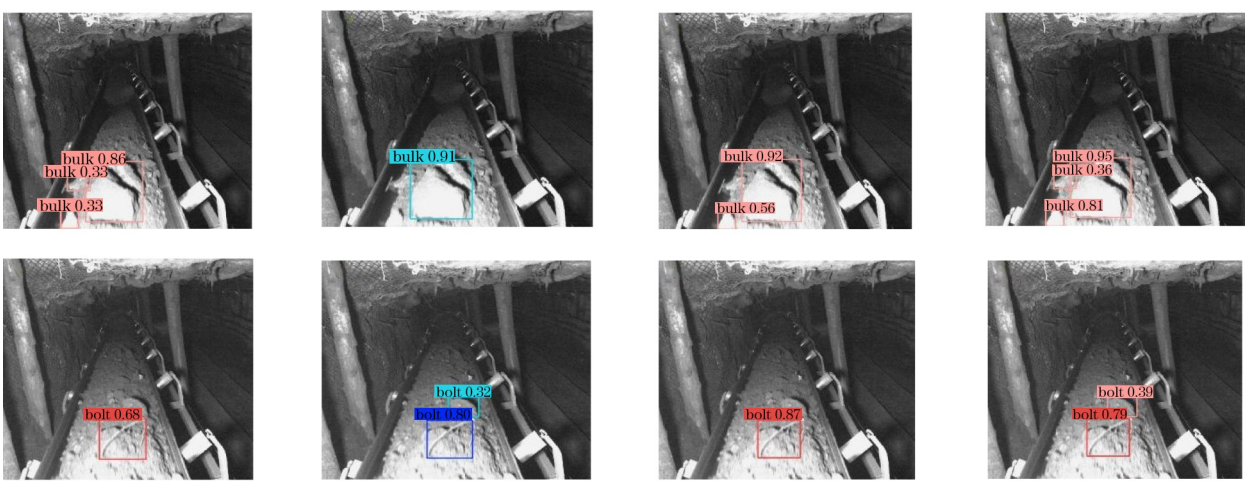


图 7 模型平均精度均值与参数量的关系

Fig.7 Relationship between model average precision mean value and the number of parameters



(a) YOLOv8n (b) YOLOv10n (c) RT-DETR (d) ES-YOLO

图 8 不同模型的可视化检测结果

Fig.8 Visual detection results of different models

3 结论

(1) 针对矿山井下传送带异物检测模型精度不高、大块矿料检测泛化能力差等问题,本文提出了ES-YOLO检测模型。相较于YOLOv8n,本文模型的$mAP@50$提升了1.5个百分点,参数量降低了18%,大块矿料类平均精度提高了4.1个百分点。与YOLOv5n、YOLOv10n及SSD等检测网络相比,本文模型在检测精度方面有明显优势。

(2) ES-YOLO网络在主干网络中引入C2f-EMA模块,以增强对目标相关特征的关注;同时构建由维度感知选择性整合模块和尺度序列特征融合模块组成的自适应特征融合颈部网络,实现了参数量、计算量与检测精度之间的最优平衡。

(3) 受试验样本数量及环境限制,当前试验结果的检测精度仍存在一定提升空间。随着训练数据集的不断扩充,所提的检测网络精度有望进一步提高。后续工作将继续聚焦于在保证模型精度不下降的前提下,实现模型在硬件资源受限矿井场景下的部署。

参考文献(References):

[1] 杨建辉,黄子洋,汪梅,等.机器视觉灰度化金字塔卷积模型的煤流异物识别[J].煤炭科学技术,2022,50(11):194-201.
Yang Jianhui, Huang Ziyang, Wang Mei, et al. Recognition of unwanted objects in coal flow based on gray pyramid convolution model of machine vision[J]. Coal Science and Technology,2022,50(11):194-201.

[2] 徐志强,吕子奇,王卫东,等.煤矸智能分选的机器视觉识别方法与优化[J].煤炭学报,2020,45(6):2207-2216.
Xu Zhiqiang, Lyu Ziqi, Wang Weidong, et al. Machine vision recognition method and optimization for intelligent separation of coal and gangue[J]. Journal of China Coal Society,2020,45(6):2207-2216.

[3] 郝帅,张旭,马旭,等.基于CBAM-YOLOv5的煤矿输送带异物检测[J].煤炭学报,2022,47(11):4147-4156.
Hao Shuai, Zhang Xu, Ma Xu, et al. Foreign object detection in coal mine conveyor belt based on CBAM-YOLOv5[J]. Journal of China Coal Society,2022,47(11):4147-4156.

[4] Huang K F, Li S Y, Cai F, et al. Detection of large foreign objects on coal mine belt conveyor based on improved[J]. Processes,2023,11(8):2469.

[5] 张新航,张雅茹,麻振华,等.基于深度学习方法的YOLO目标检测综述[J].长江信息通信,2024,37(8):52-56.
Zhang Xinhang, Zhang Yaru, Ma Zhenhua, et al. Overview of YOLO object detection based on deep learning methods [J]. Changjiang Information & Communication, 2024, 37(8):52-56.

[6] 李宗霖,王广祥,张立亚,等.基于改进YOLOv8n的煤矿带式输送异物检测研究[J].矿业安全与环保,2024,51(4):41-48.
Li Zonglin, Wang Guangxiang, Zhang Liya, et al. Research on the detection of foreign objects in coal mine belt conveying based on improved YOLOv8n[J]. Mining Safety and Environmental Protection,2024,51(4):41-48.

[7] Song K Y, Yang H, Yin Z P. Multi-scale attention deep neural network for fast accurate object detection[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2018,29(10):2972-2985.

[8] Zhang Y, Zhou A B, Zhao F K, et al. A lightweight vehicle-pedestrian detection algorithm based on attention mechanism in traffic scenarios[J]. Sensors,2022,22(21):8480.

[9] 陈佳鑫,赵国贞,程伟,等.基于动态称量和图像识别技术的井下煤矸石智能分选装置研发[J].矿业研究与开发,2023,43(2):178-183.
Chen Jiaxin, Zhao Guozhen, Cheng Wei, et al. Development of underground coal gangue intelligent sorting device based on dynamic weighing and image recognition technology[J]. Mining Research and Development,2023,43(2):178-183.

[10] 陈科圻,朱志亮,邓小明,等.多尺度目标检测的深度学习研究综述[J].软件学报,2021,32(4):1201-1227.
Chen Keqi, Zhu Zhiliang, Deng Xiaoming, et al. Deep learning for multi-scale object detection:a survey[J]. Journal of Software,2021,32(4):1201-1227.

[11] 员娇娇,胡永利,孙艳丰,等.基于深度学习的小目标检测方法综述[J].北京工业大学学报,2021,47(3):293-302.
Yuan Jiaojiao, Hu Yongli, Sun Yanfeng, et al. Survey of small object detection methods based on deep learning[J]. Journal of Beijing University of Technology,2021,47(3):293-302.

[12] Kang M, Ting C M, Ting F F, et al. ASF-YOLO:a novel YOLO model with attentional scale sequence fusion for cell instance segmentation[J]. Image and Vision Computing, 2024,147:105057.

[13] 王国新,祝建东,汝洪芳.基于改进YOLOv8s的矸石流目标检测方法[J].矿业研究与开发,2025,45(4):229-237.
Wang Guoxin, Zhu Jiandong, Ru Hongfang. An object detection method for gangue flow based on improved YOLOv8s[J]. Mining Research and Development,2025,45(4):229-237.

[14] Xu S B, Zheng S C, Xu W H, et al. HCF-Net:hierarchical context fusion network for infrared small object detection [J]. IEEE Transactions on Geoscience and Remote Sensing, 2021,59(11):9808-9820.

[15] 程德强,钱建生,郭星歌,等.煤矿安全生产视频AI识别关键技术研究综述[J].煤炭科学技术,2023,51(2):349-365.

Cheng Deqiang, Qian Jiansheng, Guo Xingge, et al. Review on key technologies of AI recognition for videos in coal mine[J]. Coal Science and Technology,2023,51(2):349-365.

[16] Zheng Z H, Wang P, Liu W, et al. Distance-IoU loss:faster and better learning for bounding box regression[C]// Proceedings of the AAAI Conference on Artificial Intelligence. Palo Alto:AAAI,2020,34(7):12993-13000.

[17] Zhu X Z, Su W J, LU L W, et al. Deformable DETR: deformable transformers for end-to-end object detection[J]. arXiv preprint, 2020:04159.

[18] Tan M X, Pang R M, Le Q V. EfficientDet:scalable and efficient object detection[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). New York:IEEE,2020:10781-10790.