



文章编号: 1671-251X(2024)03-0035-07

DOI: 10.13272/j.issn.1671-251x.2023090088

# 基于 YOLOv7-SE 的煤矿井下场景小目标检测方法

曹帅, 董立红, 邓凡, 高峰

(西安科技大学 计算机科学与技术学院, 陕西 西安 710054)

**摘要:** 目前的小目标检测方法虽然提高了小目标检测效果, 但针对的多为常规场景, 而煤矿井下环境恶劣, 在井下小目标检测过程中存在小目标特征信息提取困难的问题。针对上述问题, 提出了一种基于 YOLOv7-SE 的煤矿井下场景小目标检测方法。首先, 将模拟退火(SA)算法与 k-means++聚类算法融合, 通过优化 YOLOv7 模型中初始锚框值的估计, 准确捕捉井下小目标; 然后, 在 YOLOv7 骨干网络中增加新的检测层得到井下小目标高分辨率特征图, 减少大量煤尘对井下小目标特征表示的干扰; 最后, 在骨干网络中的聚合网络模块后引入双层注意力机制, 强化井下小目标的特征表示。实验结果表明: ① YOLOv7-SE 网络模型训练后的损失函数值稳定在 0.05 附近, 说明 YOLOv7-SE 网络模型参数设置合理。② 基于 YOLOv7-SE 网络模型的安全帽检测平均精度(AP)较 Faster R-CNN, RetinaNet, CenterNet, FCOS, SSD, YOLOv5, YOLOv7 分别提升了 13.86%, 25.3%, 16.13%, 12.71%, 15.53%, 11.59%, 12.20%。基于 YOLOv7-SE 网络模型的自救器检测 AP 较 Faster R-CNN, RetinaNet, CenterNet, FCOS, SSD, YOLOv5, YOLOv7 分别提升了 12.37%, 20.16%, 15.22%, 8.35%, 19.42%, 9.64%, 7.38%。YOLOv7-SE 网络模型的每秒传输帧数(FPS)较 Faster R-CNN, RetinaNe, CenterNet, FCOS, SSD, YOLOv5 分别提升了 42.56, 44.43, 31.74, 39.84, 22.74, 23.34 帧/s, 较 YOLOv7 下降了 9.36 帧/s。说明 YOLOv7-SE 网络模型保证检测速度的同时, 有效强化了 YOLOv7-SE 网络模型对井下小目标的特征提取能力。③ 在对安全帽和自救器的检测中, YOLOv7-SE 网络模型有效改善了漏检和误检问题, 提高了检测精度。

**关键词:** 煤矿井下; 小目标检测; 特征提取; YOLOv7; 聚类分析; 模拟退火

中图分类号: TD391 文献标志码: A

A small object detection method for coal mine underground scene based on YOLOv7-SE

CAO Shuai, DONG Lihong, DENG Fan, GAO Feng

(College of Computer Science and Technology, Xi'an University of Science and Technology, Xi'an 710054, China)

**Abstract:** Although current small object detection methods have improved the detection performance, they are mostly objected at conventional scenarios. In harsh underground environments in coal mines, there are difficulties in extracting small object feature information during the underground small object detection process. In order to solve the problem, a small object detection method for coal mine underground scenes based on YOLOv7-SE has been proposed. Firstly, the simulated annealing (SA) algorithm is integrated with the k-means++clustering algorithm to accurately capture small underground objects by optimizing the estimation of initial anchor box values in the YOLOv7 model. Secondly, a new detection layer is added to the YOLOv7 backbone network to obtain high-resolution feature maps of underground small objects, reducing the interference of a large amount of coal dust on the feature representation of underground small objects. Finally, a dual layer attention mechanism is introduced after the aggregation network module in the backbone network to enhance the

收稿日期: 2023-09-27; 修回日期: 2024-03-13; 责任编辑: 王晖, 郑海霞。

基金项目: 陕西省自然科学基础研究计划项目(2019JLM-11); 陕西省教育厅科研计划专项项目(8146119003); 陕西省自然科学基金项目(2018JQ5095)。

作者简介: 曹帅(1998—), 男, 陕西榆林人, 硕士研究生, 研究方向为煤矿智能化, E-mail: ff5225252@qq.com。

引用格式: 曹帅, 董立红, 邓凡, 等. 基于 YOLOv7-SE 的煤矿井下场景小目标检测方法[J]. 工矿自动化, 2024, 50(3): 35-41.

CAO Shuai, DONG Lihong, DENG Fan, et al. A small object detection method for coal mine underground scene based on YOLOv7-SE[J]. Journal of Mine Automation, 2024, 50(3): 35-41.



扫码移动阅读

feature representation of small underground objects. The experimental results show the following points. ① The loss function of the YOLOv7-SE network model after training is stable around 0.05, indicating that the parameter settings of the YOLOv7-SE network model are reasonable. ② The average precision (AP) of helmet detection based on the YOLOv7-SE network model has improved by 13.86%, 25.3%, 16.13%, 12.71%, 15.53%, 11.59% and 12.20% compared to Faster R-CNN, RetinaNet, CenterNet, FCOS, SSD, YOLOv5 and YOLOv7, respectively. The self rescue device detection AP based on the YOLOv7-SE network model has improved by 12.37%, 20.16%, 15.22%, 8.35%, 19.42%, 9.64% and 7.38% compared to Faster R-CNN, RetinaNet, CenterNet, FCOS, SSD, YOLOv5 and YOLOv7, respectively. The frames per second (FPS) of the YOLOv7-SE network model has increased by 42.56, 44.43, 31.74, 39.84, 22.74 and 23.34 frames/s compared to Faster R-CNN, RetinaNet, CenterNet, FCOS, SSD and YOLOv5, respectively, and decreased by 9.36 frames/s compared to YOLOv7. The YOLOv7-SE network model effectively enhances the feature extraction capability of the YOLOv7-SE network model for small underground objects while ensuring detection speed. ③ In the detection of safety helmets and self rescue devices, the YOLOv7-SE network model effectively improves missed and false detection, and improves detection precision.

**Key words:** underground coal mine; small object detection; feature extraction; YOLOv7; clustering analysis; simulated annealing

## 0 引言

据统计,近10年来煤矿安全事故率持续下降,但仍在高危行业前列<sup>[1]</sup>。为了保障井下人员安全生产,矿工在井下作业过程中必须佩戴防护设备。由于部分矿工安全意识低,对于防护设备不是很重视,不能有效佩戴防护设备,保障自身安全。目前煤矿企业主要通过人工及视频监控摄像头查看的方法来监督矿工是否佩戴防护设备。由于井下作业环境的监控摄像头位置固定,其覆盖范围广、拍摄距离远,监控画面中防护设备目标的尺寸较小,导致小型防护设备目标易受到目标尺寸和井下环境变化的影响,检测难度大大增加。因此,对煤矿井下场景的小目标(尺寸小于32×32的目标)进行检测研究,在小型防护设备监测中起至关重要的作用。

传统的小目标检测方法<sup>[2-4]</sup>难以有效提取井下小目标特征信息,随着深度学习的兴起,卷积神经网络(Convolutive Neural Networks, CNN)模型<sup>[5]</sup>逐步代替了传统小目标检测方法,基于CNN的小目标检测方法主要分为两阶段目标检测和单阶段目标检测。两阶段目标检测算法通过生成区域网络(Region Proposal Network, RPN)提取到感兴趣的特征信息后进行分类,例如R-CNN算法(Region with CNN feature)<sup>[6]</sup>、Fast R-CNN算法<sup>[7]</sup>及Faster R-CNN算法<sup>[8]</sup>等,该类算法需生成大量候选区域,检测速度慢,无法满足对小目标实时检测的要求。单阶段目标检测算法将检测归纳为回归问题,实现端到端的检测技术,如单步多框目标检测(Single Shot

MultiBox Detector, SSD)算法<sup>[9]</sup>、YOLO系列算法<sup>[10-14]</sup>,该类检测算法的检测速度较快,但会有一定的误差。基于CNN的小目标检测方法较传统方法有很大提升,但仍然存在召回率低、误检率高的问题。针对上述问题,文献[15]为了准确处理和提取小目标信息特征,在YOLOv3网络模型的特征金字塔网络中自适应融合浅层和深层特征图的局部和全局特征。文献[16]提出了一种轻量型特征提取模块,该模块采用空洞瓶颈和多尺度卷积获得更丰富的图像特征信息,增强了目标特征表达能力。文献[17]在YOLOv5s模型Backbone区域嵌入自校正卷积(Self-Calibrated Convolution, SCConv)作为特征提取网络,可更好地融合多尺度特征信息。文献[18]提出了一种结合通道和空间注意力引导的残差学习方法,用于捕捉目标的关键信息。文献[19]提出了高分辨率表示模块,通过使用多尺度特征来捕捉目标的细节信息,并将其融合到高分辨率表示模块,有助于提高目标的定位准确性。

上述研究虽然提高了小目标检测效果,但针对的多为常规场景,受煤矿井下恶劣环境影响,在检测过程中存在井下小目标特征信息提取困难的问题。针对该问题,本文提出一种基于YOLOv7-SE的井下场景小目标检测方法。首先,将模拟退火(Simulated Annealing, SA)算法与k-means++聚类算法融合,优化YOLOv7网络模型中初始锚框值。然后,在YOLOv7网络模型骨干网络中增加新的检测层,得到井下小目标信息丰富的特征图。最后,在YOLOv7网络模型骨干网络的聚合网络模块之后引入双层注意力机制<sup>[20]</sup>,强化井下小目标特征表示。

## 1 YOLOv7 网络模型

YOLOv7 网络模型<sup>[21]</sup>主要由输入层、骨干网络、检测头 3 个模块组成,如图 1 所示。

输入层将图像缩放到固定尺寸  $640 \times 640$ , 以满足骨干网络对尺寸输入的要求。骨干网络由卷积模块、聚合网络模块和最大池化模块组成, 对输入图像进行特征提取。卷积模块主要由 1 个卷积层、1 个批

量归一化层和 1 个 SiLU 激活函数构成。聚合网络模块包含 3 个  $1 \times 1$  卷积层和 4 个  $3 \times 3$  卷积层。最大池化模块由 1 个最大池化层、2 个  $1 \times 1$  卷积层和 1 个  $3 \times 3$  卷积层组成。检测头由路径聚合特征金字塔网络 (Path Aggregation Feature Pyramid Network, PAFPN) 和检测层组成，PAFPN 用于融合不同尺寸的特征图，检测层输出带有检测类别和准确度的结果。

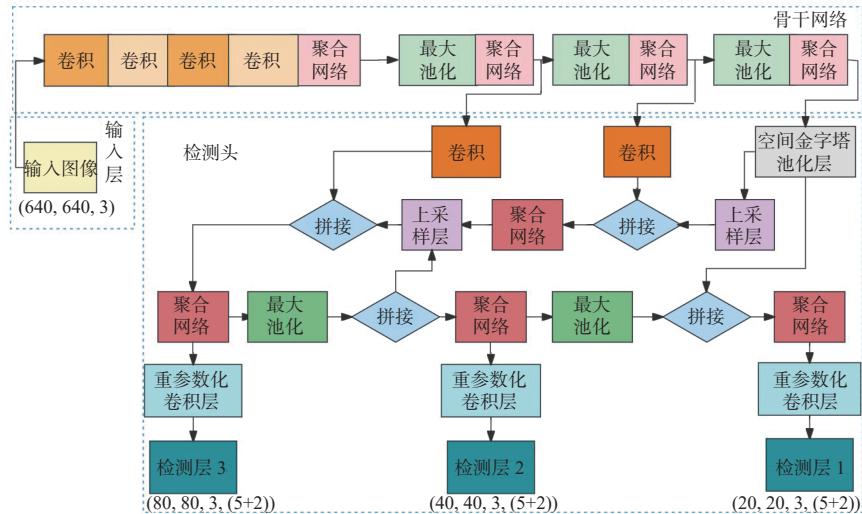


图 1 YOLOv7 网络结构

Fig. 1 YOLOv7 network structure

## 2 YOLOv7-SE 小目标检测方法

## 2.1 融合 SA 和 k-means++聚类算法

YOLOv7 网络模型的初始锚框尺寸是针对井下常规目标设计，难以适用于井下小目标检测，因此，使用 k-means++聚类算法对 YOLOv7 网络模型进行聚类分析，生成新的锚框尺寸，若随机选取的初始化聚类中心点在密度小的簇内，会使其相似性较小，聚类结果较差。因此，利用 SA 算法<sup>[22]</sup>来确定 k-means++聚类算法中最优的首个初始聚类中心点。SA 算法是寻找数据集样本中最优值的算法，先初始化 SA 算法的参数，如初始温度、终止温度和迭代次数。通过迭代的方式，不断更新聚类中心点的值，并计算目标函数的值。根据 SA 策略判断是否接受当前参数的更新，若接受，则更新聚类中心点的值，否则降低温度继续迭代下一次参数值，直到达到终止温度，返回最优的聚类中心点。

将 SA 算法与 k-means++聚类算法进行融合，并对本文数据集进行聚类分析，得到的锚框尺寸分别为 [9, 31], [9, 17], [15, 37], [16, 23], [21, 29], [26, 40], [29, 71], [39, 55], [50, 79]。

## 2.2 新的目标检测层

YOLOv7 网络模型中深层网络提取抽象的语义特征信息,用来反映并下大目标特征信息,而浅层网络提取目标的细节特征,能够保留更多的并下小目标特征信息。YOLOv7 网络模型通常通过增加下采样结构以获取更大的目标感受野,但随着下采样结构的增加,并下小目标的特征信息逐渐丢失,对于在图像中特征信息少的并下小目标不友好,不利于并下小目标检测。

为了使 YOLOv7 网络模型重点提取浅层网络中丰富的井下小目标特征信息,而抑制深层网络的特征提取,本文去掉 YOLOv7 网络模型骨干网络的 32 倍下采样结构,在 YOLOv7 网络模型中增加 1 层新的目标检测层,获取浅层网络中细节信息更明确的井下小目标特征,减少在下采样过程中丢失井下小目标特征。YOLOv7 网络模型从第 1 个聚合网络模块开始提取特征信息,经过 PAFPN 输出到检测层,得到新尺寸( $160 \times 160$ )的特征图,该特征图的井下小目标信息相对丰富。

### 2.3 引入双层注意力机制

针对煤矿复杂环境下小目标图像特征抽取不准

确的问题,本文在 YOLOv7 网络模型的聚合网络模块之后添加双层注意力机制,以强化聚合网络模块对井下小目标的特征提取能力。

双层注意力机制整体结构如图 2 所示。第 1 阶段使用重叠块嵌入,将原始的二维输入图像转换为一维的图像块。第 2 阶段到第 4 阶段使用合并块模块进行下采样操作,用于调整相应通道数,并降低输入分辨率,同时在每个阶段的操作后采用双层注意力模块(图 3)对特征图做特征变换。

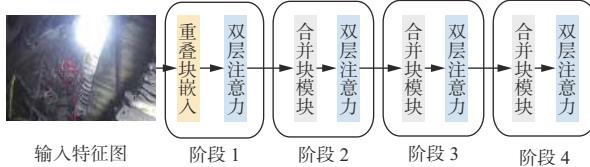


图 2 双层注意力机制整体结构

Fig. 2 Overall structure of dual layer attention mechanism

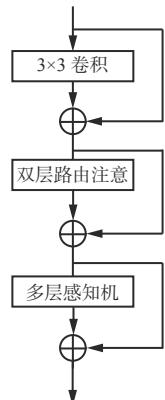


图 3 双层注意力模块

Fig. 3 Dual layer attention module

通过收集前  $n$  个相关区域中的键值对,并利用稀疏性操作忽略最不相关区域的计算来节省模块参数数量和计算量,并对收集的键值对进行注意力操作。

$$\mathbf{O} = A_{\text{attention}}(\mathbf{Q}, \mathbf{K}^T, \mathbf{V}) + L_{\text{CE}}(\mathbf{V}) \quad (1)$$

式中:  $\mathbf{O}$  为输出;  $A_{\text{attention}}(\cdot)$  为自注意力函数;  $\mathbf{Q}$  为查询;  $\mathbf{K}^T$  为键的张量;  $\mathbf{V}^T$  为值的张量;  $L_{\text{CE}}(\cdot)$  为局部上下文增强模块;  $\mathbf{V}$  为值。

$$A_{\text{attention}}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = S \left( \frac{\mathbf{Q}\mathbf{K}^T}{\sqrt{d}} \right) \mathbf{V} \quad (2)$$

式中:  $\mathbf{K}$  为键;  $S(\cdot)$  为归一化函数;  $d$  为缩放因子;  $\mathbf{Q}\mathbf{K}^T$  为  $\mathbf{Q}$  和  $\mathbf{K}$  之间的相似程度。

对 YOLOv7 进行上述改进,将改进后的网络模型命名为 YOLOv7-SE。

### 3 实验结果与分析

#### 3.1 数据集

本文数据集来源于煤矿井下多个场景的视频监

控摄像头拍摄的图像视频,包括采煤工作面、胶带机头工作面、井下巷道、井下站台、煤壁面、井下候车点等场景,如图 4 所示。



图 4 不同场景样本示例

Fig. 4 Sample examples of different scenarios

为提升数据集的多样性,通过水平、垂直翻转及随机方向旋转等方法扩充数据。整理后共有 5 622 张图像,将数据集按照 8:1:1 的比例划分为训练集、验证集和测试集。同时,使用 labelImg 工具对数据集进行标注,在满足小目标尺寸的条件下,标注类别有安全帽和自救器。

为解决井下环境中煤尘对图像的干扰,在 YOLOv7 网络模型的数据预处理阶段增加图像处理模块,对数据集进行预处理,如图 5 所示。

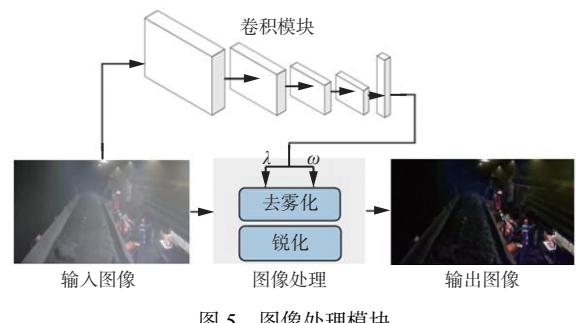


图 5 图像处理模块

Fig. 5 Image processing module

首先,采用暗通道去雾方法<sup>[23]</sup>对原始图像进行去雾,以减少井下煤尘对图像的影响。然后,采用高斯函数对去雾后的图像进行锐化,以突出图像细节,提高井下场景目标边缘与周围像素之间的反差。最后,使用卷积模块作为调优器,利用其反向传播的特性对图像处理方法中的去雾程度和锐化强度进行优化,以达到更好的增强效果。

### 3.2 评价指标

常用的小目标评价指标包括准确率  $P$ 、召回率  $R$ 、平均精度(Average Precision, AP)、所有类别的平均精度值(mean Average Precision, mAP)及每秒传输帧数(Frames Per Second, FPS)。

$$P = \frac{N_{TP}}{N_{TP} + N_{FP}} \quad (3)$$

$$R = \frac{N_{TP}}{N_{TP} + N_{FN}} \quad (4)$$

$$AP = \int_0^1 PRdR \quad (5)$$

$$mAP = \frac{\sum AP}{M} \quad (6)$$

式中:  $N_{TP}$  为预测正确的正样本数量;  $N_{FP}$  为预测错误的正样本数量;  $N_{FN}$  为预测错误的负样本数量;  $M$  为目标的类别数。

### 3.3 实验配置

本文实验在 ubuntu20.04 操作系统中搭建, 具体配置见表 1。

表 1 实验环境配置

Table 1 Experimental environment configuration

实验环境	配置
GPU	RTX 3090(24 GiB)
CPU	12 vCPU Xeon(R) Platinum 8255C
操作系统	ubuntu20.04
GPU环境	CUDA11.3 cuDNN8.2.1
深度学习框架	Pytorch1.11
编译器	Python3.8

### 3.4 结果分析

#### 3.4.1 模型训练

在模型训练前需对实验超参数进行设置, 迭代次数为 200, 初始学习率为 0.015, 批量大小为 32, 选取 640×640 的图像作为模型的输入。在模型训练过程中损失函数值随迭代次数变化曲线如图 6 所示。

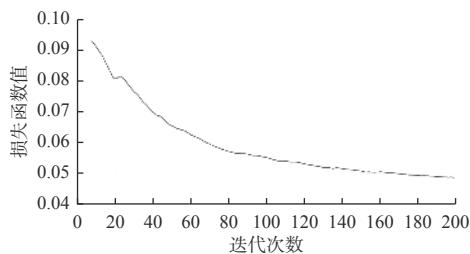


图 6 模型训练过程

Fig. 6 Model training process

由图 6 可看出在模型训练过程前 80 次迭代, 损失函数值下降十分明显, 第 80—180 次迭代时, 损失

函数值下降趋势趋于平缓, 最后 20 次迭代的损失函数值已逐渐稳定。模型训练过程中的最终损失函数值低于 0.05, 说明本文模型的训练参数设置合理, 模型学习效果较好。

#### 3.4.2 对比实验

为了衡量 YOLOv7-SE 网络模型的检测性能, 将其与 Faster R-CNN, RetinaNet, CenterNet, FCOS (Fully Convolutional One-Stage), SSD, YOLOv5, YOLOv7 目标检测模型进行对比, 结果见表 2。

表 2 各模型对比结果

Table 2 Comparison results of each model

模型	AP/%		mAP/%	FPS/(帧·s <sup>-1</sup> )
	安全帽	自救器		
Faster R-CNN	58.64	52.11	55.38	19.28
RetinaNet	47.20	44.32	45.76	17.41
CenterNet	56.37	49.26	52.82	30.10
FCOS	59.79	56.13	57.96	22.00
SSD	56.97	45.06	51.02	39.10
YOLOv5	60.91	54.84	57.88	38.50
YOLOv7	60.30	57.10	58.70	71.20
YOLOv7-SE	72.50	64.48	68.49	61.84

由表 2 可看出, YOLOv7-SE 网络模型的安全帽检测 AP 较 Faster R-CNN, RetinaNet, CenterNet, FCOS, SSD, YOLOv5, YOLOv7 分别提升了 13.86%, 25.3%, 16.13%, 12.71%, 15.53%, 11.59%, 12.20%。YOLOv7-SE 网络模型的自救器检测 AP 较 Faster R-CNN, RetinaNet, CenterNet, FCOS, SSD, YOLOv5, YOLOv7 分别提升了 12.37%, 20.16%, 15.22%, 8.35%, 19.42%, 9.64%, 7.38%。YOLOv7-SE 网络模型的 mAP 较 Faster R-CNN, RetinaNet, CenterNet, FCOS, SSD, YOLOv5, YOLOv7 分别提升了 13.11%, 22.73%, 15.67%, 10.53%, 17.47%, 10.61%, 9.79%。YOLOv7-SE 网络模型的 FPS 较 Faster R-CNN, RetinaNet, CenterNet, FCOS, SSD, YOLOv5 分别提升了 42.56, 44.43, 31.74, 39.84, 22.74, 23.34 帧/s, 较 YOLOv7 下降 9.36 帧/s。说明 YOLOv7-SE 网络模型的检测性能更佳。

#### 3.4.3 消融实验

为了验证不同改进方法对 YOLOv7 网络模型性能的影响, 设计了 6 组实验, 改进的骨干网络实验为添加新的目标检测层和引入双层注意力机制, 实验结果见表 3。

由表 3 可看出, 使用改进的 k-means++ 方法重新聚类分析锚框值, 安全帽检测 AP、自救器检测 AP、

表 3 消融实验结果  
Table 3 Results of ablation experiment

模型	AP/%		mAP/%	FPS/(帧·s <sup>-1</sup> )
	安全帽	自救器		
YOLOv7	60.30	57.10	58.70	71.20
YOLOv7+改进的k-means++	63.21	60.70	61.95	74.20
YOLOv7+改进骨干网络	70.70	62.32	66.51	63.18
YOLOv7-SE	72.50	64.48	68.49	61.84

mAP、FPS 分别为 63.21%、60.7%，61.95%，74.2 帧/s，较 YOLOv7 分别提升了 2.91%，3.6%，3.25%，3 帧/s；改进 YOLOv7 网络模型骨干网络后，安全帽检测 AP、自救器检测 AP、mAP 分别为 70.7%，62.32%，66.51%，较 YOLOv7 分别提升了 10.4%，5.22%，7.81%，FPS 下降 8.02 帧/s，为 63.18 帧/s；YOLOv7-SE 网络模型的安全帽 AP、自救器 AP、mAP 分别为 72.5%，64.48%，68.49%，较 YOLOv7 分别提升了 12.2%，7.38%，9.79%，FPS 为 61.84 帧/s，说明 YOLOv7-SE 网络模型在保证检测速度的同时，有效强化了 YOLOv7-SE 网络模型对井下小目标的特征提取能力。

#### 3.4.4 检测效果对比分析

为更加直观地体现 YOLOv7-SE 网络模型的优越性，在井下采煤工作面、胶带机头工作面、井下巷道及井下站台等场景中与 YOLOv7 网络模型对比，对比结果如图 7 所示。



图 7 检测效果对比

Fig. 7 Comparison of detection effects

由图 7 可看出，在对安全帽和自救器的检测中，YOLOv7 网络模型出现漏检和误检的问题，而 YOLOv7-SE 网络模型有效改善了该问题，提高了检测精度。因此，YOLOv7-SE 网络模型可满足井下小目标检测任务。

## 4 结论

1) 针对煤矿井下场景中目标尺寸较小、环境存在大量煤尘导致小目标特征提取困难等问题，将 SA 算法与 k-means++聚类算法融合，在 YOLOv7 骨干网络中增加新的目标检测层，同时将双层注意力机制嵌入聚合网络模块之后。YOLOv7-SE 网络模型安全帽检测 AP、自救器检测 AP、mAP 分别为 72.50%，64.48%，68.49%，较 YOLOv7 网络模型分别提升了 12.2%，7.38%，9.79%，FPS 为 61.84 帧/s。

2) 将 YOLOv7-SE 网络模型与 Faster R-CNN, RetinaNet, CenterNet, FCOS, SSD, YOLOv5, YOLOv7 进行对比，实验结果表明，YOLOv7-SE 网络模型对安全帽和自救器的检测精度最高。

3) 在对安全帽和自救器的检测中，YOLOv7-SE 网络模型有效改善了漏检和误检问题，提高了检测精度。

#### 参考文献(References):

- [1] 许鹏飞. 2000—2021 年我国煤矿事故特征及发生规律研究[J]. 煤炭工程, 2022, 54(7): 129-133.  
XU Pengfei. Characteristics and occurrence regularity of coal mine accidents in China from 2020 to 2021[J]. Coal Engineering, 2022, 54(7): 129-133.
- [2] LOWE D G. Distinctive image features from scale-invariant keypoints[J]. International Journal of Computer Vision, 2004, 60(2): 91-110.
- [3] DALAL N, TRIGGS B. Histograms of oriented gradients for human detection[C]. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, San Diego, 2005. DOI: 10.1109/CVPR.2005.177.
- [4] PLATT J. Sequential minimal optimization: a fast algorithm for training support vector machines [EB/OL]. (1998-04-21) [2023-07-05]. <https://api.semanticscholar.org/CorpusID: 577580>.
- [5] 杨锋, 丁之桐, 邢蒙蒙, 等. 深度学习的目标检测算法改进综述[J]. 计算机工程与应用, 2023, 59(11): 1-15.  
YANG Feng, DING Zhitong, XING Mengmeng, et al. Review of object detection algorithm improvement in deep learning[J]. Computer Engineering and Applications, 2023, 59(11): 1-15.
- [6] 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, Columbus, 2014: 580-587.
- [7] GIRSHICK R. Fast R-CNN[C]. IEEE International Conference on Computer Vision, Santiago, 2015: 1440-1448.

- [8] REN Shaoqing, HE Kaiming, 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, 2017, 39(6): 1137-1149.
- [9] LIU Wei, ANGUELOV D, ERHAN D, et al. SSD: single shot multiBox detector[C]. The14th European Conference on Computer Vision, Amsterdam, 2016: 21-37.
- [10] 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, Las Vegas, 2016: 779-788.
- [11] REDMON J, FARHADI A. YOLO9000: better, faster, stronger[C]. IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, 2017: 7263-7271.
- [12] REDMON J, FARHADI A. YOLOv3: an incremental improvement[EB/OL]. (2018-04-08) [2023-07-05]. <https://arxiv.org/abs/1804.02767>.
- [13] BOCHKOVSKIY A, WANG C Y, LIAO H. YOLOv4: optimal speed and accuracy of object detection[EB/OL]. (2020-04-23) [2023-07-05]. <https://arxiv.org/abs/2004.10934>.
- [14] GE Z, LIU S, WANG F, et al. YOLOX: exceeding yolo series in 2021[EB/OL]. (2021-07-18) [2023-07-05]. <https://arxiv.org/abs/2107.08430>.
- [15] SUN Wei, DAI Liang, ZHANG Xiaorui, et al. RSOD: real-time small object detection algorithm in uav-based traffic monitoring[J]. Applied Intelligence, 2022, 52(8). DOI: [10.1007/S10489-021-02893-3](https://doi.org/10.1007/S10489-021-02893-3).
- [16] 翟国栋,任聰,王帥,等.多尺度特征融合的煤矿救援机器人目标检测模型[J].工矿自动化,2020,46(11):54-58.  
ZHAI Guodong, REN Cong, WANG Shuai, et al. Object detection model of coal mine rescue robot based on multi-scale feature fusion[J]. Industry and Mine Automation, 2020, 46(11): 54-58.
- 
- (上接第34页)
- [84] 邢震.数字孪生驱动的煤矿多元业务全局动态协同管控[J].工矿自动化,2023,49(7):60-66,82.  
XING Zhen. Global dynamic collaborative management and control of diversified business in coal mines driven by digital twins[J]. Journal of Mine Automation, 2023, 49(7): 60-66, 82.
- [85] 郭一楠,杨帆,葛世荣,等.知识驱动的智采数字孪生主动管控模式[J].煤炭学报,2023,48(增刊1):334-344.  
GUO Yinan, YANG Fan, GE Shirong, et al. Novel
- [17] 沈科,季亮,张袁浩,等.基于改进YOLOv5s模型的煤矿目标检测[J].工矿自动化,2021,47(11):107-111,118.  
SHEN Ke, JI Liang, ZHANG Yuanhao, et al. Research on coal and gangue detection algorithm based on improved YOLOv5s model[J]. Industry and Mine Automation, 2021, 47(11): 107-111, 118.
- [18] ZHANG Dawei, ZHENG Zhonglong, LI Minglu, et al. CSART: channel and spatial attention-guided residual learning for real-time object tracking[J]. Neurocomputing, 2021, 436(14): 260-272.
- [19] ZHANG Dawei, ZHENG Zhonglong, WANG Tianxiang, et al. HROM: learning high-resolution representation and object-aware masks for visual object tracking[J]. Sensors, 2020, 20(17). DOI: [10.3390/s20174807](https://doi.org/10.3390/s20174807).
- [20] ZHU Lei, WANG Xinjiang, KE Zhanghan, et al. BiFormer: vision transformer with bi-level routing attention[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, Vancouver, 2023: 10323-10333.
- [21] WANG Chenyao, BOCHKOVSKIY A, LIAO H. YOLOv7: trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, Vancouver, 2023: 7464-7475.
- [22] 曹鹏.基于模拟退火算法的大螺旋钻机自主钻进控制系统研究[J].煤矿机械,2023,44(10):194-196.  
CAO Peng. Research on autonomous drilling control system of large screw drilling rig based on simulated annealing algorithm[J]. Coal Mine Machinery, 2023, 44(10): 194-196.
- [23] HE Kaiming, SUN Jian, TANG Xiao'ou. Single image haze removal using dark channel prior[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2011, 33(12): 2341-2353.
- 

knowledge-driven active management and control scheme of smart coal mining face with digital twin[J]. Journal of China Coal Society, 2023, 48(S1): 334-344.

- [86] 尤秀松,葛世荣,郭一楠,等.智采工作面三机数字孪生驱动控制架构[J/OL].煤炭学报:1-11[2024-01-11]. <http://kns.cnki.net/kcms/detail/11.2190.TD.20230926.1756.004.html>.

YOU Xiusong, GE Shirong, GUO Yinan, et al. Digital twin-driven control construction for three machines of smart coal mining face[J/OL]. Journal of China Coal Society:1-11[2024-01-11]. <http://kns.cnki.net/kcms/detail/11.2190.TD.20230926.1756.004.html>.