

基于改进YOLOv8的矿井复杂环境小目标检测算法

彭森^{1,2}, 张珂珂³, 杨一群², 魏巍¹

1. 西北工业大学计算机学院, 陕西 西安 710129;
2. 北京东方国信科技股份有限公司, 北京 100012;
3. 联通在线信息科技有限公司, 北京 100031

摘要

针对煤矿井复杂环境中因图像灰暗模糊、低照明导致小目标难以检测和误检等问题, 提出了一种改进的YOLOv8模型用于煤矿井下安全帽佩戴检测, 以提高检测精度。首先, 在YOLOv8的主干网络中引入了SPD-Conv卷积层, 增强了小目标的特征提取能力。然后, 再将SPPF替换为SPPFCSPC模块, 进一步增强多尺度特征信息。最后, 在颈部网络中引入了注意力机制CoTAttention, 以提高对关键特征的关注度。在矿井下图像数据集CUMT-HelmeT上进行训练和测试的结果表明, 改进后的Model3模型相比原始YOLOv8n模型, 在F1、P、R、mAP50、mAP50-95这5个关键指标上分别提升了9.08%、17.7%、3.8%、3.5%、1.9%。这些改进显著增强了模型在低照明、灰暗模糊条件下对小目标的检测能力, 有效减少了漏检误检情况的发生。与YOLOv10-X模型相比, 改进后的YOLOv8-X模型在F1、P、R、mAP50、mAP50-95方面分别提升了6.23%、4.7%、7%、3.9%、1.8%, 进一步验证了改进模型在煤矿井下复杂环境中的检测性能优势。

关键词

YOLOv8; SPD-Conv; SPPFCSPC; 注意力机制

中图分类号: TP312

文献标志码: A

doi: 10.11959/j.issn.2096-0271.2025075

Small target detection algorithm in a mine complex environment based on improved YOLOv8

PENG Sen^{1,2}, ZHANG Keke³, YANG Yiqun², WEI Wei¹

1. School of Computer Science, Northwestern Polytechnical University, Xi'an 710129, China
2. Business-Intelligence of Oriental Nations Corporation Ltd., Beijing 100012, China
3. China Unicom Online Information Technology Co., Ltd., Beijing 100031, China

Abstract

To address the challenges of detecting small safety helmets in coal-mine tunnels—where images are dim, blurry and poorly lit—a refined YOLOv8 model is proposed. First, an SPD-Conv layer is inserted into the backbone to strengthen feature extraction for small objects. Next, the SPPF module is replaced with SPPFCSPC to enrich multi-scale features. Finally, the CoTAttention mechanism is introduced in the neck to focus on salient cues. Trained and tested on the

underground image set CUMT-HelmeT, the improved Model3 surpasses the original YOLOv8n by 9.08% F1, 17.7% Precision, 3.8% Recall, 3.5% mAP50 and 1.9% mAP50-95, markedly reducing missed and false detections under low-illumination and hazy conditions. Compared with YOLOv10-X, the enhanced YOLOv8-X further improves by 6.23% F1, 4.7% Precision, 7% Recall, 3.9% mAP50 and 1.8% mAP50-95, confirming its superiority in complex underground environments.

Key words

YOLOv8, SPD-Conv, SPPFCSPC, attention mechanism

0 引言

煤矿井下作业环境复杂,通常伴随着低照明、灰暗模糊等问题,使得安全帽佩戴检测面临诸多挑战。安全帽作为保护矿工头部安全的重要防护装备,其佩戴情况直接关系到矿工的生命安全。因此,煤矿井下环境中安全帽佩戴情况的自动检测具有重要意义^[1-4]。然而,在传统的检测方法中,由于图像质量差和小目标难以检测等问题,存在漏检和误检现象,难以满足实际应用需求^[5]。

近年来,随着深度学习和计算机视觉技术的发展,经典目标检测算法主要分为两类,即一阶段目标检测算法和两阶段目标检测算法。一阶段目标检测算法包括 YOLO^[6-9]、SSD^[10]、RetinaNet^[11]等。两阶段目标检测算法包括 R-CNN^[12]、Fast R-CNN^[13]、Faster R-CNN^[14]和 Mask R-CNN^[15]等。两阶段目标检测算法由于需要生成和处理大量的候选区域,计算复杂度较高,检测速度较慢。YOLO 系列模型因其高效的目标检测能力,无须生成候选区域,因而具有较快的检测速度,被广泛应用于各类检测任务中。YOLOv8 作为该系列的重大更新版本,在检测精度和速度上都有显著提升。然而,尽管 YOLOv8 在多尺度特征融合上有所改进,但其对小

目标的处理仍不够理想^[16]。小目标在特征图上占据的像素较少,在特征融合过程中容易被忽略或弱化。YOLO 模型在提取特征时,通过多次下采样会导致特征图分辨率较低^[17]。这使得小目标的特征信息可能丢失或难以被有效提取,从而降低了小目标的检测精度。

1 研究现状

矿井环境下的视频监控环境存在光照昏暗、分布不均、粉尘干扰等情况,同时矿井背景复杂多变,检测目标尺度各异,这些因素给安全帽的检测带来了很高的难度^[18]。刘晓慧等^[19]等使用 Hu 矩进行特征提取,它是一种形状特征描述符,能够捕捉图像中物体的形状特征,结合分析安全帽与头发之间的颜色差异,进而有效地判断工人是否佩戴了安全帽。Park 等^[20]提取了安全帽的方向梯度直方图(histogram of oriented gradient, HOG)特征,通过捕捉图像中边缘和纹理的方向分布,有效地描述了物体的形状和轮廓;随后,将这些提取到的特征输入支持向量机(support vector machine, SVM)模型中进行分类,区分是否佩戴了安全帽。Mneymneh 等^[21]利用图像中的特征点,提取工人与安全帽的空间特征,通过特征匹配算法,确定这些特征点在工人和安全帽

之间的对应关系。上述检测方法通常涉及大量的特征提取和匹配计算,导致其实时性较差,难以在现场推理环境中实现高效的实时检测。手工特征提取方法还依赖于特定的场景配置和拍摄角度,在更广泛的应用场景中,往往需要对检测算法进行调整和优化,限制了其通用性和适用范围。

为了克服这些问题,部分学者通过深度学习技术自动学习特征,在模型参数量较小的情况下提高检测精度和效率,并增强算法的适应性和鲁棒性。许凯等^[22]使用一种参数量较少的轻量化模型 YOLOv3 进行安全帽检测,通过增加网络层数来增强模型的特征提取能力,使其能够捕捉更多的细节和特征信息。与此同时,引入注意力机制有助于使模型聚焦于更有价值的特征,减少噪声和冗余信息对检测结果的干扰。但在实际应用中,特别是面对密集分布的场景时,仍然存在较多漏检的情况。杨永波等^[23]通过将 YOLOv5s 的初始网络替换为 MobileNetV3,实现了一种轻量化的安全帽检测算法。但 MobileNetV3 在减少计算量的同时,丢失了一些细节信息,导致检测精度下降,在处理光照变化、遮挡等复杂条件时,轻量化模型的鲁棒性通常较差。赵敏等^[24]在 YOLOv7 的基础上设计一种轻量化 Tiny-BiFPN 的特征金字塔网络,以此改进模型的特征融合能力。尽管 Tiny-BiFPN 结构在提升特征融合效率和减少模型参数量方面表现出色,但在某些特定情况下,如反光场景中,模型仍然存在漏检的情况。上述改进算法的检测精度仍然不足,背景特征提取仍不够丰富,未能应对矿井实际场景中的复杂性和多样性。YOLO 模型在矿井复杂条件下捕捉小目标的能力未能大幅改善,仍然容易出现漏检和误检的问题,凸显了现有算法在实际应用中的局限性。为了解决这些问题,本文

提出一种改进的 YOLOv8 算法,该算法能更好地应对煤矿井下复杂多变的环境,提高小目标的检测性能,减少漏检和误检的发生^[25-27]。这不仅有助于提高矿井安全生产管理的效率,还能在实际应用中提供更加可靠的安全保障。

2 YOLOv8 概述

YOLO 系列模型以其高效的实时目标检测能力著称,YOLOv8 是一种先进的目标检测模型,由 Ultralytics 公司开发和发布。相比于之前的 YOLOv5^[28],YOLOv8 引入了 C2f 模块,代替了 YOLOv5 中的 C3 模块。C2f 模块通过更丰富的特征连接方式,增强了梯度流,使特征提取过程更加高效和灵活。检测头被升级为解耦合头(decoupled head),这意味着分类和定位任务被分开处理时,可以更加专注地优化每个任务的损失函数,从而加快收敛速度。

在 YOLOv8 中,根据模型的网络深度和宽度的不同,划分为 5 种不同大小的模型结构。具体来说,这些模型从小到大依次为 n、s、m、l 和 x,这种划分的主要目的是满足不同计算资源和应用场景的需求。其中,n 是最小的模型,实时性高,具有最少的参数和计算量,为方便部署到嵌入式设备,选用 YOLOv8n 模型。YOLOv8 网络结构如图 1 所示。

3 改进的 YOLOv8 算法

3.1 SPD-Conv 模块

卷积神经网络(CNN)在许多计算机视觉任务(如图像分类和对象检测)中取得了巨大成功,但它们在处理特定问题时,

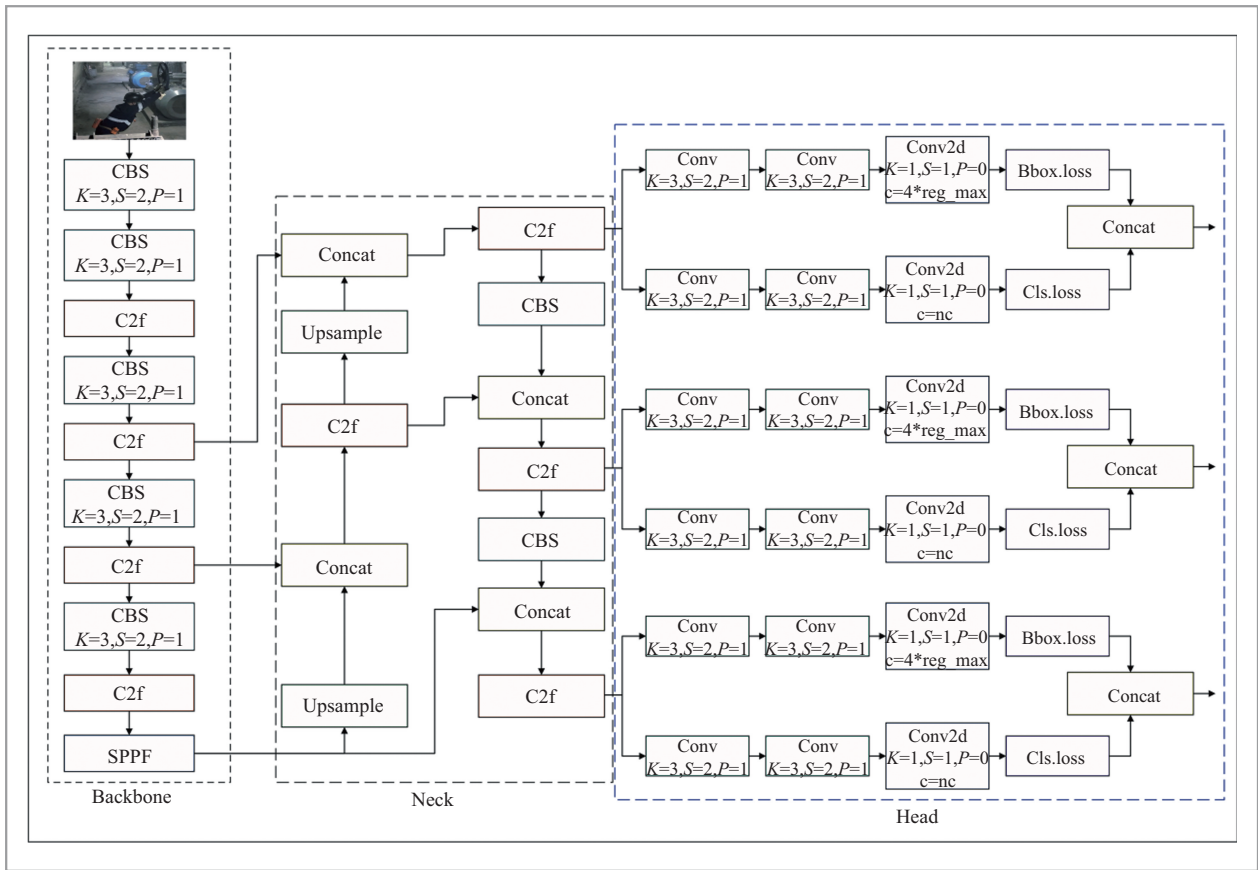


图1 YOLOv8网络结构

如低分辨率图像或小目标物体检测，效果会下降。这主要是因为CNN难以有效捕捉和保留细节信息。

为了解决上述问题，本文在YOLOv8的主干网络中引入SPD-Conv模块^[29]，该模块是将输入张量的空间维度转化为更高的通道数，然后在转换后的张量上执行标准的卷积操作，通过减少空间分辨率并增加通道数，可在保持或增加特征信息量的同时，减少特征图的空间大小，有助于捕捉多尺度特征^[30]。通过重新组织空间信息到通道的维度，模型能够更好地整合和处理不同尺度的特征，将空间信息折叠进通道，模型的表达能力得到增强^[31]。这种特征的丰富性可以帮助模型更好地学习和区

分复杂模式和细节，提高网络对小目标的检测能力。SPD-Conv模块结构如图2所示。

考虑任意大小为 $S \times S \times C$ 的中间特征映射 X ，特征映射 X 分割为一系列子特征映射，SPD组件按照一定的步长scale对特征图进行切分，将特征图分成一系列的子特征图。每个子特征图通过选取原始特征图中特定位置的元素来构成。子特征映射的分割方式如下：

$$f_{0,0} = X[0:S:\text{scale}, 0:S:\text{scale}], f_{0,1} = X[1:S:\text{scale}, 0:S:\text{scale}], \dots, \quad (1)$$

$$f_{\text{scale}-1,0} = X[\text{scale}-1:S:\text{scale}, 0:S:\text{scale}]; \quad (2)$$

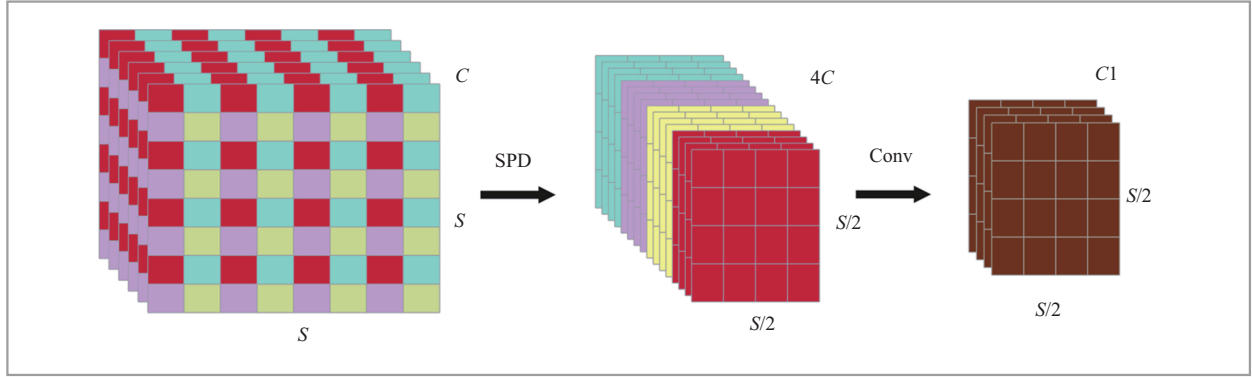


图2 SPD-Conv模块结构

$$f_{0,1} = X[0:S:\text{scale}, 1:S:\text{scale}], f_{1,1} = X[1:S:\text{scale}, 1:S:\text{scale}], \dots, \quad (3)$$

$$f_{\text{scale}-1,1} = X[\text{scale}-1:S:\text{scale}, 1:S:\text{scale}]; \quad (4)$$

$$f_{0,\text{scale}-1} = X[0:S:\text{scale}, \text{scale}-1:S:\text{scale}], f_{1,\text{scale}-1}, \dots, \quad (5)$$

$$f_{\text{scale}-1,\text{scale}-1} = X[\text{scale}-1:S:\text{scale}, \text{scale}-1:S:\text{scale}], \dots, \quad (6)$$

通常情况下，每个子特征图 $f_{x,y}$ 包含原始特征图中那些横纵坐标 (i,j) 满足 $i+x$ 和 $i+y$ 分别对 scale 整除的元素。具体来说，对于给定的步长 scale ，特征图 X 被分割成 $\text{scale} \times \text{scale}$ 个子特征图。每个子特征图的空间尺寸被缩小为 $1/\text{scale}$ ，但其通道数保持不变。

当 $\text{scale}=2$ 时，特征图被分割成4个子特征图 $f_{0,0}$ 、 $f_{0,1}$ 、 $f_{1,0}$ 、 $f_{1,1}$ ，每个子特征图的尺寸为 $(\frac{S}{2}, \frac{S}{2}, C)$ ，并且每个子特征图都对原始特征图进行了2倍的下采样。这些子特征图在通道维度上被拼接，形成新的特征图 X 。拼接后的特征图的空间维度减少为原来的 $\frac{S}{\text{scale}}$ ，而通道数增加为原来的 scale 的2倍。

3.2 CoTAttention 注意力机制

在传统的 Self-Attention 机制中，输入特征图的每个位置都会生成一个查询 (query) 和一个键 (key)，分别表示该位置的特征和与其他位置的关联^[32]。然而，这种机制的 Query-Key 关系是独立学习的，仅关注各自所在位置的特征，缺乏对全局上下文信息的有效建模。在视觉任务中，缺乏上下文信息的挖掘可能会导致特征表达能力的不足，从而限制模型性能的提升。为了克服这些问题，CoTAttention 模块将上下文信息的挖掘与自注意力学习结合到一个统一的结构中，如图3所示。

通过这种设计，CoTAttention 不仅能够捕捉特征之间的相互依赖关系，还能充分利用上下文信息，显著增强特征的表达能力。假设输入特征图 X 的尺寸为 $H \times W \times C$ ，其工作原理如下。

(1) 上下文信息编码

在 CoT Block 中，键 (Key)、查询 (Query) 和值 (Value) 仍由输入特征图 X 生成：

$$Q = XW_Q, K = XW_K, V = XW_V \quad (7)$$

不同于传统自注意力机制中使用 1×1 卷积对 Key 进行编码，CoTAttention 首先

在每个键的空间位置上应用 $k \times k$ 组卷积，覆盖一个局部的 $H \times W \times C$ 网格。这一卷积操作引入了邻域信息，使生成的上下文键 K 不仅包含自身特征，还反映了其邻域内的上下文关系。此时， K 的尺寸仍然为 $H \times W \times C$ ，但其特征表示更丰富。

(2) 局部注意力计算

CoT Block 的核心是基于查询特征 Q 和具有上下文信息的键特征 K 来学习局部注意力矩阵，而非传统的逐点 Query-Key 计算。这种方法在每个位置的 Query-Key 关系中加入了上下文指导，使注意力机制更具全局视角。

(3) 上下文增强的特征表达

在上下文信息的辅助下，CoTAttention 能够捕捉到更广泛的像素间依赖关系，同时增强目标区域的特征表达。通过将卷积提取的上下文信息 K 融合到 Value 中，特征图 V 的更新不仅包含 Query-Key 注意力得分，还能显式体现局部邻域的上下文信息。

3.3 SPPFCSPC 模块

在 YOLOv8 的主干网络中，SPPF (spatial pyramid pooling feature) 模块结构相对简单，采用固定大小 5×5 的卷积核多尺度池化。该结构虽然能够捕获多尺度特征，但因为没有进一步的特征融合操

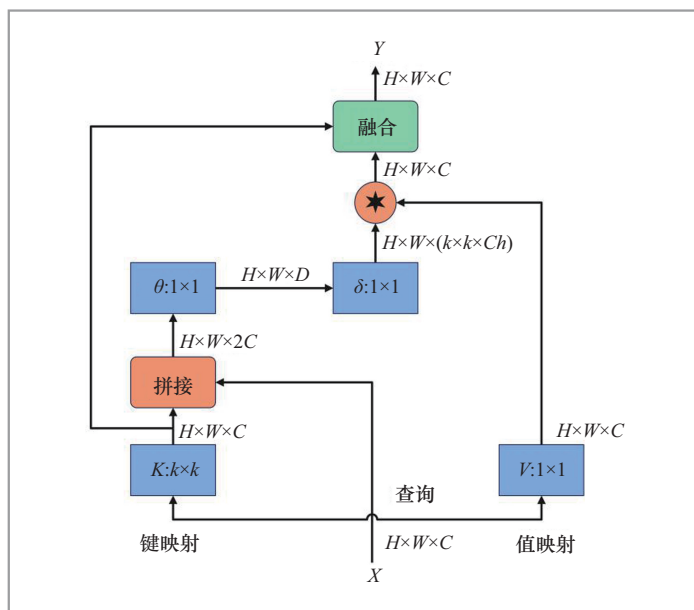


图3 CoT Block网络结构

作，未能引入跨阶段部分连接（cross-stage-partial-connections, CSP）来进行特征的重新融合和增强，因此在处理复杂特征时表现不佳。

为了解决原始YOLOv8中SPPF模块在复杂特征处理中的不足，本文将其替换为SPPFCSPC模块，如图4所示。该模块通过引入更多卷积层和跨阶段连接，提升了特征提取与融合能力，从而显著提高了模型的准确性和鲁棒性。SPPFCSPC模块首先对输入特征图进行3层卷积操作，内核尺寸分别为 1×1 、 3×3 和 1×1 。第一层

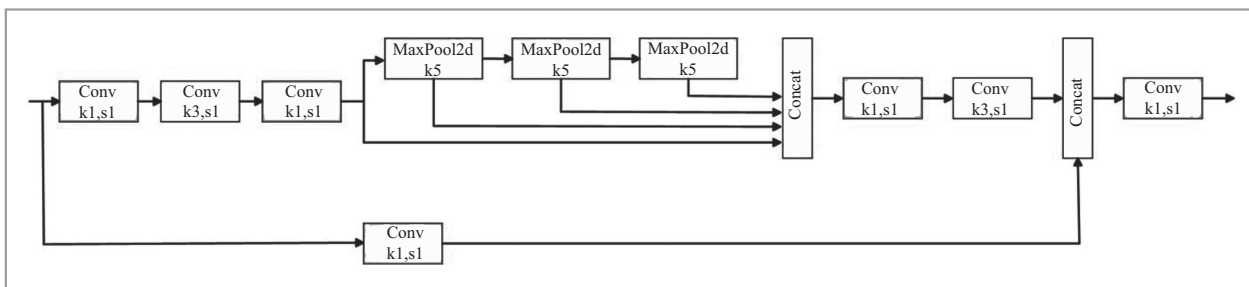


图4 SPPFCSPC网络结构

卷积将通道数减半,降低计算复杂度,并压缩冗余信息;第二层卷积在不改变通道数的前提下,向下取整缩小特征图的空间尺寸,提取更具代表性的多尺度特征;第三层卷积保持特征图的通道数与尺寸不变,进一步融合通道信息与空间信息^[33]。这3层卷积在保持空间尺寸的同时,有效提取了多尺度、多通道的特征,为后续的池化和特征融合操作提供了丰富的特征表达。

卷积后的特征图依次通过3层最大池化操作,每层池化的内核尺寸为 5×5 ,步长为1,填充为内核尺寸的一半。这样处理后,池化输出的特征图尺寸与输入特征图尺寸保持一致,确保信息不丢失^[34]。此外,在每层最大池化操作之前,特征图会

通过一条分支路径直接保留下来并被输送至特征融合层。这种设计保留了每层特征图的空间信息,同时增加了特征中的通道和上下文信息,为后续特征融合提供了更加全面的表达能力。在完成多尺度池化操作后,特征图经过尾部多层卷积操作,将通道数还原至初始状态。最终,这些卷积后的特征图与原始输入特征图在特征融合层中进行整合,输出更丰富、更精确的特征表示^[35]。

3.4 改进的YOLOv8模型结构

本文提出了一种改进的YOLOv8算法,如图5所示。针对复杂场景下的目标检测

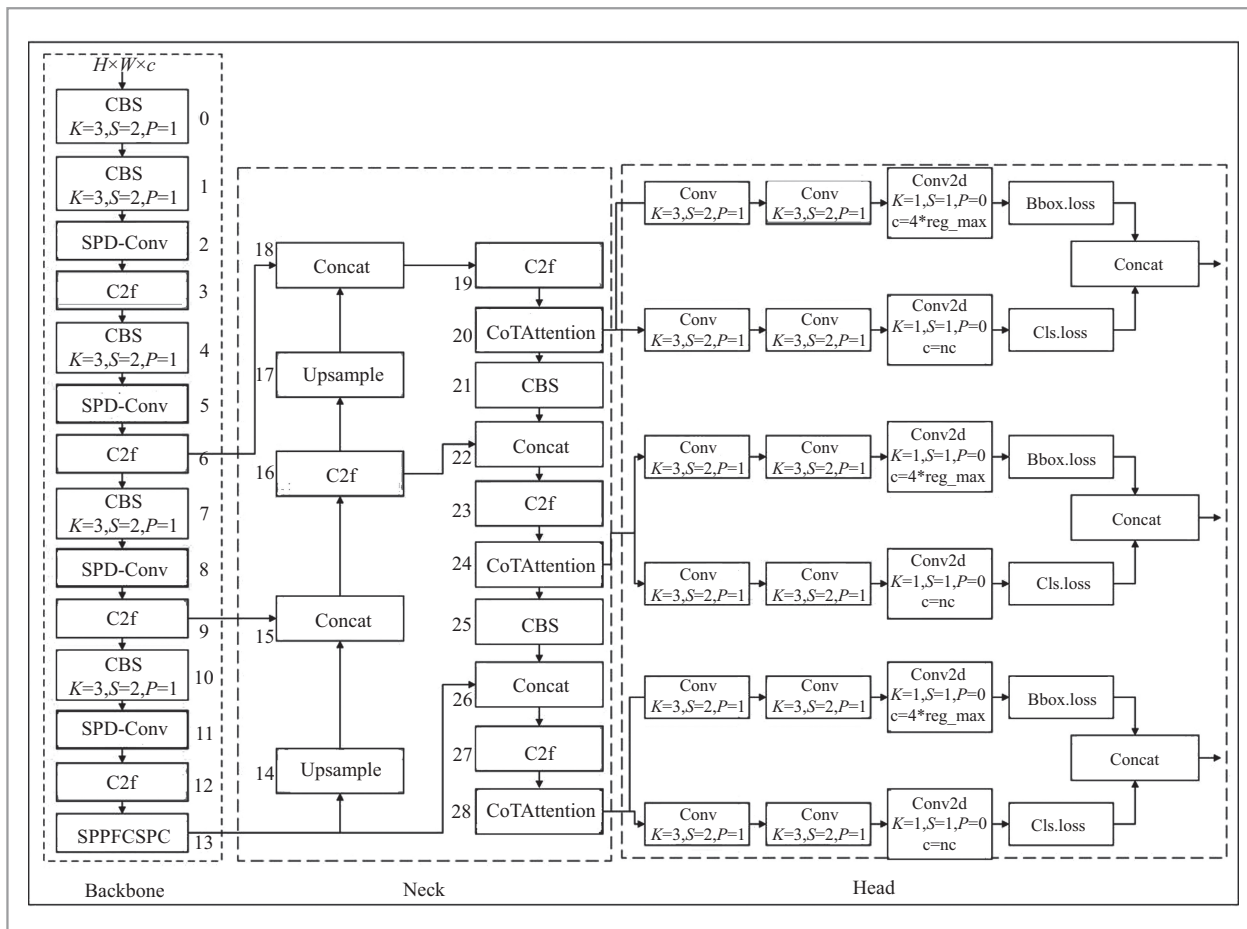


图5 改进的YOLOv8网络结构

任务，在模型的主干网络、颈部结构等关键模块进行了优化设计，显著提升了检测性能。改进的核心内容包括引入 SPD-Conv 模块、SPPFCSPC 模块以及 CoTAttention 模块，具体描述如下。

(1) SPD-Conv 模块

SPD-Conv 模块将特征图的空间维度转换为通道维度，重新排列特征图中的每个像素块，使特征图在保持总信息量不变的情况下，空间分辨率降低、通道深度增加。例如，对于一个形状为 (H, W, C) 的特征图，进行 2×2 的 SPD-Conv 操作后，输出特征图的形状将变为 $(H/2, W/2, C \times 4)$ 。这种操作能够压缩特征图的空间信息，同时通过增加通道深度，使每个通道捕捉到不同的细节特征，从而增强模型的判别能力。在第 2、5、8 和 11 层分别插入 SPD-Conv 模块，使模型能够在不同尺度上处理特征信息，捕捉多尺度目标的特征，特别适合检测小目标。该设计显著提升了模型对多尺度物体的适应能力和检测效果。

(2) SPPFCSPC 模块

在主干网络中将原 YOLOv8 的 SPPF 模块替换为 SPPFCSPC 模块。该模块通过对特征的分离和重新融合，更好地保留和整合了不同层次的特征，增强了特征的多样性和丰富性。这种改进提高了模型的特征表达能力，使其能够提取更丰富、更细致的特征，从而在多目标检测任务中表现出更高的检测精度和鲁棒性。

(3) CoTAttention 模块

在模型的颈部结构中，特别是在第 20、24 和 28 层，引入了 CoTAttention 模块。该模块在特征图处理过程中，重点关注包含目标物体的重要区域，同时忽略不相关的背景信息，提高了模型对目标区域的敏感度，显著减少了误报的概率。在检测包含多种物体的复杂场景中，

CoTAttention 模块通过增强特征图的语境信息，使模型能够更好地理解目标物体与周围环境的关系，从而更准确地识别和区分不同物体。此外，CoTAttention 模块在特征融合过程中也发挥着关键作用。通过加强多尺度特征的表达能力，该模块能够确保在特征融合过程中保留更多的有用信息，从而更有效地捕捉不同大小目标的特征。

4 实验结果与分析

4.1 实验环境

实验平台的配置如下：操作系统为 Linux，搭载 Intel(R) Xeon(R) Silver 4114 CPU 处理器，主频为 2.20 GHz，并配备 Tesla V100-SXM2-32GB 显卡，提供强大的图形处理能力，实验使用的深度学习框架是 Pytorch 2.0.1，CUDA 11.8 实现计算加速。

实验的参数配置如下：选择 16 作为每次训练的批次大小，初始学习率设为 0.01，训练过程将进行 500 次迭代，采用了随机梯度下降 (stochastic gradient descent, SGD) 作为优化算法，动量设为 0.937，权重衰减为 0.0005。

4.2 实验数据集

本文采用数据集 CUMT-HelmeT 用于网络模型训练，数据集中的图像采集自矿井地面和地下的实际场景。这些图像通常包括工人在不同工作环境中佩戴和不佩戴安全帽的情况。由于煤矿场景中的目标相对单一，为了提高数据的多样性和复杂性，数据集不仅包含煤矿场景的图像，还包括从其他安全帽数据集中筛选并处理后的图

片, 这些处理旨在模拟更加复杂的工作环境, 以提高模型的泛化能力。CUMT-HelmeT 数据集在打乱顺序后, 按照 7 : 2 : 1 的比例随机划分为训练集、验证集和测试集。

4.3 评价指标

本文为验证改进 YOLOv8 模型的性能, 采用多个指标进行评估, 这些指标包括精确率 (P)、召回率 (R)、平均精度 (AP)、平均精度均值 (mAP)、F1 分数。

(1) 精确率 (P)

P 为正确检测到的目标占有所有检测到的目标的比例。它衡量了模型的准确性, 即检测出的目标中有多少是正确的, 计算式如下:

$$P = \frac{TP}{TP + FP} \times 100\% \quad (8)$$

(2) 召回率 (R)

R 表示模型能够正确识别出所有实际存在的正类样本中的比例, 计算式如下:

$$R = \frac{TP}{TP + FN} \times 100\% \quad (9)$$

(3) 平均精度 (AP)

AP 是精确率和召回率在不同阈值下的加权平均, 通常表示为精确率-召回率曲线下的面积。AP 是单一类别的综合评估指标, 用于衡量模型在一个类别上的检测性

能, 计算式如下:

$$AP = \int_0^1 P(R) dR \quad (10)$$

(4) 平均精度均值 (mAP)

mAP 是所有类别的 AP 的平均值, 反映了模型在所有类别上的总体检测性能, 计算式如下:

$$mAP = \frac{1}{n} \sum_{i=1}^n AP(i) \times 100\% \quad (11)$$

(5) F1 分数

F1 分数是精确率和召回率的调和平均值, 计算式如下:

$$F1 = \frac{2 \times P \times R}{P + R} \times 100\% \quad (12)$$

4.4 消融实验

为了验证提出及设计模块的有效性, 本文以 YOLOv8n 为基准模型, 设计了 4 组实验, 分别对提出的 3 个改进模块进行不同组合下的效果验证和评估, 分析引入的模块对模型性能的提升效果。最后, 整合了所有改进模块, 得到了完整的改进模型 Model3, 并与基准模型进行性能对比。实验结果见表 1。

从表 1 可以看出, 在 YOLOv8n 基准模型中引入 SPD-Conv 模块后, Model1 的整体性能显著提升, F1、 P 、 R 、mAP50、mAP50-95 分别提升了 3.21%、3.6%、2.9%、1.9%、0.6%。在此基础上, 引入

表1 消融实验结果

Model	SPD-Conv	CoTAttention	SPPFCSPC	F1	P	R	mAP50	mAP50-95
YOLOv8n	—	—	—	53.61%	58.6%	49.4%	52.4%	31.6%
Model1	✓	—	—	56.82%	62.2%	52.3%	54.3%	32.2%
Model2	✓	✓	—	61.77%	74.8%	52.6%	55.0%	33.2%
Model3	✓	✓	✓	62.69%	76.3%	53.2%	55.9%	33.5%

CoTAttention 注意力机制, Model2 性能进一步提升, $F1$ 、 P 、 R 、 $mAP50$ 、 $mAP50-95$ 分别提升了 4.95%、12.6%、0.3%、0.7%、1.0%。最后, 在 Model2 的基础上再引入 SPPFCSPC 模块, Model3 性能继续提升, $F1$ 、 P 、 R 、 $mAP50$ 、 $mAP50-95$ 分别提升了 0.92%、1.5%、0.6%、0.9%、0.3%。可见, SPD-Conv 和 CoTAttention 模块对 Precision 和 $F1$ 分数的提升尤为显著。通过对比可知, 应用了 3 个模块优化后的 Model3 性能最佳, 体现了这些模块在提高模型性能方面的有效性。

从图 6 可以更加直观地看出 Model3 和 YOLOv8n 在 $mAP50$ 和 $mAP50-95$ 的评估指标上的对比。YOLOv8n 和 Model3 的 $mAP50$ 随着 epoch 的增加均在逐渐上升, 说明两个模型的检测精度都在不断提高, 但在整个训练过程中, Model3 的 $mAP50$ 、 $mAP50-95$ 一直高于 YOLOv8n。而且, 在大约 25 个 epoch 之后, Model3 的表现开始超过 YOLOv8n。在 200 个 epoch 之后, 性能曲线都趋于平稳, 说明模型已经逐渐收敛, 并达到了各自的最佳表现。可见 Model3 在检测精度上显著优于 YOLOv8n。

为了更加深入地理解改进前后的模型对不同类别的提升程度, 图 7 (a) Model3 在 “people” “Carry objects” “helmet” 和 “No-helmet” 4 个类别上均优于 YOLOv8n 模型, 尤其是在 Carry objects 和 No-helmet 这两个类别上表现非常显著。图 7 (b) 在 $mAP50-95$ 的比较中, Model3 在所有类别上的表现同样优于 YOLOv8n 模型, 特别是在 “No-helmet” 类别中, 提升更加显著, Model3 在 “No-helmet” 类别上的 $mAP50-95$ 接近 43%, 远高于 YOLOv8n。

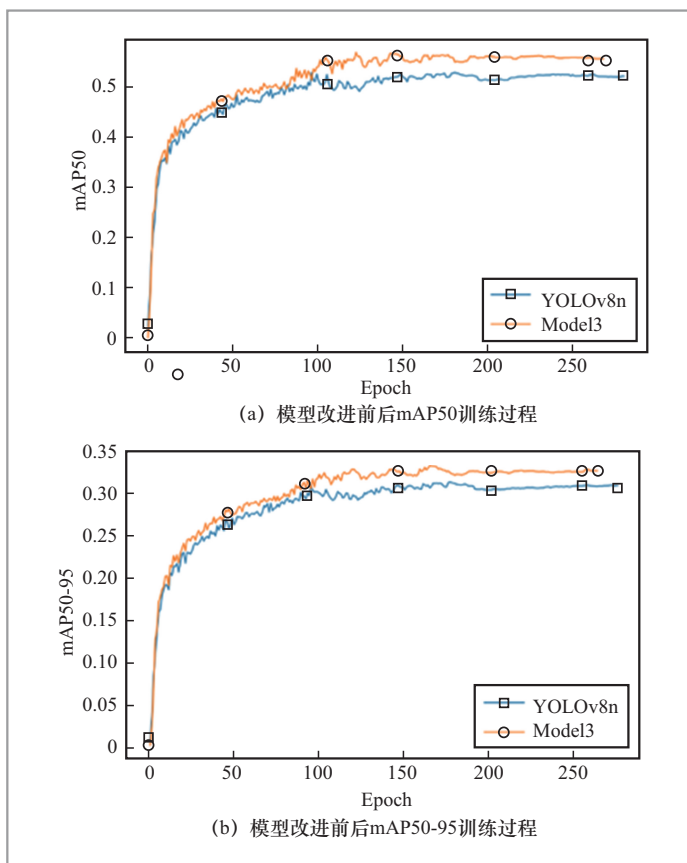


图 6 模型改进前后 $mAP50$ 和 $mAP50-95$ 对比

从图 8 可直观感受模型改进前后检测性能对比。图 8 (a) 和图 8 (c) 为煤矿井下场景, 检测对象处于低照明、灰暗模糊条件, YOLOv8n 模型将安全帽错误地识别为 “people”, 未能正确检测到 “helmet”。改进后的 Model3 能够成功检测到工人的安全帽, 将其标记为 “helmet” 且置信度为 0.81, 显示出 Model3 对小目标的检测效果有显著提升, 减少了漏检现象。图 8 (b) 和图 8 (d) 为误修正场景, YOLOv8n 模型误将旁边的白色物体检测为 “helmet”, 出现了误检。Model3 模型有效消除了这个误检, 仅检测到实际存在的安全帽。

为了分析煤矿数据集在不同任务场景

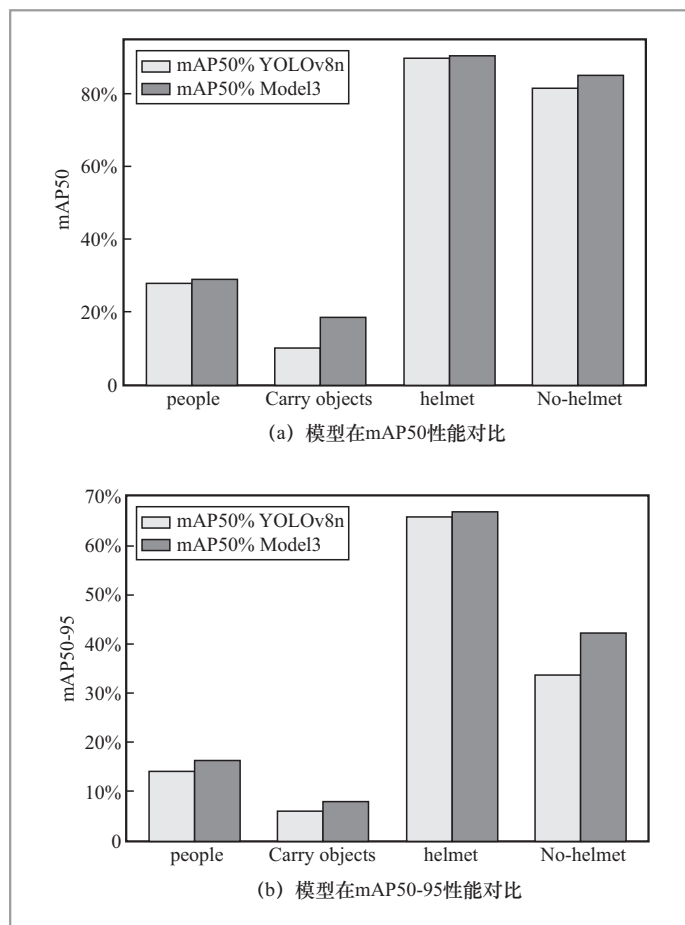


图7 模型改进前后不同类别的性能对比

下的检测情况, 本文选取某操作室和某变电所两个典型场景进行对比分析, 如图9所示。图9上半部分为YOLOv8n的检测结果, 下半部分为改进后的Model3检测结果。在图9(a)和图9(c)操作室场景中, 环境较为明亮, 背景干扰较少。两种模型均能够正确检测出目标对象, 但Model3的检测置信度显著高于YOLOv8n, 具体从0.32提升至0.73, 提升了0.41。这表明, 在相对简单的场景中, 改进后的Model3依然表现出更高的检测精度和可靠性。在图9(b)和图9(d)变电所场景中, 环境相对复杂, 低照明且灰暗模糊, YOLOv8n受背景干扰影响较大, 检测置信度明显下降。而Model3在相同场景下, 依然能够保持较高的检测置信度, 准确检测出目标。人员检测置信度具体从0.75提升至0.86, 提升了0.11; 安全帽检测置信度具体从0.61提升至0.73, 提升了0.12。这充分体现了Model3在复杂场景下的优势, 其特征提取和目标区分能力得到了显著提升。

4.5 不同模型实验结果

为了横向验证改进模型的性能, 将改进模型与YOLOv5、YOLOv6、YOLOv8、YOLOv10这4种模型进行对比, 使用相同的数据集进行训练和测试, 实验结果见表2。

从表2可以看出, 本文模型在各个版本中整体表现均优于YOLOv5、YOLOv6、YOLOv8和YOLOv10。在所有F1得分、召回率、mAP50和mAP50-95的综合指标上, 本文模型均表现出色, 尤其在F1得分上, 本文模型在所有版本中都取得了最高分, 特别是在大型模型(X)中达到了64.81%, 远高于YOLOv10-X的58.58%, 表明本文模型在精度和召回率之间取得了

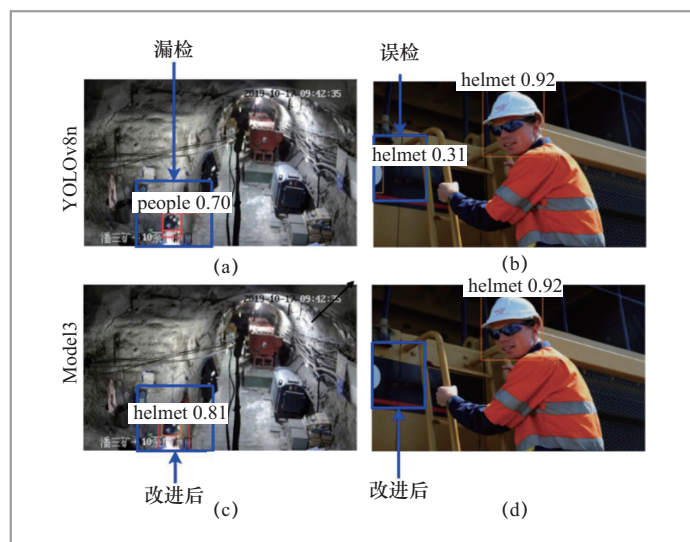


图8 YOLOv8n和Model3模型检测结果

最佳平衡。同时，本文模型在 mAP50 和 mAP50-95 两个关键指标上也表现突出，本文模型的 mAP50 达到了 59.2%，mAP50-95 达到了 36.0%，进一步显示了本文模型在检测精度上的优势。虽然 YOLOv10 在某些版本中精度表现突出，例如 YOLOv10-L 的精度达到 82.2%，但总体来看，本文模型在 F1 得分和其他指标上的平衡性更好，综合性能更强。本文模型在所有版本中召回率普遍较高，尤其是在大型模型中，召回率达到 58.4%，这表明

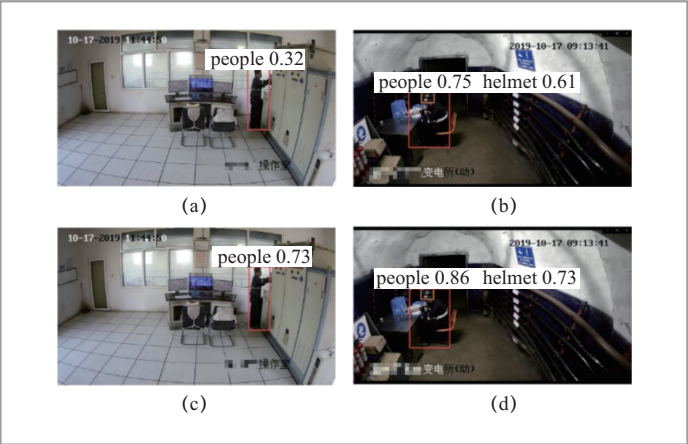


图9 不同场景的检测结果

表2 不同模型的实验结果

模型	F1	P	R	mAP50	mAP50-95
YOLOv5-N	58.29%	71.3%	49.3%	53.5%	32.0%
YOLOv6-N	55.18%	61.1%	50.3%	51.9%	31.1%
YOLOv8-N	53.61%	58.6%	49.4%	52.4%	31.6%
YOLOv10-N	59.50%	78.5%	47.9%	53.4%	31.5%
本文模型	62.69%	76.3%	53.2%	55.9%	33.5%
YOLOv5-S	58.82%	71.0%	50.2%	55.3%	33.1%
YOLOv6-S	63.39%	80.7%	52.2%	56.1%	34.9%
YOLOv8-S	60.96%	71.2%	53.3%	56.1%	33.4%
YOLOv10-S	61.92%	77.4%	51.6%	55.6%	33.4%
本文模型	63.58%	76.3%	54.5%	57.5%	35.5%
YOLOv5-M	60.76%	63.2%	58.5%	57.7%	35.3%
YOLOv6-M	63.22%	78.1%	53.1%	56.4%	33.9%
YOLOv8-M	61.87%	72.6%	53.9%	56.7%	35.3%
YOLOv10-M	62.86%	76.0%	53.6%	56.2%	34.5%
本文模型	63.34%	76.8%	53.9%	57.8%	35.8%
YOLOv5-L	62.49%	77.6%	52.3%	58.0%	35.1%
YOLOv6-L	57.12%	63.5%	51.9%	53.1%	32.6%
YOLOv8-L	62.14%	71.4%	55.0%	57.8%	35.8%
YOLOv10-L	63.17%	82.2%	51.3%	56.7%	34.3%
本文模型	63.27%	76.8%	53.8%	57.4%	35.8%
YOLOv5-X	61.20%	69.3%	54.8%	57.4%	35.6%
YOLOv6-X	62.11%	81.7%	50.1%	53.7%	33.0%
YOLOv8-X	62.18%	72.2%	54.6%	56.6%	35.9%
YOLOv10-X	58.58%	68.1%	51.4%	55.3%	34.2%
本文模型	64.81%	72.8%	58.4%	59.2%	36.0%

本文模型在捕捉目标时具有更高的准确性和覆盖率。可见，本文模型在综合性能上优于YOLO系列的其他版本，验证了模型在复杂环境中的优越性能。

5 结束语

本文通过在YOLOv8模型中引入SPD-Conv卷积层、将SPPF替换为SPPFCSPC模块，以及在颈部网络中引入CoTAttention注意力机制，显著提升了模型对煤矿井下低照明、灰暗模糊环境中小目标的检测能力，满足了煤矿井下复杂环境的实际应用需求。改进后的YOLOv8模型通过提升检测精度和稳定性，可以在煤矿场景下加强对矿工规范佩戴安全帽的有效监管，具备较强的实际应用价值，为进一步提高矿井安全监控提供了有力支持，为煤矿井下安全生产提供了技术保障。

参考文献：

- [1] LIN B Y. Safety helmet detection based on improved YOLOv8[J]. IEEE Access, 2024, 12: 28260-28272.
- [2] DONG Z B, ZHANG Q. Safety helmet detection: adding attention mechanism to YOLOv8 to improve detection accuracy [C]//Proceedings of the 2024 7th International Conference on Artificial Intelligence and Big Data. Piscataway: IEEE Press, 2024: 448-454.
- [3] CHEN L Y, MAO Y H, ZHANG H Y, et al. Improved the detection algorithm of safety helmet wearing based on YOLOv8 [C]//Proceedings of the 2023 3rd International Conference on Electronic Information Engineering and Computer Communication. Piscataway: IEEE Press, 2023: 1-4.
- [4] FU Z B, LING J R, YUAN X P, et al. YOLOv8n-FADS: a study for enhancing miners' helmet detection accuracy in complex underground environments[J]. Sensors, 2024, 24(12): 3767.
- [5] LOU H T, DUAN X H, GUO J M, et al. DC-YOLOv8: small-size object detection algorithm based on camera sensor[J]. Electronics, 2023, 12(10): 2323.
- [6] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection[C]//Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2016: 779-788.
- [7] REDMON J, FARHADI A. YOLO9000: better, faster, stronger[C]//Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE Press, 2017: 6517-6525.
- [8] REDMON J, FARHADI A. YOLOv3: an incremental improvement[EB]. arXiv preprint, 2018, arXiv: 1804.02767.
- [9] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. YOLOv4: optimal speed and accuracy of object detection[EB]. arXiv preprint, 2020, arXiv: 2004.10934.
- [10] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot MultiBox detector[M]//Computer Vision - ECCV 2016. Cham: Springer International Publishing, 2016: 21-37.
- [11] WANG Y Y, WANG C, ZHANG H, et al. Automatic ship detection based on RetinaNet using multi-resolution Gaofen-3 imagery[J]. Remote Sensing, 2019, 11(5): 531.
- [12] BHARATI P, PRAMANIK A. Deep learning techniques: r-CNN to mask R-

- CNN: a survey[C]//Computational Intelligence in Pattern Recognition. Singapore: Springer, 2020: 657-668.
- [13] REN S 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, 2017, 39(6): 1137-1149.
- [14] GIRSHICK R. Fast R-CNN[C]//Proceedings of the 2015 IEEE International Conference on Computer Vision. Piscataway: IEEE Press, 2015: 1440-1448.
- [15] HE K M, GKIOXARI G, DOLL R P, et al. Mask R-CNN[C]//Proceedings of the 2017 IEEE International Conference on Computer Vision. Piscataway: IEEE Press, 2017: 2980-2988.
- [16] ZHANG L J, FANG J J, LIU Y X, et al. CR-YOLOv8: multiscale object detection in traffic sign images[J]. IEEE Access, 2023, 12: 219-228.
- [17] HU W C, HU Q, PI J Y, et al. Dense pedestrian detection algorithm based on multiscale feature fusion in YOLOv8[C]//Proceedings of the 2024 5th International Conference on Computer Vision, Image and Deep Learning. Piscataway: IEEE Press, 2024: 1362-1366.
- [18] BARIK A, SAHOO J K, MOHANTY S, et al. An autonomous helmet detection model in diversified application areas using DNN-YOLOv8 technique[C]//Proceedings of the 2024 IEEE International Students' Conference on Electrical, Electronics and Computer Science. Piscataway: IEEE Press, 2024: 1-7.
- [19] 刘晓慧, 叶西宁. 肤色检测和Hu矩在安全帽识别中的应用[J]. 华东理工大学学报(自然科学版), 2014, 40(3): 365-370.
- LIU X H, YE X N. Skin color detection and hu moments in helmet recognition research[J]. Journal of East China University of Science and Technology (Natural Science Edition), 2014, 40(3): 365-370.
- [20] PARK M W, ELSAFTY N, ZHU Z H. Hardhat-wearing detection for enhancing on-site safety of construction workers[J]. Journal of Construction Engineering and Management, 2015, 141(9): 04015024.
- [21] MNEYMNEH B E, ABBAS M, KHOURY H. Automated hardhat detection for construction safety applications[J]. Procedia Engineering, 2017, 196: 895-902.
- [22] 许凯, 邓超. 基于改进YOLOv3的安全帽佩戴识别算法[J]. 激光与光电子学进展, 2021, 58(6): 615002.
- XU K, DENG C. Research on helmet wear identification based on improved YOLOv3[J]. Laser & Optoelectronics Progress, 2021, 58(6): 615002.
- [23] 杨永波, 李栋. 改进YOLOv5的轻量级安全帽佩戴检测算法[J]. 计算机工程与应用, 2022, 58(9): 201-207.
- YANG Y B, LI D. Lightweight helmet wearing detection algorithm of improved YOLOv5[J]. Computer Engineering and Applications, 2022, 58(9): 201-207.
- [24] 赵敏, 杨国亮, 王吉祥, 等. 改进YOLOv7-tiny的安全帽实时检测算法[J]. 无线电工程, 2023, 53(8): 1741-1749.
- ZHAO M, YANG G L, WANG J X, et al. Improved YOLOv7-tiny algorithm for safety helmet real-time detection[J]. Radio Engineering, 2023, 53(8): 1741-1749.
- [25] ZHENG X L, ZOU J X, DU S, et al. Small target detection in refractive panorama surveillance based on improved YOLOv8[J]. Sensors, 2024, 24(3): 819.
- [26] QU S M, CUI C, DUAN J L, et al. Under-water small target detection under YO-

LOv8-LA model[J]. Scientific Reports, 2024, 14: 16108.

[27] XU W Y, CUI C, JI Y C, et al. YOLOv8-MPEB small target detection algorithm based on UAV images[J]. Heliyon, 2024, 10(8): e29501.

[28] WU W T, LIU H, LI L L, et al. Application of local fully Convolutional Neural Network combined with YOLO v5 algorithm in small target detection of remote sensing image[J]. PLoS One, 2021, 16(10): e0259283.

[29] LIU J W, LI Y C, NING T, et al. Football player identification based on YOLOv5 backbone and SPD-Conv[C]//Proc SPIE 12715, Eighth International Conference on Electronic Technology and Information Science (ICETIS 2023), 2023, 12715: 695-701.

[30] SUN R Y, FAN H, TANG Y J, et al. Research on small target detection algorithm for UAV inspection scene based on SPD-conv[C]//Proceedings of the Fourth International Conference on Computer Vision and Data Mining. [S.l.]: SPIE, 2024: 81.

[31] GU Z C, ZHU K, YOU S T. YOLO-SSFS: a method combining SPD-conv/STD/LIM-FPN/SIoU for outdoor small target vehicle detection[J]. Electronics, 2023, 12(18): 3744.

[32] GENG Q C, LIU H, GAO T L, et al. An ECG classification method based on multi-task learning and CoT attention mechanism[J]. Healthcare, 2023, 11(7): 1000.

[33] YANG B, ZHANG L N, HE J P. Research on rice disease recognition based on improved SPPFCSPC-G YOLOv5 network [J]. PLoS One, 2023, 18(12): e0295661.

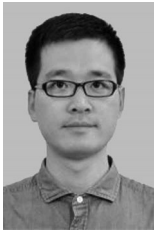
[34] YU G C, SHAO B R. Garbage classification and detection based on improved YOLOv7 network[C]//Proceedings of the 2023 International Conference on Pattern Recognition, Machine Vision and Intelligent Algorithms. Piscataway: IEEE Press, 2023: 103-107.

[35] WAN W Q, WANG L, WANG B B, et al. Space to depth convolution bundled with coordinate attention for detecting surface defects[J]. Signal, Image and Video Processing, 2024, 18(5): 4861-4874.

作者简介



彭森（1990-），女，西北工业大学计算机学院博士生，主要研究方向为计算机视觉、深度学习、多模态大模型等。



张珂珂（1985-），男，联通在线信息科技有限公司高级工程师，主要研究方向为深度学习、云化技术、通信网络等。



杨一群（1994-），男，北京东方国信科技股份有限公司人工智能算法工程师，主要研究方向为计算机视觉、多模态大模型等。



魏巍（1979-），男，博士，西北工业大学计算机学院教授、博士生导师，主要研究方向为计算机视觉、遥感图像处理、机器学习、人工智能等。

收稿日期: 2025-02-08
通信作者: 魏巍, weiweinwpu@nwpu.edu.cn
基金项目: 西安市科技计划重点产业链技术攻关项目(No.24ZDCYJSGG0003);2021年产业技术基础公共服务平台一建设面向区域化人工智能应用发展的公共服务平台项目(CEIEC-2021-ZM02-0166)
Foundation Items: Xi'an Science and Technology Plan Key Industrial Chain Technology Research General Item (No.24ZDCYJSGG0003), 2021 Industrial Technology Basic Public Service Platform – Construction of A Public Service Platform for the Development of Regional Artificial Intelligence Applications Project (No. CEIEC-2021-ZM02-0166)