詹辛慧, 姚瑞, 祁云, 等. 基于 HIC－YOLO 的井下安全帽检测算法研究[J]. 安全与环境学报, 2026, 26(6): 2151－2161.
ZHAN X H, YAO R, QI Y, et al. Research on underground safety helmet detection algorithm based on HIC－YOLO[J]. Journal of Safety and Environment, 2026, 26(6): 2151－2161.

基于 HIC－YOLO 的井下安全帽检测算法研究*

詹辛慧[1],姚　瑞[1],祁　云[2],薛凯隆[3],井雪艳[4],齐庆杰[5]

(1 山西大同大学煤炭工程学院,山西大同 037003;2 内蒙古科技大学矿业与煤炭学院,内蒙古包头 014010;
3 西安科技大学安全科学与工程学院,西安 710054;4 云南师范大学西南联合研究生院,昆明 650092;
5 煤炭科学研究总院有限公司应急科学研究院,北京 100013)

摘　要:针对煤矿井下光照不足、遮挡频发和目标尺度偏小导致的安全帽检测漏检、误检及实时性下降问题,研究提出基于 YOLOv5 的改进模型 HIC－YOLOv5。首先,在原始模型上增加高分辨率小目标检测头(Small Object Detection Head,SODH),以增强小目标特征提取能力;其次,在骨干网络与颈部网络之间引入通道特征融合模块(Channel Feature Fusion with Involution,CFFI),以缓解 1×1 降维造成的语义流失,强化跨通道信息交互,使更多关键信息传递至深层网络;最后,在骨干网络末端集成轻量级注意力机制(Convolutional Block Attention Module,CBAM),通过空间与通道双重加权突出安全帽区域特征。结果显示,在 IoU 为 0.5 的条件下,HIC－YOLOv5 的 mAP@50 达 88.33%,较基线 YOLOv5 提升 4.73 百分点,mAP@50:95 提升至 51.89%,P、R 分别为 90.81%、82.67%。在资源开销方面,参数量为 8.4 MB,FLOPs 为 30.6 GB,推理速度为 102 帧/s。消融试验结果显示,3 项改进具备协同增益,其中 SODH 对小目标召回提升最为显著。与 YOLOv7、YOLOv8、YOLOv9、YOLOv11、YOLOv12 对比,HIC－YOLOv5 模型综合性能保持领先。HIC－YOLOv5 在低照度、强遮挡与小目标密集的矿井安全帽检测场景中实现了更高的检测准确性与稳定性,并以中等计算量维持较好的实时性能,具备边缘部署可行性。

关键词:安全工程;矿井安全;安全帽检测;YOLOv5;小目标检测头;通道特征融合;注意力机制
中图分类号:X936　　**文献标志码**:A　　**文章编号**:1009-6094(2026)06-2151-11
DOI: 10.13637/j.issn.1009-6094.2025.1419

0　引　言

在矿山等高危作业场景中,安全帽的规范佩戴直接关系到作业人员的生命安全。传统人工巡检在大规模、复杂环境下效率低且易受主观因素影响,因此基于计算机视觉的自动安全帽检测技术逐渐成为安全管理的新趋势[1]。然而,井下场景光照昏暗、作业面杂乱、监控视角固定,且安全帽目标体积较小、易被遮挡,加剧了检测难度[2-3]。上述因素削弱了目标检测模型的鲁棒性与精度,导致安全帽检测中的漏检率与误检率较高。为此,开展面向井下场景的深度学习自动检测算法研究,对提升矿山安全管理水平具有重要意义。

近年来,单阶段目标检测算法(You Only Look Once,YOLO)由于检测速度快、效果好而广泛应用于安全帽检测等工业场景。为了进一步提高检测性能,学者提出了多种改进方向。2024 年,Huang 等[4]提出在 YOLOv7 框架中嵌入高效注意力模块与特征融合结构的概念,建立了轻量化的双层路由注意力机制融合模型(YOLOv7 with AFPN and Bi-level Routing Attention,FB－YOLOv7)模型,以增强对小目标的识别能力;2024 年,Lian 等[5]首次提出融合姿态特征的高分辨率单阶段目标检测模型(High-Resolution You Only Look Once,HR－YOLO),通过对象检测分支与姿态检测分支协同工作,结合 Laplace 感知注意力模块增强了多尺度特征融合能力,同时设计姿态辅助非极大值抑制(Pose-Assisted Non-maximum Suppression,PA－NMS)投票算法提升了遮挡目标的识别率;2023 年,Jakubec 等[6]最先提出一种基于改进 YOLOv7 的自动头盔识别模型 YOLOv7-AM(YOLOv7 with CBAM),即通过集成卷积块注意力模块增强小目标特征提取能力;2023 年,白培瑞等[7]提出了 DS－YOLOv5(YOLOv5 with Deep SORT)算法,即通过在主干网络中集成简化的 Transformer 模块并在颈部网络应用双向特征金字塔网络,加强了对图像全局信息和多尺度特征的捕获;2024 年,Xu 等[8]

* **收稿日期**: 2025－08－11
作者简介: 詹辛慧,讲师,博士,硕士生导师,从事矿山环境灾害控制方面的研究,zxhzhanxinhui@163.com;祁云(通信作者),副教授,博士,从事矿井灾害防治、矿山安全评价、应急技术与管理等研究,qiyun_sx@163.com。
基金项目:山西大同大学 2025 年研究生学术与实践创新项目(25SJCX23)

提出的 YOLOv8 - MPEB(YOLOv8 with MobileNetV3, EMA and BiFPN)模型采用 MobileNetV3 轻量骨干及专用小目标检测层,在特征融合模块中集成了高效多尺度注意力机制与双向特征金字塔网络,显著提升了在复杂场景下的检测精度;2025 年,Li 等[9]提出一种基于改进 YOLOv8 的矿工安全帽实时检测算法 MH - YOLO(YOLO with Multi-scale Feature Fusion Module),即通过集成卷积块注意力机制优化 C2f 模块特征提取能力,采用 MaxPooling 模块降低样本不平衡干扰,并引入小目标检测层及 ZoomCat - Scalseq 模块增强多尺度特征融合。此外,2020 年,Carion 等[10]首次提出端到端检测框架(DEtection TRansformer,DETR),该方法基于 Transformer 的全局特征建模思路,摆脱了传统的预设锚框和非极大值抑制等过程。总体而言,尽管上述方法均取得进展,但在井下低照度、复杂背景与超高实时性要求下,小目标检测精度、弱光鲁棒性与推理效率之间仍存在不足。

针对上述问题,本文提出 HIC - YOLOv5 模型,并在原始模型的核心结构上进行调整:一是增设小目标专用检测头(Small Object Detection Head, SODH),提供更高分辨率的特征图以实现精细定位;二是引入通道特征融合模块(Channel Feature Fusion with Involution,CFFI),在骨干网络与颈部网络之间强化通道信息交互,提升特征表达能力;三是采用轻量化注意力机制(Convolutional Block Attention Module,CBAM),在通道与空间维度上突出关键特征、抑制冗余信息。研究旨在增强模型对井下复杂背景与微小目标的适应性,即在兼顾检测精度与实时效率的前提下,为井下安全帽自动检测提供理论依据与技术支持。

1 HIC - YOLOv5 算法设计

1.1 HIC - YOLOv5 网络结构

YOLOv5 整体流程可以概括为输入图像首先经骨干网络提取多尺度特征,然后经颈部网络融合后由多个检测头输出检测框与类别置信度。骨干网络采用 CSPDarknet53(Cross Stage Partial Darknet - 53),其中包含 CBS(Convolution - BatchNorm - SiLU)、C3(CSP Bottleneck with 3 Convolutions)和 SPPF(Spatial Pyramid Pooling - Fast)等模块,以增强特征提取能力,保持计算效率。颈部网络部分使用 PANet(Path Aggregation Network)进行自上而下和自下而上的多尺度融合[11]。

在此基础上构建的 HIC - YOLOv5 仍属于单阶段检测框架,即围绕井下环境中的低照度、小目标与遮挡难题提出 3 项关键改进。首先,针对传统的 YOLOv5 模型在井下小目标检测中易漏检的问题,增加 SODH 小目标检测分支,在原有 80 像素 ×80 像素、40 像素 ×40 像素、20 像素 ×20 像素 3 个尺度预测的基础上,增加一条 160 像素 ×160 像素高分辨率特征图的输出分支,即直接连接颈部网络末端特征,经 1 ×1 卷积进行通道融合后生成面向小目标的锚框预测。其输出的高分辨率特征图可使安全帽在特征图上占据更多像素,增强微小目标的可分辨性和定位精准度。其次,在骨干网络与颈部之间插入 CFFI 模块,即原始 YOLOv5 在颈部网络起始处对骨干网络输出特征做 1 ×1 卷积以减少通道数,易造成关键信息流失,而 CFFI 模块先将骨干网络末端的 1 024 通道、20 像素 ×20 像素的特征图作为输入,通过位置专属、通道共享的核生成方式自适应分配空间权重并进行邻域聚合,再结合 1 ×1 卷积融合以弥补通道压缩导致的信息缺失,进而以可变感受野增强局部建模与通道间信息交互,进一步提升特征表达能力。最后,在骨干网络末端集成轻量级注意力机制 CBAM 模块,即先对输入特征在通道维度进行全局平均与最大池化以生成通道注意力图,并与原特征相乘,随后在空间维度进行平均与最大池化并拼接,经卷积与 Sigmoid 函数得到空间注意力图再次加权。该设计在 20 像素 ×20 像素特征图上运行,额外开销占比极小,能有效突出安全帽区域特征、抑制背景与阴影干扰,从而强化低照度与遮挡场景下的鲁棒性。上述改进模块均采用轻量化设计,在保持 YOLOv5 原有实时检测性能的同时,显著提升了小目标识别能力[12]。改进模型结构见图 1。

1.2 小目标检测头 SODH

在目标检测任务中,检测头作为模型的输出模块,负责对候选区域进行分类与边界框回归。在 YOLOv5 中,检测头采用基于锚框的多尺度预测机制,旨在精准覆盖不同尺寸的目标。具体而言,YOLOv5 在特征金字塔网络与路径聚合网络的基础上,构建 3 个尺度的输出特征图,分别对应小、中、大目标。各尺度的检测头均通过 1 ×1 卷积将通道数调整为指定的输出维度,并对每个锚框的边界框参数、置信度与类别概率进行回归操作。通过联合优化分类损失与回归损失,检测头能够有效捕获目标的空间位置与语义信息,从而实现端到端的高效目标检测[13]。

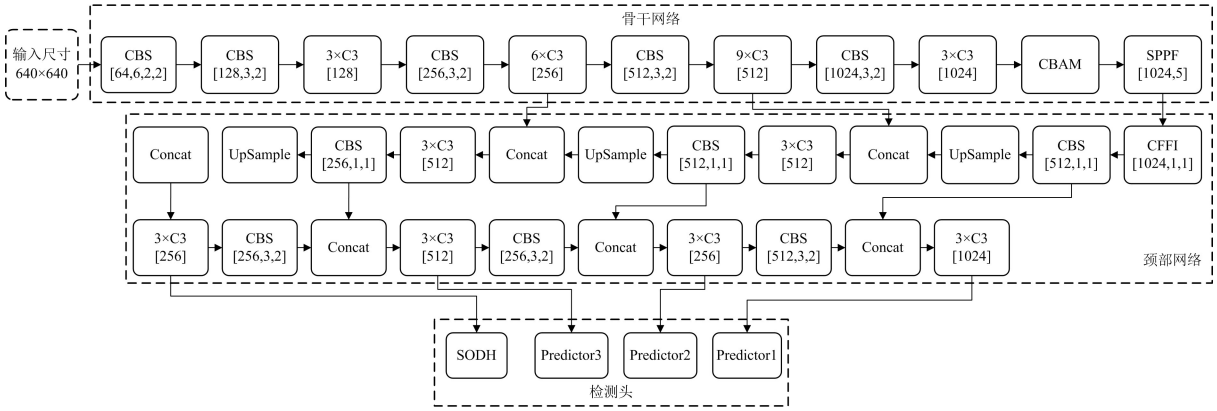


图 1　HIC－YOLOv5 结构

Fig. 1　HIC－YOLOv5 structure

为了进一步提升检测头对多尺度,尤其是小目标的识别能力,HIC－YOLOv5 在原有 YOLOv5 检测头的基础上引入 SODH 模块。该模块通过融合动态卷积、注意力机制与跨尺度信息融合策略,从结构层面增强了模型在复杂背景下对不同尺度目标特征的建模能力[14]。在特征尺度方面,SODH 新增一条 160 像素×160 像素的高分辨率检测分支,从而在特征图中保留更多微小目标信息,显著提升对远距或尺寸极小目标的感知与可分辨性。与之配套,改进模型将原有金字塔进一步扩展并与 FPN、PANet 协同工作,引入额外的浅层特征融合,使语义信息与空间细节在不同层级间形成更强互补。为了实现与目标尺度的精准匹配,SODH 采用 K－means 聚类生成的小尺寸锚框,在训练过程中联合优化分类与定位损失,使输出对微小目标的尺寸与形状更加贴合,进一步降低了漏检与误检风险。在复杂度控制方面,SODH 作为一支额外预测头,相较于引入大量 Transformer 或复杂注意力模块的方案[15],只带来较小的计算开销。整个 SODH 以残差形式嵌入 YOLOv5 的检测分支,具备良好的可插拔性与轻量化特征。在中低层特征融合阶段,小目标检测头能更有效地提取并强化小目标(如远处或局部遮挡的安全帽)信息,即在保持端到端实时性的前提下,显著提升检测精度与目标定位能力。面向矿井安全帽检测这类工人分布密集,姿态各异且风筒、电缆与照明光源等背景干扰突出的场景,SODH 模块能够有效缓解传统模型易出现的漏检与误检问题,体现出在低照度与遮挡条件下的稳健性。改进模型与原始模型检测头结构对比见图 2。

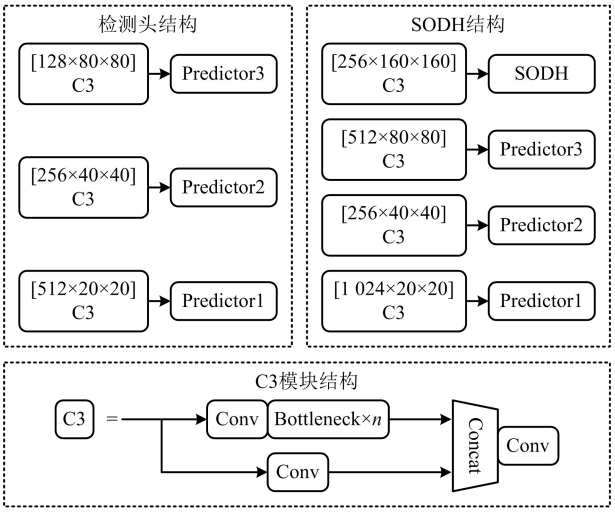


图 2　SODH 与 YOLOv5 检测头对比

Fig. 2　Comparison between SODH and YOLOv5 detection heads

1.3　通道特征融合模块

为了增强井下弱光、积尘、遮挡等复杂工况下的小目标表征,在 YOLOv5 的 Backbone 与 Neck 之间引入 CFFI 模块。原始 YOLOv5 的 Neck 在起始阶段采用 1×1 卷积降维提速,但会削弱通道信息并影响后续金字塔融合对小目标的刻画。为此,在 Backbone 与 Neck 的连接处加入 CFFI 模块,以减少信息损失、提升特征融合效果,这对小尺寸目标尤为有利,同时 CFFI 对不同空间位置的视觉模式具有更好的自适应性[16]。

改进模块的核心在于按空间位置自适应地生成卷积核并在通道分组内共享。设输入特征 $\boldsymbol{X} \in \mathbf{R}^{H\times W\times C}$,同时为每个位置$(i,j)$构造核张量 $\boldsymbol{H} \in \mathbf{R}^{H\times W\times K\times K\times G}$($G$ 取 1)。其中,K 为核尺寸,G 为分组数,组内通道共享同一位置核。输出的第 k 个通道在(i,j)的计算为

$$\boldsymbol{Y}_{i,j,k} = \sum_{(u,v)\in\Delta_K} \boldsymbol{H}_{i,j,u+\lfloor K/2\rfloor,v+\lfloor K/2\rfloor,\lceil kG/C\rceil} \boldsymbol{X}_{i+u,j+v,k} \quad (1)$$

式中 (u,v) 为局部邻域内的空间偏移量；Δ_K 表示以 (i,j) 为中心的 $K\times K$ 邻域的所有偏移坐标 (u,v)，$\Delta_K = \left\{-\frac{K-1}{2},-\frac{K-3}{2},\cdots,\frac{K-1}{2}\right\}^2$。该算子将单像素通道维度中的判别信息隐式扩散至其空间邻域，从而获得更丰富的有效感受域以为后续多尺度融合提供更厚实的特征底座。

基于上述机制，CFFI 在进入 Neck 之前即增强通道语义并向空间外扩，使得高、低层特征的自顶向下与自底向上融合更具判别力，进而提升小目标召回与定位稳定性。在总体结构层面，相较于直接重构 Neck 会显著增加计算量的方案，本文方法以轻量级模块完成增强，兼顾精度与效率。CFFI 模块结构见图 3。

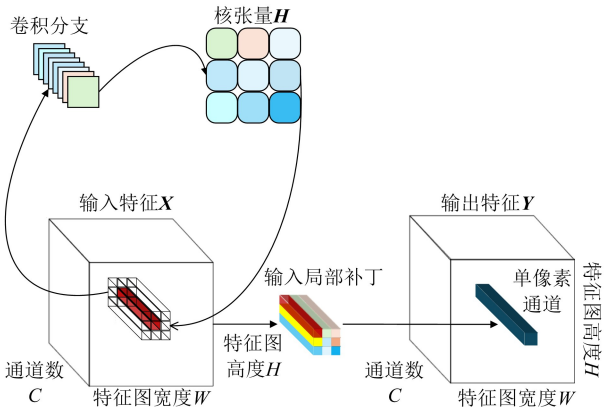


图 3 通道特征融合模块结构

Fig. 3 Structure of channel feature fusion with involution

1.4 轻量级 CBAM 注意力机制

现有研究大多在生成特征金字塔的 Neck 模块中集成 CBAM，但该模块所处理的特征图尺寸较大，导致模型参数量和计算开销显著增加，且训练难度加大[17]。为了解决该问题，本文将 CBAM 引入 Backbone 网络，即在骨干特征提取阶段突出关键信息，而不是在 Neck 阶段构建金字塔。此时，CBAM 的输入特征图尺寸仅为 20 像素 ×20 像素，约为 640 像素 ×640 像素特征图的 1/32，因此计算成本得到了有效控制。

CBAM 模块由通道注意力模块与空间注意力模块串联组成，可在两个正交维度上自适应地为关键特征分配更大的权重，从而增强特征表达的判别性[18]。对于中间特征图 $\boldsymbol{F}\in\mathbf{R}^{C\times H\times W}$，首先通过通道注意力机制对其进行加权，通道注意力计算为

$$M_c(\boldsymbol{F}) = \sigma((\mathrm{MLP}(\mathrm{AvgPool}(\boldsymbol{F})) + \mathrm{MLP}(\mathrm{MaxPool}(\boldsymbol{F})))) \quad (2)$$

式中 AvgPool 和 MaxPool 分别表示对输入特征图在空间维度上进行平均池化与最大池化操作；MLP 表示共享的多层感知机制；σ 表示 Sigmoid 激活函数。通过该模块生成的通道注意力权重 $M_c\in\mathbf{R}^{C\times1\times1}$，可用于增强显著通道特征，得到加权后的特征图 $\boldsymbol{F}'$，见式(3)。

$$\boldsymbol{F}' = M_c(\boldsymbol{F})\odot\boldsymbol{F} \quad (3)$$

之后，通过空间注意力机制进一步对 $\boldsymbol{F}'$ 进行空间维度加权，计算方式为

$$M_s(\boldsymbol{F}') = \sigma(f^{7\times7}([\mathrm{AvgPool}(\boldsymbol{F}');\mathrm{MaxPool}(\boldsymbol{F}')])) \quad (4)$$

式中 $f^{7\times7}$ 表示 7 × 7 卷积操作；(AvgPool($\boldsymbol{F}'$)；MaxPool($\boldsymbol{F}'$))表示通道维度上的拼接操作。最终，空间注意力输出 $M_s\in\mathbf{R}^{1\times H\times W}$ 与 $\boldsymbol{F}'$ 相乘，得到最终增强后的特征图为

$$\boldsymbol{F}'' = M_s(\boldsymbol{F}')\odot\boldsymbol{F}' \quad (5)$$

将 CBAM 模块嵌入 HIC - YOLOv5 模型的 Backbone 结构关键节点后，网络能够更加关注安全帽所在区域，特别是在人员密集、小目标及遮挡场景下，仍能保持较高的识别精度。CBAM 模块结构见图 4。

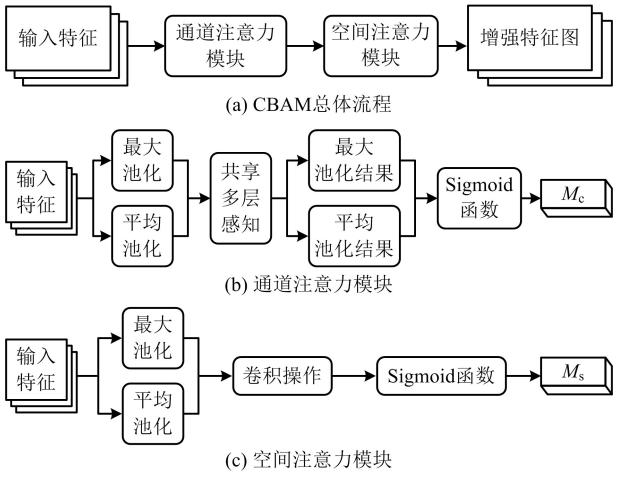


图 4 CBAM 模块结构

Fig. 4 CBAM module structure

1.5 损失函数

HIC - YOLOv5 在 YOLOv5 的基础上优化了其损失函数设计，并综合考虑了目标框定位、分类精度与目标存在性的置信度评估。具体而言，HIC - YOLOv5 的总损失函数 L_{total} 由三部分构成，见式(6)。

$$L_{total} = \lambda_{cls}L_{cls} + \lambda_{box}L_{box} + \lambda_{obj}L_{obj} \quad (6)$$

式中　L_{cls}表示分类损失,即采用改进的 Focal Loss 以缓解矿井中正负样本不平衡问题;L_{box}表示定位损失,即使用 CIoU Loss 来优化目标框与真实框之间的重叠度、中心距离与长宽比;L_{obj}表示置信度损失,即利用 Binary Cross-Entropy Loss 对目标存在性进行评估;λ_{cls}、λ_{box}、λ_{obj}分别为分类损失权重、定位损失权重及置信度损失权重[19]。

2　试验设计

2.1　数据集制作

选用中国矿业大学公开发布的 CUMT－HelmeT 数据集进行模型训练与验证。该数据集由井下煤矿的监控视频抽帧构成,共收集了约 10 000 张图像。本文按目标类别标注为佩戴安全帽和未佩戴安全帽两类,每类图像约 5 000 张。为了保证检测的可靠性,采用随机划分策略:将 90% 的样本(共 9 000 张,含各类别 4 500 张)作为训练集,10% 的样本(共 1 000 张,含各类别 500 张)作为测试集。数据集中的图像拍摄于不同井下环境,光照条件、摄像角度及场景布局多样,但部分图像中安全帽目标尺寸较小,且存在低光照、不清晰或遮挡等情况,易导致检测模型出现漏检和误检。部分数据集见图 5。

2.2　数据增强方法

在将 HIC－YOLOv5 应用于矿井安全帽检测任务时,针对地下复杂环境与多目标密集分布的特点,为了提升模型的泛化能力和鲁棒性,首先,在传统 YOLOv5 的 Mosaic 增强基础上,引入分层图像复合策略,按照矿井现场的光照与背景特征,将 4 幅原始图像分层拼接,以丰富目标的相对位置分布,同时保留矿井特有的低照度与高对比度场景信息;其次,采用实例级 Mix Up 增强,通过跨样本像素级插值融合,使模型更好地应对安全帽部分遮挡与多目标重叠情况;再次,结合随机仿射变换(包括旋转 ±15°、缩放 80% ~120%、平移 ±10%)与水平翻转,以模拟摄像头安装角度变化及巷道运动模糊情况;最后,在色彩空间(Hue-Saturation-Value,HSV)上施加亮度(±30%)、对比度(±20%)和色相(±10°)扰动,并加入高斯噪声与运动模糊,以再现矿灯不稳定照明与运输车辆运动时的视觉干扰。该数据增强方法能够显著提升 HIC－YOLOv5 在采掘面、输送巷道等典型矿井工况下对安全帽目标的召回率和精确度,实现了多目标实时检测的高可靠性[20]。

2.3　试验环境及参数配置

试验基于 PyTorch 1.8.1 构建深度学习网络框架,运行环境包括 Visual Studio Code、Python 3.8、CUDA 11.1 与 cuDNN 8.0.5,操作系统为 Windows 11。硬件平台配置为 Xeon(R) Gold6430 处理器、32 GB 内存及 NVIDIA GeForce RTX4090 显卡(显存为 24 GB)。为了确保试验结果的可比性与一致性,所有试验在相同的环境和超参数设置下进行,训练过程采用 SGD 优化器与余弦退火学习率调整策略。当连续 100 次迭代各项性能指标无明显变化时,训练过程自动终止。相关超参数设置详见表 1。

2.4　评价指标

为了全面评价 HICYOLOv5 在 CUMTHelmeT 煤矿安全帽检测任务中的性能,选取精确率(Precision,P)、召回率(Recall,R)、F_1作为指标进行评价,见式

图 5　部分数据集

Fig. 5　Part of the dataset

表 1 超参数设置

Table 1 Hyperparameter setting

参数	数值
图像尺寸	640 像素 ×640 像素
初始学习率	0.01
动量系数	0.937
批量处理	16
训练周期	150
权重衰减系数	0.005

(7) ~ (9)。

$$P = \frac{T_P}{T_P + F_P} \tag{7}$$

$$R = \frac{T_P}{T_P + F_N} \tag{8}$$

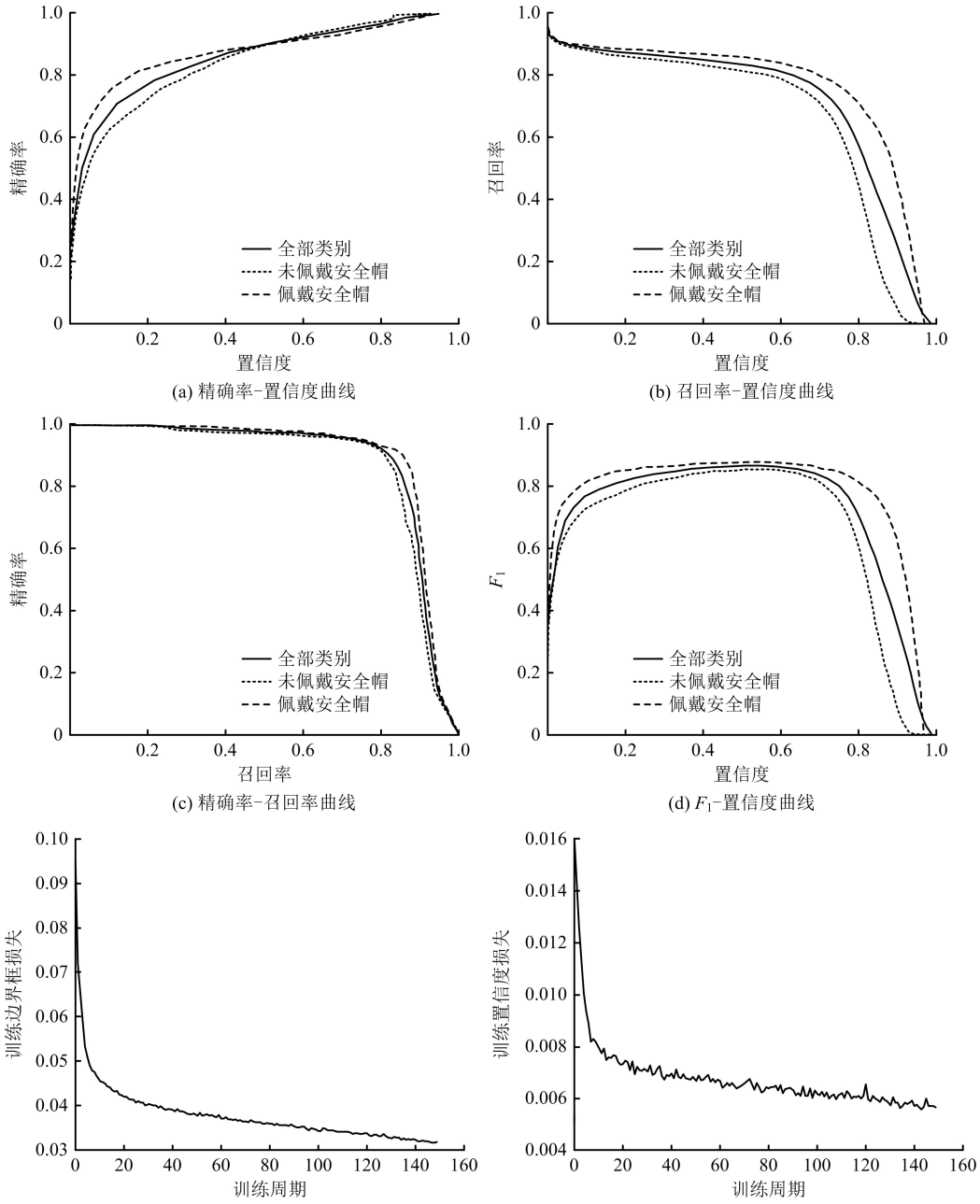
$$F_1 = 2 \times \frac{PR}{P + R} \tag{9}$$

式中 T_P表示真正例数,F_P表示假正例数,F_N表示假负例数;P 用于衡量预测结果中真正为安全帽的比例;R 用于衡量所有安全帽目标中被正确检测的比例;F_1用于平衡 P 与 R 的权重。

3 结果与讨论

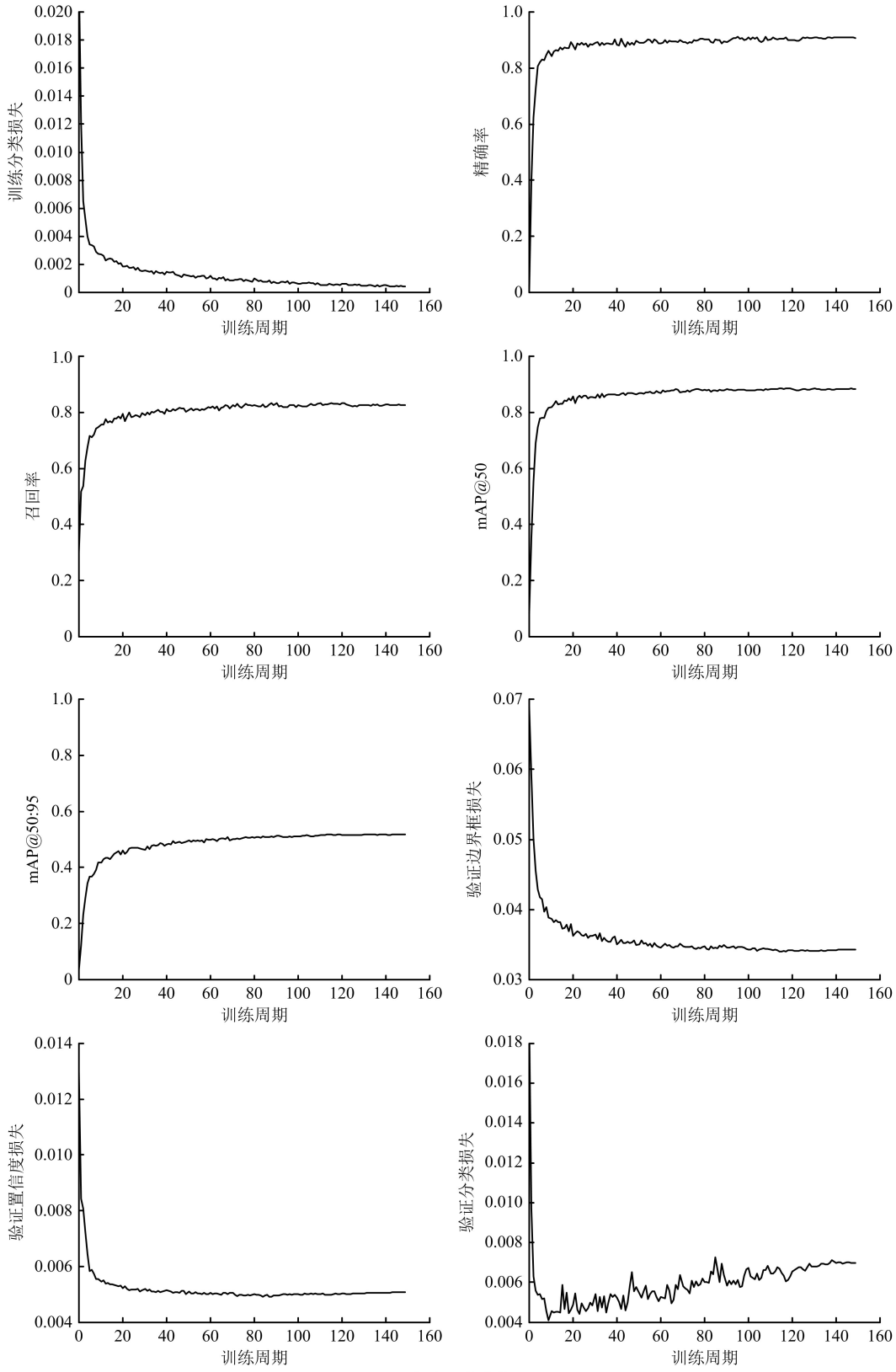
3.1 训练结果可视化

为了全面验证 HIC - YOLOv5 在井下安全帽检测任务中的有效性,采用 CUMT - Helmet 公开数据集进行训练与测试。图 6 为改进模型的可视化训练



(a) 精确率-置信度曲线 (b) 召回率-置信度曲线

(c) 精确率-召回率曲线 (d) F_1-置信度曲线



(e) 损失函数曲线

图 6　训练结果可视化

Fig. 6　Visualization of training results

图 7 HIC－YOLOv5 的安全帽检测结果

Fig. 7 Helmet detection results of HIC－YOLOv5

结果,图 7 为改进模型对安全帽的检测结果。结果显示:各项损失函数在训练过程中均稳定下降,精确率、召回率和 mAP 均稳定提升,表明 HIC－YOLOv5 模型在安全帽检测任务中的准确性较高。同时,验证集上的损失表现稳定,表明改进模型具有较好的泛化能力。

3.2 消融试验

为了验证所提出各改进模块在提升模型检测性能中的具体作用,本文设计并开展系统性的消融试验,逐步剔除或替换模型结构中的关键改进组件,分析各模块对整体检测效果的影响。消融试验在 CUMT－Helmet 数据集上进行,保持训练参数一致,并以 mAP@50、P、R 及模型参数量、推理速度为主要指标评价,消融试验结果见表 2。由表 2 可知,基础模型 YOLOv5 的 mAP@50 为 83.60%,mAP@50:95 为 47.13%,P 和 R 分别为 87.19% 与 79.71%。在此基础上逐一引入各模块后,各性能指标均有不同程度的提升。在引入 SODH 检测头后,模型在小目标识别能力方面表现显著增强,R 提升至 81.65%,mAP@50 增长至 86.43%,mAP@50:95 增至 48.94%,表明 SODH 在提升远距离人员头部目标的检测精度方面效果突出,尤其适应于井下复杂背景中小尺寸目标的高效定位。尽管参数量略增至 7.4 MB,FPS 降至 114 帧/s,但仍维持良好的实时检测性能。当加入 CBAM 注意力机制后,P 下降至 86.24%,R 下降至 75.33%,mAP@50 降至 82.65%,这是因为 CBAM 更专注于特征增强而非空间定位,单独引入 CBAM 可能造成注意力分布偏差,导致对复杂背景中的目标识别能力下降。但该方案在模型轻量化方面具有优势,参数量仅为 6.8 MB,推理速度可达 125 帧/s,适合对资源受限场景的快速部署需求。在引入 CFFI 通道融合模块后,模型的 P 和 R 分别为 85.42% 和 77.23%,mAP@50 为 83.85%,mAP@50:95 为 47.42%。相较于 CBAM 模块,CFFI 侧重于多尺度特征重构与通道语义增强,提升了整体的语义表达能力,适用于处理中等尺寸和边缘目标,其推理速度略有下降,但在模型准确性与速度之间实现了良好的权衡。总体来看,3 种改进模块均在特定维度发挥了积极作用,SODH 强化小目标识别能力,CBAM 提升关键特征响应强度,CFFI 优化通道信息表达与语义聚合。

3.3 对比试验

为了验证 HIC－YOLOv5 改进模型在井下作业人员安全帽佩戴检测任务中的综合性能优势,设计多组对比试验,并选取当前主流的目标检测算法作为对照模型(涵盖 YOLOv5、YOLOv7、YOLOv8、YOLOv9、YOLOv11、YOLOv12),从多个维度进行性能评估,试验结果见表 3 和图 8。在检测精度方面,HIC－YOLOv5 的表现最为优异,精确率达到 90.81%,召回率为 82.67%,mAP@50 高达 88.33%,相较于

表 2　改进模型消融试验结果

Table 2　Results of the ablation experiment on the improved model

模型	*P*/%	*R*/%	mAP@50/%	mAP@50:95/%	参数量/MB	FPS/(帧·s^{-1})
Baseline	87.19	79.71	83.60	47.13	7.0	120
+SODH	89.68	81.65	86.43	48.94	7.4	114
+CBAM	86.24	75.33	82.65	46.09	6.8	125
+CFFI	85.42	77.23	83.85	47.42	7.2	118
+SODH+CBAM	88.47	80.87	86.85	48.32	7.5	108
+SODH+CFFI	89.53	81.73	87.32	50.78	8.2	105
HIC－YOLOv5	90.81	82.67	88.33	51.89	8.4	102

表 3　对比试验结果

Table 3　Comparison test results

模型	*P*/%	*R*/%	mAP@50/%	参数量/MB	FLOPs/GB
YOLOv5	87.19	79.71	83.60	7.00	16.0
YOLOv7	87.75	80.63	84.67	36.49	103.2
YOLOv8	87.01	77.58	84.25	3.01	8.2
YOLOv9	84.45	77.75	83.81	9.60	38.7
YOLOv11	86.40	77.70	84.27	2.58	6.3
YOLOv12	86.10	75.81	82.80	2.52	6.0
HIC－YOLOv5	90.81	82.67	88.33	8.40	30.6

基础模型 YOLOv5 提升了 4.73 百分点,也显著优于 YOLOv7、YOLOv8、YOLOv9、YOLOv11 与 YOLOv12,表明 HIC－YOLOv5 在矿井复杂环境中具备更强的目标识别能力,减少了误检与漏检的发生。从模型复杂度来看,HIC－YOLOv5 的参数量为 8.4 MB,虽然高于 YOLOv8、YOLOv11 及 YOLOv12,但远低于 YOLOv7 和 YOLOv9,表明改进模型在兼顾检测精度提升的同时实现了结构轻量化的平衡。在计算资源消耗方面,HIC－YOLOv5 的浮点计算量为 30.6 GB,低于 YOLOv7 的 103.2 GB,但高于 YOLOv5 与 YOLOv12,表明 HIC－YOLOv5 仅以中等程度的计算资源换取了性能上的突破,具有良好的工程部署适应性。

整体而言,HIC－YOLOv5 在保持相对紧凑模型结构的基础上,实现了精度、准确率与召回率的全面提升,尤其在矿井等低照度、高遮挡、小目标密集等复杂场景中展现出更强的鲁棒性与应用价值。试验结果表明了其结构设计的有效性,这为矿山安全智能监测提供了一种性能优越、部署友好且实用性强的解决方案。

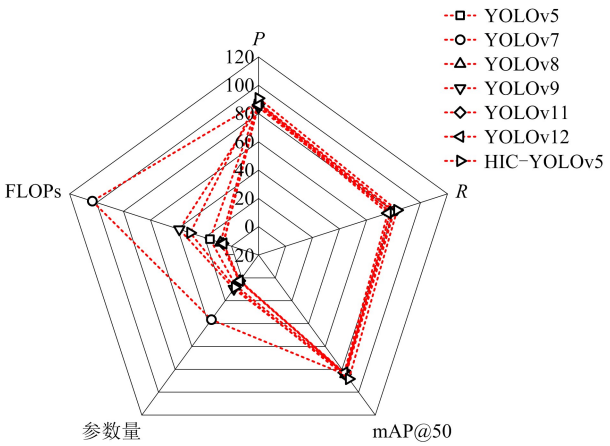


图 8　对比试验结果

Fig. 8　Comparison test results

4　结　论

(1)在 YOLOv5 基础上引入 3 种关键结构改进模块:小目标检测头(SODH)、通道特征融合模块(CFFI)以及轻量级注意力机制(CBAM),有效增强了模型对小尺度目标与复杂背景的判别能力。消融

试验结果显示:改进后的模型在保持较低参数量的同时,mAP@50 提升至 88.33%,召回率达到 82.67%,表明了各模块对检测性能的协同增益作用。同时,HIC-YOLOv5 模型整体参数量控制在 8.4 MB,计算量为 30.6 GB,保持了较高帧率,优于大部分模型在推理速度与部署灵活性方面的表现,具备较强的边缘部署适配能力,适用于矿山实际监测环境中的实时应用需求。

(2)对比试验显示,HIC-YOLOv5 在多个主流 YOLO 系列模型中综合表现最佳,在真实矿井图像数据集 CUMT-Helmet 中实现了更高的准确率与稳定性。结合模型轻量化结构与检测鲁棒性,HIC-YOLOv5 为矿山作业人员违规行为识别与智能安全监控提供了可行、高效的技术路径。

(3)尽管 HIC-YOLOv5 在矿井安全帽检测任务中展现出良好性能,但仍存在进一步优化空间,后续研究可进一步压缩模型参数与计算复杂度,提升部署效率,且该模型用于安全帽单一目标检测,后续可扩展至烟火、作业服、人员摔倒等多类矿井行为识别任务,即通过融合多源视频信息与时序特征,实现更高层次的行为理解与风险预警。

参考文献(References):

[1] HAYAT A, MORGADO-DIAS F. Deep learning-based automatic safety helmet detection system for construction safety[J]. Applied Sciences, 2022, 12(16): 8268.

[2] 王媛彬, 韦思雄, 吴华英, 等. 基于改进 YOLOv5s 的矿井下安全帽佩戴检测算法[J]. 煤炭科学技术, 2025, 53(增刊1): 1-12.
WANG Y B, WEI S X, WU H Y, et al. Detection algorithm for wearing safety helmet under mine based on improved YOLOv5s[J]. Coal Science and Technology, 2025, 53(Suppl. 1): 1-12.

[3] 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.

[4] HUANG Y G, FANG M J, PENG J. Lightweight helmet target detection algorithm combined with Effici-Bi-Level Routing Attention[J]. PLoS One, 2024, 19(5): e0303866.

[5] LIAN Y F, LI J, DONG S H, et al. HR-YOLO: a multi-branch network model for helmet detection combined with high-resolution network and YOLOv5[J]. Electronics, 2024, 13(12): 2271.

[6] JAKUBEC M, LIESKOVSKA E, BREZANI A, et al. Deep learning-based automatic helmet recognition for two-wheeled road safety[J]. Transportation Research Procedia, 2023, 74: 1171-1178.

[7] 白培瑞, 王瑞, 刘庆一, 等. DS-YOLOv5: 一种实时的安全帽佩戴检测与识别模型[J]. 工程科学学报, 2023, 45(12): 2108-2117.
BAI P R, WANG R, LIU Q Y, et al. DS-YOLOv5: a real-time detection and recognition model for helmet wearing[J]. Chinese Journal of Engineering, 2023, 45(12): 2108-2117.

[8] 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.

[9] LI J, XIE S H, ZHOU X Y, et al. Real-time detection of coal mine safety helmet based on improved YOLOv8[J]. Journal of Real-Time Image Processing, 2025, 22(1): 26.

[10] CARION N, MASSA F, SYNNAEVE G, et al. End-to-end object detection with transformers[C]// European conference on computer vision. Cham: Springer International Publishing, 2020: 213-229.

[11] TAN S, LU G, JIANG Z, et al. Improved YOLOv5 network model and application in safety helmet detection[C]//IEEE International Conference on Intelligence and Safety for Robotics (ISR), March 4-6, 2021, Nagoya, Japan. Piscataway, NJ, USA: IEEE, 2021: 330-333.

[12] TANG S Y, ZHANG S, FANG Y N. HIC-YOLOv5: improved YOLOv5 for small object detection[C]//IEEE International Conference on Robotics and Automation (ICRA), May 13-17, 2024, Yokohama, Japan. Piscataway, NJ, USA: IEEE, 2024: 6614-6619.

[13] CAO S, WANG T, LI T, et al. UAV small target detection algorithm based on an improved YOLOv5s model[J]. Journal of Visual Communication and Image Representation, 2023, 97: 103936.

[14] ZHU X K, LYU S C, WANG X, et al. TPH-YOLOv5: improved YOLOv5 based on transformer prediction head for object detection on drone-captured scenarios[C]// Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCVW), October 11-17, 2021, Montreal, Canada. Los Alamitos, California: ICCVW, 2021: 2778-2788.

[15] WANG A, PENG T, CAO H, et al. TIA-YOLOv5: an improved YOLOv5 network for real-time detection of crop and weed in the field[J]. Frontiers in Plant Science, 2022, 13: 1091655.

[16] LI D, HU J, WANG C H, et al. Involution: inverting the inherence of convolution for visual recognition[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (ICCVW), October 11-

17, 2021, Montreal, Canada. Los Alamitos, California: ICCVW, 2021: 12321－12330.

[17] LI R, WU Y P. Improved YOLOv5 wheat ear detection algorithm based on attention mechanism [J]. Electronics, 2022, 11(11): 1673.

[18] WOO S, PARK J, LEE J Y, et al. CBAM: convolutional block attention module [C]//Proceedings of the European Conference on Computer Vision (ECCV), September 8－14, 2018, Munich, Germany. Cham, Switzerland: ECCV, 2018: 3－19.

[19] ZHENG Z H, WANG P, REN D W, et al. Enhancing geometric factors in model learning and inference for object detection and instance segmentation [J]. IEEE Transactions on Cybernetics, 2021, 52(8): 8574－8586.

[20] 孙建波, 王丽杰, 麻吉辉, 等. 基于改进 YOLOv5s 算法的光伏组件故障检测[J]. 红外技术, 2023, 45(2): 202－208.
SUN J B, WANG L J, MA J H, et al. Photovoltaic module fault detection based on improved YOLOv5s algorithm [J]. Infrared Technology, 2023, 45(2): 202－208.

Research on underground safety helmet detection algorithm based on HIC－YOLO

ZHAN Xinhui[1], YAO Rui[1], QI Yun[2], XUE Kailong[3], JING Xueyan[4], QI Qingjie[5]

(1 School of Coal Engineering, Shanxi Datong University, Datong 037003, Shanxi, China;
2 College of Mining and Coal, Inner Mongolia University of Science and Technology, Baotou 014010, Inner Mongolia, China;
3 College of Safety Science and Engineering, Xi'an University of Science and Technology, Xi'an 710054, China;
4 Southwest United Graduate School, Yunnan Normal University, Kunming 650092, China;
5 Emergency Science Research Institute, CCTEG Chinese Institute of Coal Science, Beijing 100013, China)

Abstract: To enhance the accuracy and stable real-time performance of underground safety helmet detection under a unified evaluation protocol, we propose an improved model named HIC－YOLOv5. Experiments were conducted on the public CUMT－Helmet dataset released by the China University of Mining and Technology, which was divided into training and test sets at a ratio of 9:1. The model was implemented and trained using the PyTorch framework with CUDA acceleration. For data preprocessing, Mosaic augmentation was combined with instance-level Mix Up, random affine transformations, horizontal flipping, HSV jitter, Gaussian noise, and motion blur to realistically simulate low-illumination and occlusion scenarios. In terms of architectural improvements, a 160 × 160 Small Object Detection Head (SODH) was added on top of the original three-scale prediction heads of YOLOv5 to provide fine-grained coverage for small targets. A Channel Feature Fusion with Involution (CFFI) module was then inserted between the backbone and neck to alleviate semantic loss caused by 1 × 1 channel compression and enhance locally adaptive receptive fields. Finally, a lightweight Convolutional Block Attention Module (CBAM) was integrated at the end of the backbone to emphasize helmet regions through joint channel-wise and spatial-wise attention. Experimental results showed that at IoU = 0.5, HIC－YOLOv5 achieved precision, recall, mAP@50, and mAP@50:95 of 90.81%, 82.67%, 88.33%, and 51.89%, respectively, with mAP@50 improving by 4.73 percentage points over the baseline. The model comprises 8.4 MB of parameters and 30.6 GB of FLOPs, achieving an inference speed of 102 frame/s, thus establishing a controllable trade-off between accuracy and efficiency. Ablation studies indicated that SODH significantly contributed to the recall of small objects, while CBAM incurred minimal additional cost but yielded limited accuracy gains. In contrast, CFFI provided robust improvements in channel-wise semantic reconstruction; the combination of all three modules delivered the best overall performance. Compared with YOLOv7, YOLOv8, YOLOv9, YOLOv11, YOLOv12, HIC－YOLOv5 maintained leading comprehensive accuracy and robustness with moderate inference speed. HIC－YOLOv5 achieves a stable balance between accuracy and real-time performance under low illumination, occlusion, and densely distributed small-object scenarios, making it suitable for deployment on edge devices. Future work will focus on further compressing the model via structural pruning and knowledge distillation, as well as extending it to multi-task collaborative detection without altering the current evaluation protocol.

Key words: safety engineering; mine safety; safety helmet detection; YOLOv5; small object detection head; channel feature fusion; attention mechanism