基于多维扩张轻量化YOLOv9的矿用安全帽检测方法*

钟彬，余波，白水全，杨亚伟，张帅，郝建龙

（国电建投内蒙古能源有限公司， 内蒙古 鄂尔多斯市 017209）

摘要：针对矿井复杂环境下人员安全帽佩戴情况检测精度低、漏检和误检等问题，提出一种基于YOLOv9模型的多维扩张轻量化模型MD-YOLO。为解决图像信息清晰度低、受粉尘遮挡等问题，通过引入多维扩张模块改进原有特征提取模块，使模型在检测安全帽时获取更大的感受野，以增强捕捉全局特征的能力；为应对小目标和密度目标漏检情况，构建了融合注意力机制的动态检测头（PD-Dyhead），从尺度、空间、任务感知3个方面增强了模型的特征提取能力；通过采用指数移动平均样本加权损失函数（EMA-SlideLoss）改善了YOLOv9模型在处理样本不平衡问题上的不足。试验结果表明，相较于YOLOv9模型，MD-YOLO模型的检测精度达到92.5%，提高了2.8个百分点，参数量减少了26.6%，检测速度提高了27.0%。MD-YOLO模型在各种复杂环境下均能有效监测安全帽佩戴情况，综合性能最优。

关键词：安全帽检测；YOLOv9；多维扩张；轻量化；消融试验

中图分类号：TD76 **文献标识码**：A **文章编号**：1005-2763(2025)08-0236-10

Mining Safety Helmet Detection Method Based on Multi-Dimensional Expansion and Lightweight YOLOv9

ZHONG Bin, *YU Bo*, *BAI Shuiquan*, *YANG Yawei*, *ZHANG shuai*, *HAO Jianlong*

(Guodian Construction Investment Inner Mongolia Energy Co., Ltd., Ordos, Inner Mongolia 017209, China)

Abstract: A multi-dimensional expansion lightweight model MD-YOLO based on YOLOv9 model was proposed to address the issues of low detection accuracy, missed detection and false detection of personnel wearing safety helmets in complex mining environments. To solve the problems of low image clarity and dust obstruction, the original feature extraction module was improved by introducing a multi-dimensional expansion module, which enabled the model to obtain a larger receptive field when detecting safety helmets and enhanced the ability to capture global features. To address the missed detection of small and density targets, a dynamic detection head (PD Dyhead) with attention mechanism was constructed, which enhanced the feature extraction ability of the model from three aspects of scale, space and task perception. The use of exponential moving average sample weighted loss function (EMA SledeLoss) improved the shortcomings of YOLOv9 model in dealing with sample imbalance problems. The experimental results show that compared with the YOLOv9 model, the detection accuracy of the MD-YOLO model reaches 92.5%, with an increase of 2.8 percentage points. The parameter quantity decreases by 26.6%, and has a 27.0% increase in detection speed. The MD-YOLO model can effectively monitor the wearing of safety helmets in various complex environments, with the best comprehensive performance.

Key words: Safety helmet detection, YOLOv9, Multi-dimensional expansion, Lightweight, Ablation experiment

0 引言

在矿井等高风险昏暗环境中，正确佩戴安全帽是保护工作人员生命安全的重要措施[1]。安全帽不仅能有效预防高处坠物对头部造成的直接伤害，还能在意外坠落时显著减轻头部受到的冲击力。然而在实际工作中，安全帽的佩戴情况受多种因素的影响，如工作人员因疏忽或不规范佩戴安全帽而削弱其保护作用。由于井下作业环境复杂，仅依靠人工检测安全帽佩戴情况难度较大。因此，采用智能化设备对安全帽佩戴情况进行精准检测，对于降低事故发生率和保障工作人员安全具有重要意义。

安全帽检测任务属于目标检测领域，在深度学习算法兴起之前，该任务主要通过机器学习视觉框

* **收稿日期**：2024-08-25

基金项目：国家自然科学基金项目(52004120)

作者简介：钟彬(1987—)，男，陕西定边县人，工程师，主要研究方向为安全工程。E-mail:64787168@qq.com

架完成。刘晓慧[2]提出一种基于肤色定位和 Hu 矩的安全帽检测方法,通过肤色检测定位目标人脸,然后使用 Hu 矩作为特征向量,比较支持向量机和神经网络的检测效果。虽然该方法能够实现安全帽识别,但未考虑矿井环境中工作人员脸部可能存在污渍的问题,其检测稳定性较差。SHRESTHA 等[3]针对建筑安全领域设计了一种可视化检测系统,该系统首先通过闭路电视获取目标图像,然后采用 Haar-like 方法提取面部特征,最后通过边缘检测方法识别安全帽轮廓信息。然而,这种分阶段处理方式导致检测任务的实时性较差。李晓宇等[4]提出一种基于超像素特征提取与支持向量机分类的矿井人员安全帽分割方法,该方法采用简单线性迭代聚类模型将人员图像粒化为若干内部像素颜色特征相似且空间位置相近的超像素。然而,当安全帽像素与背景像素被划分为同一超像素区域时,易出现漏分割问题。传统的机器学习算法虽然能够实现安全帽检测任务,但需要先定位目标,再进行检测,这种处理方式影响了系统的实时性。此外,传统方法多通过建模的方式进行目标定位,未能充分考虑矿井环境中的实际影响因素,因此在检测稳定性方面表现不佳。

随着人工智能技术的发展,基于深度学习算法的安全帽检测技术取得了新的进展[5]。提出了以 R-CNN(Regions-Convolutional Neural Networks)、Faster R-CNN 等为代表的两阶段检测算法以及以 YOLO 系列[6]为代表的单阶段检测算法,相较于传统方法,该算法在检测精度和速度方面均有显著提升,为矿山环境中的安全帽检测提供了更加有效的解决方案。在两阶段检测算法中,徐守坤等[7]针对已有算法检测部分遮挡、尺寸不一和小目标时难度大、准确率低的问题,提出了一种基于 Faster R-CNN 的改进安全帽佩戴检测算法,该算法通过多尺度训练和增加锚点数量,增强了网络在检测不同尺寸目标时的鲁棒性,并采用多部件结合的方法剔除误检目标。尽管该方法在准确度上较传统算法有所提高,但由于仍采用两阶段检测方法,实时性较差。韩忠利[8]提出了一种基于深度卷积网络的矿用安全帽检测方法,并结合 VGG-16(Visual Geometry Group-16)框架进行检测。该方法在检测精度上有所提升,但由于参数量较大,检测速度表现不佳。在 YOLO 系列算法中,李忠飞等[9]提出了一种融合坐标注意力与多尺度的轻量级模型,通过对 YOLOv4 特征融合网络进行重构,引入注意力模块以提高其安全帽佩戴检测精度。但该模型受限于训练数据集规模大和最终检测精度不高的问题。孙光灵等[10]提出一种基于 KD-YOLO(Knowledge Distillation-YOLO)的安全帽检测算法,将注意力机制引入 YOLOv5s,并使用引导训练的方法提高其准确度,但该模型在矿井下的检测效果较差。相比于传统方法,基于深度学习的检测算法在矿井下安全帽识别中表现出更优的准确性和环境适应性。其中,YOLO 系列框架在该领域展现出突出的技术优势。这些方法能够适应相对良好的环境条件,并提供较好的检测结果。然而,目前大多数矿用安全帽检测算法忽略了矿井下的环境因素,导致其在低光照、粉尘弥漫等能见度低的场景下,实时性较差且检测精度不高。此外,轻量化算法模型难以适应现阶段高帧率视频流的实时检测任务,这在一定程度上影响了检测效率和准确性。

针对上述问题,本文提出了一种基于多维扩张轻量化 YOLOv9 的矿用安全帽检测模型 MD-YOLO(Multi-dimensional expansion lightweight model-YOLO)。本研究的创新性工作主要体现在以下 3 个方面。

(1) 针对矿井工作环境中存在的低光照和粉尘弥漫等问题,在 YOLOv9 模型的基础上引入 MEMCSPELAN4 模块。该模块使用多维扩张模块(Multi-dimensional Expansion Module,MEM)对特征提取模块 RepNCSPELAN4 进行改进,使模型在检测安全帽时获取更大的感受野,增强捕获全局特征的能力。通过获得更多的全局上下文信息,模型能够在复杂矿井环境下更精准地定位安全帽目标,提高检测的准确性和鲁棒性。

(2) 构建了融合注意力机制的动态检测头(PD-Dyhead),有效增强了小目标的检测能力,并减少遮挡情况下的漏检问题,提高模型的感知能力和检测精度,使模型在复杂环境下更具稳定性。

(3) 采用指数移动平均样本加权损失函数(EMA-SlideLoss)替代 YOLOv9 网络的交叉熵损失函数。EMA-SlideLoss 旨在解决样本不平衡问题,通过为困难样本分配更高的权重,增强了模型在矿井环境中处理遮挡情况时的分类能力,提高了训练的稳定性和鲁棒性。

1 YOLOv9 模型

YOLOv9[11]模型结构如图 1 所示。YOLOv9

模型由 Backbone(骨干)、Neck(颈部)和 Head(头部)部分组成。相比于之前的 YOLO 模型,YOLOv9 在训练过程中引入了一种可编程梯度信息,能够进一步提升网络的稳定性和准确度。

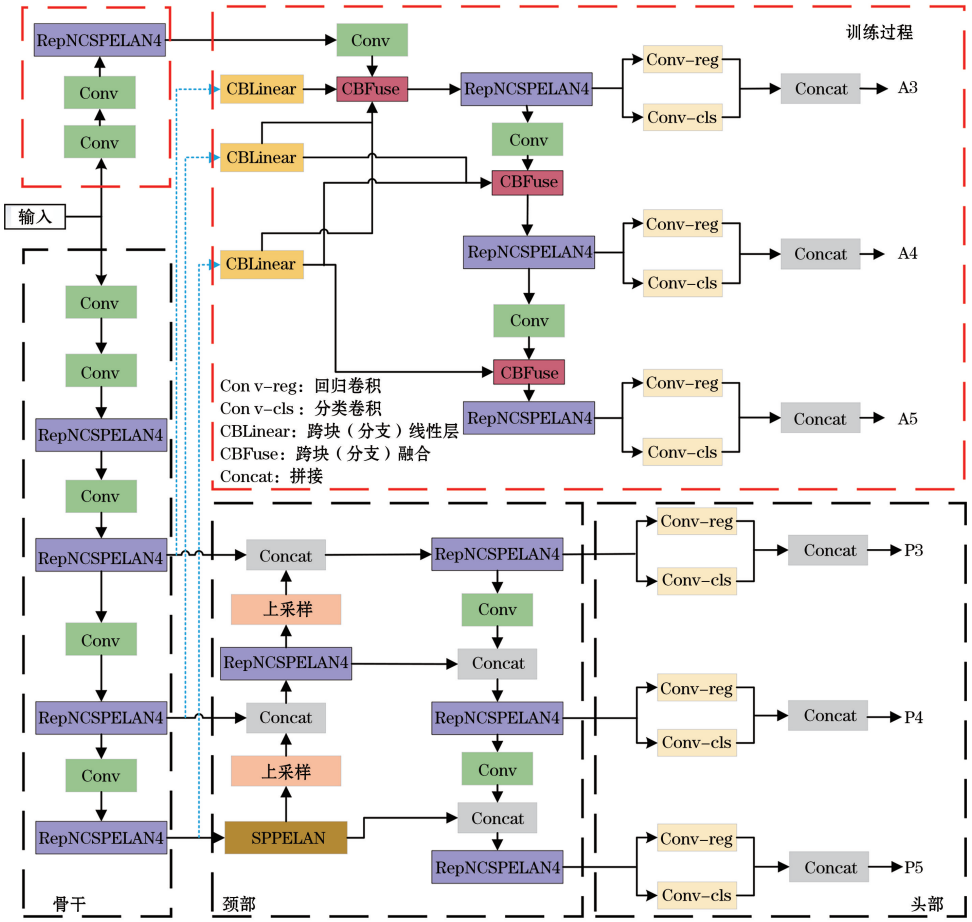


图 1 YOLOv9 模型结构

Fig.1 Structure of YOLOv9 model

YOLOv9 模型的 Backbone 骨干网络源于 YOLOv5[12]的 CSPDarknet53,并在此基础上进行了改进。主干网络主要包括 Conv 模块(基础卷积模块)、RepNCSPELAN4 模块和 ELAN(高效层聚合网络)增强空间金字塔池化模块(Spatial Pyramid Pooling Enhanced with ELAN, SPPELAN)。其中,RepNCSPELAN4 模块结合了 RepN(Re-parameter Network,重参数网络)、CSPNet(Cross Stage Partial Network,跨阶段局部网络)和 YOLOv7 的 ELAN 模块,有助于提升网络的特征提取能力和检测精度,使 YOLOv9 在处理复杂的检测任务时表现更加优异。Neck 结构与 YOLOv8[13]模型相同,旨在进行特征的融合和增强,同时提升模型的语义理解和定位能力。Head 部分采用了解耦的输出形式,每个分支通过两个卷积模块分别计算分类损失和回归损失。

2 本文 MD-YOLO 模型

2.1 模型结构

针对在矿井等环境中安全帽检测任务的现实需求,以 YOLOv9 算法为基础,提出了一种基于多维扩张轻量化的 MD-YOLO 模型,模型结构如图 2 所示。对网络中的 RepNCSPELAN4 模块进行改进,引入多维扩张模块组成了轻量化 MEMCSPELAN4 模块。为了更好地整合图像中不同尺度变化带来的特征尺度多样性,捕捉不同尺度和形状之间的内在空间关系,采用改进后的动态检测头(PD-Dyhead)替换掉原有的解耦头,使得模型能够更有效地聚焦于目标特征。在预测网络部分,采用 EMA-SlideLoss 作为类别分类损失函数中的一部分,以进一步提高安全帽佩戴检测的精度。

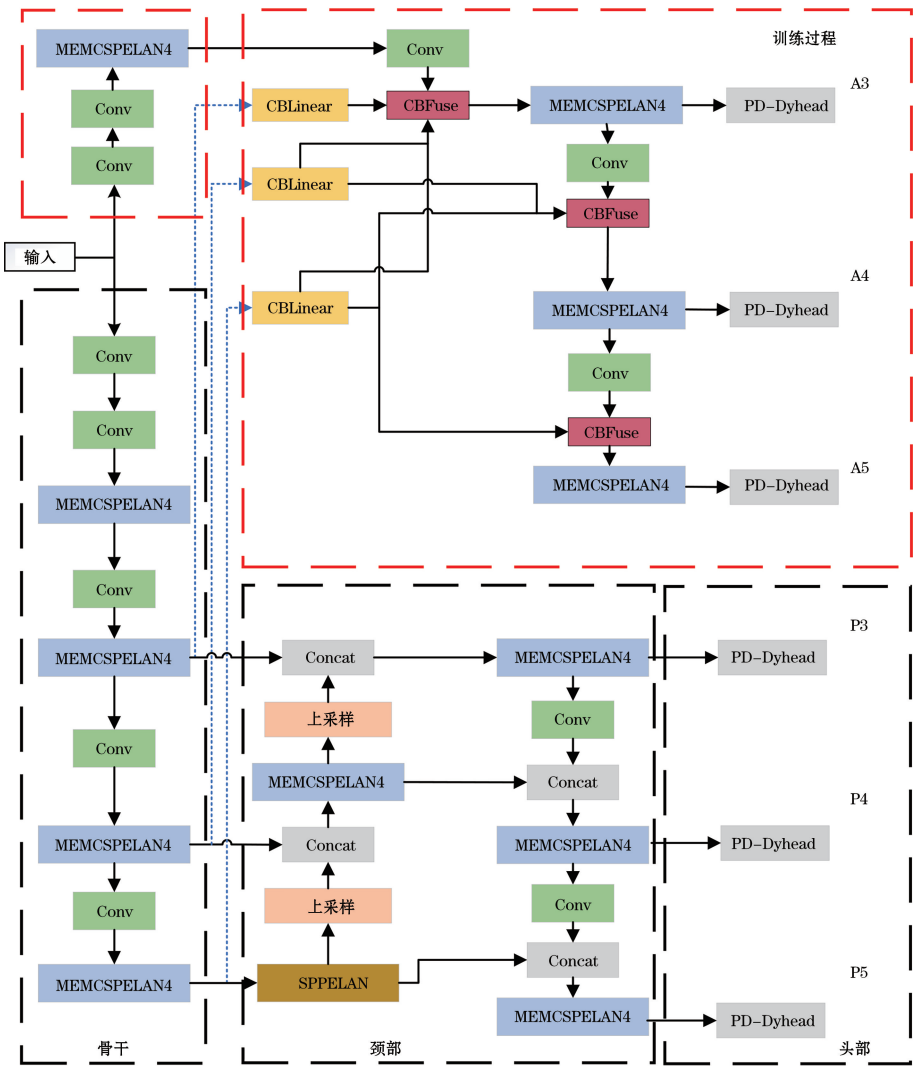


图 2　MD-YOLO 模型结构

Fig.2　Structure of MD-YOLO model

2.2　MEMCSPELAN4 模块

复杂的矿井作业环境中能见度较低,加之粉尘弥漫等恶劣条件,收集到的佩戴安全帽图像往往受到不同程度的遮挡,存在图像不清晰等问题。这些因素导致原始 YOLOv9 模型在准确捕捉并清晰提取安全帽特征方面表现不佳,进而影响检测效率和准确率,难以满足高标准的安全监管需求。

针对这一问题,本文提出了轻量化 MEMCSPELAN4 模块,取代了特征提取模块 RepNCSPELAN4,提高模型性能的同时降低了推理成本。MEMCSPELAN4 模块结构如图 3 所示。该模块使用多维扩张模块替代 RepNBottleneck 中的 RepConvN 模块,有助于模型在处理图像时捕获更大的感受野,从而增强了模型捕获全局特征的能力。同时使得模型能够获得更多的全局上下文信息,有助于在复杂环境中更精确地定位目标,提高了检测精度。

传统的卷积网络在处理小目标和密集目标时表现较差。在相同深度的卷积网络中,使用较大尺度的卷积单元相比于小卷积单元具有更广阔的感受野。然而,大的卷积单元在训练和测试中需要的参数量且计算成本较高,导致使用多尺度卷积的网络实时性较差。

为了解决这一问题,本文提出一种基于扩张卷积[14]的多维扩张模块,兼顾了感受野大和参数量少的需求,具体结构如图 4 所示。该模块融合了一个非扩张的小尺度卷积和不同尺度的扩张卷积,以增强网络对不同大小目标的特征提取能力。其中,扩张率(Dilation)采用动态调整机制,可根据尺度大小灵活变化,初始值设置为(1,1,2,3)。此外,为了提

高其训练稳定性和保持数据的连贯性,每个扩张参数首先经 BN 层(Batch Normalization)集成到前置卷积单元中,再使用元素相加操作将所扩张的维度相加,在增大覆盖面积的同时提高模型在深层次空间特征的捕捉能力。大量试验结果表明,相较于多尺度卷积模块,MEMCSPELAN4 模块参数量明显下降,能够在消耗较少计算成本的基础上提高模型的感知力。

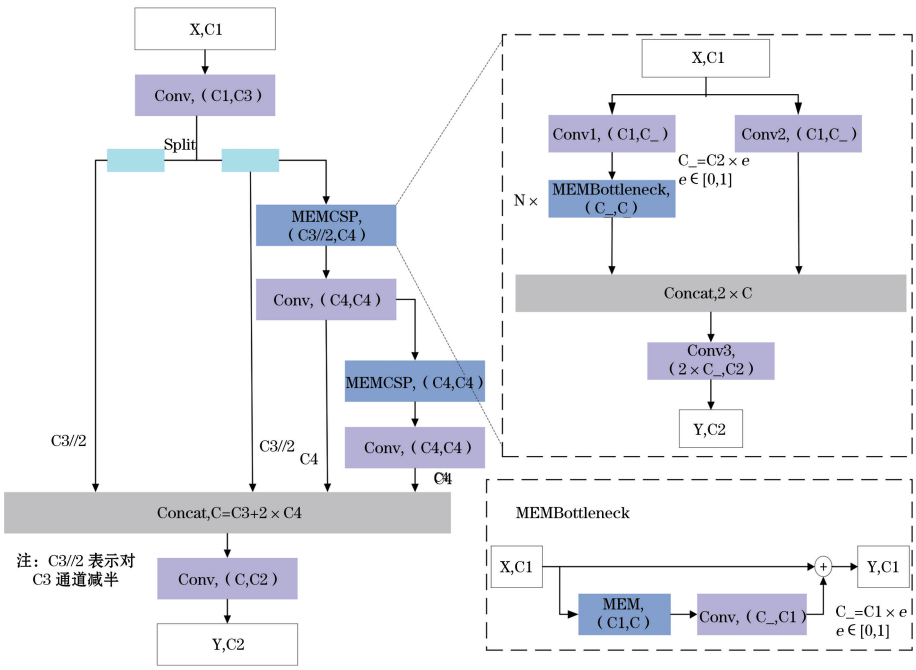


图 3 MEMCSPELAN4 模块结构

Fig.3 Structure of MEMCSPELAN4 module

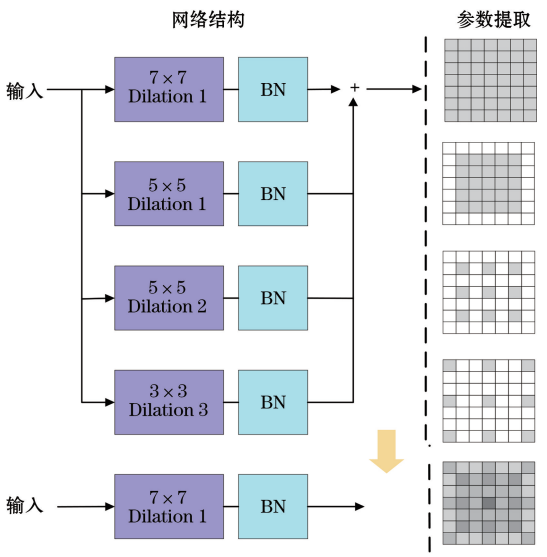


图 4 多维扩张模块结构

Fig.4 Structure of multi-dimensional expansion module

2.3 融合注意力机制的动态检测头

真实矿井环境中,监控设备与工人实地作业位置较远,检测目标多为小目标,且作业时,工作人员较为密集,容易出现遮挡情况。针对上述问题,YOLOv9 模型的解耦头难以达到理想的检测效果,容易出现漏检和误检的情况。

因此,本文提出一种融合注意力机制的动态检测头(PD-Dyhead)代替原有网络的解耦头。PD-Dyhead 由可形变卷积网络[15](Deformable Convolution Networks v3, DCNv3)和动态头[16](Dynamic Head, Dyhead)组成。DCNv3 能够弥补传统卷积在长距离建模能力和自适应空间聚合能力上的不足,通过引入可学习的偏移量,使卷积核能够在空间上自适应地变形,从而更好地适应目标的形状变化。PD-Dyhead 采用 Parametric ReLU(PReLU)[17]激活函数,使不同层具有不同的负斜率,增强小目标的检测能力,并减少遮挡情况下的漏检问题,从而更准确地识别安全帽目标。PD-Dyhead 结构如图 5 所示。

由图 5 可知,PD-Dyhead 将尺度感知注意力模块、空间感知注意力模块和任务感知注意力模块相结合,三者分别关注不同尺度的特征、空间位置信息和通道信息。首先采用尺度感知获取目标的大小信息,然后采用通道感知对比不同通道下目标突出性,

最后采用空间感知深度提取特征信息,提高准确度。其中,Offset 表示偏移量;(1,0,0,0)为超参数,用于控制阈值。这种集成显著增强了 PD-Dyhead 的特征表示能力,从而获得更准确的检测结果。计算见式(1)。

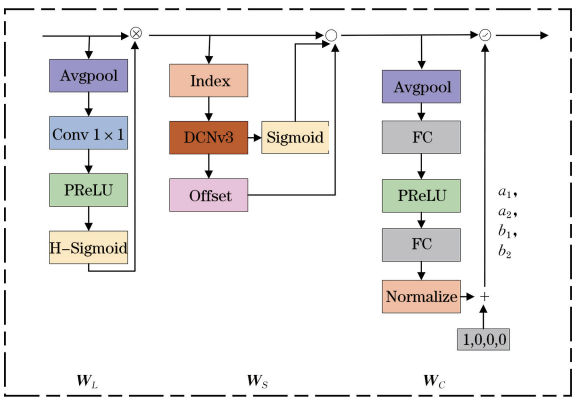


图 5 PD-Dyhead 结构

Fig.5 PD-Dyhead structure

$$A(\boldsymbol{F})=W_C(W_S(W_L(\boldsymbol{F})\cdot\boldsymbol{F})\cdot\boldsymbol{F})\cdot\boldsymbol{F} \quad (1)$$

式中:$\boldsymbol{F}$ 为三维张量输入特征,$\boldsymbol{F}\in\boldsymbol{R}^{L\times S\times C}$;$L$、$S$、$C$ 分别为特征图的层次结构、特征图的宽高积、特征图的通道数;W_L、W_S、W_C 分别为尺度感知注意力模块、空间感知注意力模块、任务感知注意力模块。

在尺度感知注意力模块中,首先通过平均池化对输入特征进行噪声平滑,随后采用 1×1 大小的卷积单元配合 PReLU 激活函数提取不同尺度下的边缘特征,最后引入 Hard-Sigmoid 函数增强尺度权重分布,提高计算效率,并通过元素相乘将权重特征融合至主干网络。计算见式(2)。

$$W_L(\boldsymbol{F})\cdot\boldsymbol{F}=\sigma\left(f\left(\frac{1}{S\times C}\sum_{S,C}\boldsymbol{F}\right)\right)\cdot\boldsymbol{F} \quad (2)$$

式中:f 为卷积核为 1×1 的卷积;σ 为 H-Sigmoid 函数,见式(3)。

$$\sigma(x)=\max\left(0,\min\left(1,\frac{x+1}{2}\right)\right) \quad (3)$$

式中,max、min 表示对输入内容取最大值和最小值的函数。

空间感知注意力模块通过多层级特征融合机制增强空间表征能力。该模块首先利用 DCNv3 实现特征图的稀疏采样,随后在相同的空间位置聚合各层级的特征,最后在该位置添加偏置使模型获得该特征的形变表征能力。具体见式(4)。

$$W_S(\boldsymbol{F})\cdot\boldsymbol{F}=\frac{1}{L}\sum_{l=1}^{L}\sum_{g=1}^{G}\sum_{k=1}^{K}w_{l,g}\cdot\boldsymbol{F}(l;p_k+p_{gk};c)\cdot m_{gk} \quad (4)$$

式中:G 为卷积组的数量;K 为样本点的个数;$w_{l,g}$ 为每组的共享投影权值;m_{gk} 为第 g 组中第 k 个采样点归一化后因子调制的标量;p_{gk} 为第 g 组中网格采样点位置 p_k 对应的偏移量;p_k 为基准采样点位置。

任务感知注意力模块通过动态特征选择机制实现多任务适配。首先采用平均池化降低特征维数,随后通过两个全连接层配合归一化层进行特征变换,最终采用移位的 Sigmoid 函数输出归一化为[-1,1],实现特征通道的动态控制。具体见式(5)。

$$W_C(\boldsymbol{F})\cdot\boldsymbol{F}=\max(a_1(\boldsymbol{F})\cdot\boldsymbol{F}_C+b_1(\boldsymbol{F}),a_2(\boldsymbol{F})\cdot\boldsymbol{F}_C+b_2(\boldsymbol{F})) \quad (5)$$

式中:$\boldsymbol{F}_C$ 为通道 C 的特征切片;θ 为控制阈值的超参数,$\theta=[a_1,a_2,b_1,b_2]$。

2.4 改进损失函数

针对复杂矿井环境下的安全帽佩戴检测任务,在选择损失函数时主要面临两个问题:一是环境中的粉尘、雾气等干扰因素降低了监控视频的清晰度,加之作业人员密集造成目标重叠和遮挡现象,导致样本特征模糊且分布失衡,影响检测准确性,增加检测难度;二是充分考虑目标大小以及后期检测的尺度、空间等因素,使得原 YOLOv9 的损失函数难以最大化提升网络的检测效率。

YOLOv9 的类别预测损失沿用了 YOLOv8 的二元交叉熵损失(Binary Cross Entropy Loss, BCE Loss),但是 BCE Loss 对样本难易失衡问题较为敏感,当训练集中存在样本失衡问题时,会降低试验结果的准确性。为了应对这一问题,研究引入 SlideLoss[18] 解决易样本和难样本之间的不平衡问题,该函数根据预测框和真实框之间的交并比(Intersection of Union, IoU)大小来区分样本难度。由于 SlideLoss 在调整平均值 μ 时无法实现真正平滑的“滑动”,在此基础上引入了指数移动平均[19](Exponential Moving Average, EMA)技术,设计了新的损失函数 EMA-SlideLoss。该损失函数通过为困难样本分配更高的权重,促使模型学习更多不同的特征,从而提高模型的检测性能。

首先根据预测框和真实框的 IoU 值来调整对正负样本的权重,见式(6)。

$$f(x)=\begin{cases}1, & x\leqslant\mu-0.1\\ e^{(1-\mu)}, & \mu-0.1<x<\mu\\ e^{(1-x)}, & x\geqslant\mu\end{cases} \quad (6)$$

式中,μ 为锚框的平均 IoU 值。

由于模型对困难样本的识别能力较差,

SlideLoss 会放大困难样本的相对损失,同时强调被误分类的样本,使模型能够有效利用简单样本的有限特征。

针对现实场景中视频目标的移动模糊问题,引入 EMA 方法动态调整均值 μ_{auto},实现对损失的自适应调整。这种方法增强了模型对运动模糊和形变的适应能力,提升了整体的训练效果。具体见式(7)。

$$V_{EMA}(t)=d\cdot(1-e^{-\frac{t}{t_{all}}})V_{EMA}(t-1)+(1-d)\cdot\mu_{auto} \tag{7}$$

式中:d 为衰减系数,取 0.999;t、t_{all}分别为迭代次数和总迭代次数。

EMA-SlideLoss 提供了一种平滑的机制来平衡历史数据和当前观测数据对损失的影响。最终损失函数的计算见式(8)。

$$f(x)=\begin{cases}1, & x\leqslant V_{EMA}(t)-0.1\\ e^{(1-V_{EMA}(t))}, & V_{EMA}(t)-0.1<x<V_{EMA}(t)\\ e^{(1-x)}, & x\geqslant V_{EMA}(t)\end{cases} \tag{8}$$

3 试验结果与分析

3.1 数据集与评价指标

为了应对井下复杂多变的工作环境,试验采用开源的安全帽佩戴数据集(Safety Helmet Wearing Dataset, SHWD)[20]进行模型的训练与测试。原始数据集包含了 5 种标签类别,但本文仅关注矿井下工人安全帽佩戴状态,因此,对数据集进行筛选,将标签简化为两类检测对象,分别为佩戴安全帽的人和未佩戴安全帽的人。数据集包含 2 000 张图片,按照 8∶1∶1 的比例划分训练集、验证集和测试集。

为验证 MD-YOLO 模型的性能并与其他目标检测模型进行比较,试验采用关键性能指标评价模型表现:采用平均精度(Average Precision, AP)及平均精度均值(mean Average Precision,mAP)衡量模型的检测性能;采用准确度(Precision, P)和召回率(Recall, R) 综合评估检测结果的准确度[13];采用每秒帧数(Frames Per Second, FPS)衡量模型的检测速度。具体计算公式如下。

$$M=\frac{\sum_{i=1}^{N}A_i}{N} \tag{9}$$

$$V_{FPS}=\frac{1}{T} \tag{10}$$

式中:M 为平均精度均值 mAP;N 为类别数;A_i 为第 i 个类别下的平均精度 AP;V_{FPS} 为 FPS 值;T 为每帧处理时间,s。

3.2 试验环境与参数配置

试验环境配置见表 1。在模型训练阶段,设置学习率为 0.000 1,迭代次数为 150 次,训练的批次大小为 8。SHWD 数据集训练损失变化曲线如图 6 所示。由图 6 可知,本文提出的 MD-YOLO 模型能够快速收敛并保持稳定的训练过程。

表 1 试验环境配置

Table 1 Experimental environment setup

环境	类别	规格
硬件环境	GPU	NVIDIA GeForce RTX 4090 24GB
	CPU	Intel(R) Core(TM) i7-11600H
	RAM	32 GB
软件环境	Windows	11
	Python	3.7
	CUDA	11.3.1
	PyTorch	1.7.1

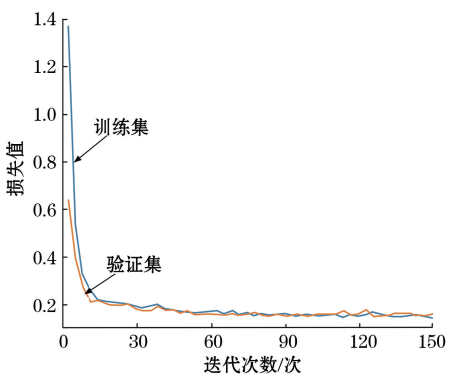


图 6 损失变化曲线

Fig.6 Loss variation curves

3.3 不同目标检测模型对比

基于标准数据集对 MD-YOLO 与主流目标检测模型进行对比试验,试验结果见表 2。从表 2 可以看出,相较于 Faster RCNN 模型[7]、SSD 模型[21]以及其他 YOLO 系列模型,MD-YOLO 模型在检测精度和实时性上均展现出了明显的优势。

表 2 不同目标检测模型性能对比

Table 2 Comparison of different object detection models property

模型	P/%	R/%	mAP/%	参数量/($\times10^6$)	FPS
Faster RCNN	76.4	79.6	80.2	136.7	34
SSD	78.7	75.3	76.2	23.7	47
YOLOv7	84.6	78.9	85.1	37.2	63
YOLOv8	89.2	83.2	88.4	27.8	69
YOLOv9	89.8	90.1	89.7	31.9	74
MD-YOLO	91.7	90.4	92.5	23.4	94

与 YOLOv9 算法相比,MD-YOLO 在准确度、召回率和检测精度上分别提升了 0.9 个百分点、0.3 个百分点和 2.8 个百分点。MD-YOLO 模型的检测精度达到 92.5%,检测速度提高了 27.0%,参数量减少了 26.6%,证明了其在复杂环境下的实用性和检测效率。

3.4 可视化试验

为直观验证 MD-YOLO 模型在实际复杂场景中的有效性,在真实矿井环境下对比了 YOLOv7、YOLOv8、YOLOv9、MD-YOLO 模型的安全帽佩戴检测性能,试验结果如图 7 所示。

由图 7(a)可知,在低照度矿井环境中,MD-YOLO 模型展现了其在处理低光环境问题上的优越性,检测效果显著优于未改进的 YOLO 模型基准模型;由图 7(b)可知,面对目标密集且相互遮挡的复杂场景,除了 YOLOv7 模型外,各模型均正确识别出了目标,MD-YOLO 模型在检测性能上更佳,置信度相较于其他模型最高;由图 7(c)可知,在小目标检测场景中,MD-YOLO 模型虽存在误检,但其整体检测效果依然显著优于其他模型;由图 7(d)可知,在探照灯照明场景中(4 个图像目标),YOLOv7、YOLOv8 模型仅检测到 3 个目标,漏检了 1 个目标,YOLOv9、MD-YOLO 模型未出现漏检情况。试验结果表明,MD-YOLO 在各种复杂环境下均能准确检测安全帽佩戴情况,综合性能最优。

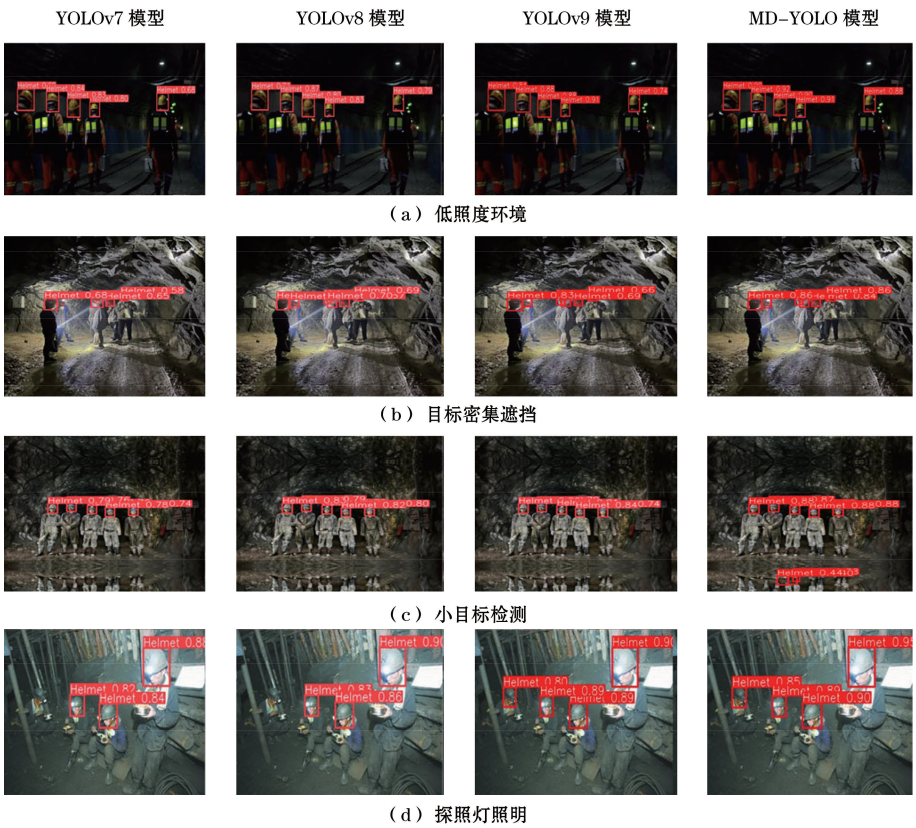


图 7 模型检测结果对比

Fig.7 Comparison of model detection results

3.5 消融试验

复杂的井下环境是影响模型检测结果的关键所在,低光照、粉尘弥漫等问题增加了检测难度。为此,通过在 YOLOv9 模型中引入 MEMCSPELAN4 模块(A)、PD-Dyhead(B)、EMA-SlideLoss 函数(C),验证各增强模块的有效性。在相同数据集和参数设置下,对 MD-YOLO 模型设计了相应的消融试验,结果见表 3。可以看出,MD-YOLO 模型在处理这些复杂场景时能够显著提高检测的准确性和鲁棒性。

表 3 消融试验结果

Table 3 Results of ablation experiment

模型	A	B	C	mAP/%	参数量/($\times10^6$)	FPS
YOLOv9	—	—	—	89.7	31.9	74
模型 1	√	—	—	90.1	26.8	80
模型 2	—	√	—	91.2	24.6	85
模型 3	—	—	√	90.5	31.9	78
模型 4	√	√	—	90.8	23.4	87
模型 5	√	—	√	91.1	26.8	83
模型 6	—	√	√	91.7	24.6	90
MD-YOLO	√	√	√	92.5	23.4	94

由模型 1 试验结果可知,MEMCSPELAN4 模块将 YOLOv9 模型的检测精度从 89.7%提高至 90.1%,表明该模块可有效增强模型的感受野,有助于模型捕捉复杂和模糊的矿井图像信息。同时,该模块在提高检测性能的同时不增加额外的成本,从而实现了精度和参数量之间的良好平衡。由模型 2 试验结果可知,PD-Dyhead 通过引入特征提取和注意力机制增强了模型对被检测目标大小的适应性,相比于 YOLOv9 模型,模型 2 的检测精度提高了 1.5 个百分点,模型检测速度也有所提高。由模型 3 试验结果可知,将原始的损失函数替换为 EMA-SlideLoss,模型的检测精度有所提高。

由模型 4 至模型 6 试验结果可知,分别加入 3 个改进功能中的两个,模型检测精度和检测速度均表现出不同程度的提升,验证了每个改进功能的必要性。结合 3 个改进功能的 MD-YOLO 模型则展现出最好的检测性能。

4 结论

针对矿井下安全帽识别准确度低和稳定性较差的问题,本文提出了基于 YOLOv9 算法的多维扩张轻量化实时检测模型 MD-YOLO,得到如下主要结论。

(1) 针对矿井工作环境中存在的低光照和粉尘弥漫等问题,引入 MEMCSPELAN 模块增强模型捕获全局特征的能力;采用动态检测头(PD-Dyhead)应对复杂环境中的小目标和密集目标,从尺度、空间和任务感知 3 个方面增强了模型的特征提取能力;使用 EMA-SlideLoss 损失函数解决了 YOLOv9 在处理样本不平衡问题上的不足。

(2) 在标准数据集上对 MD-YOLO 与主流目标检测模型进行对比试验。结果表明,与 YOLOv9 模型相比,MD-YOLO 模型的检测精度达到 92.5%,提升了 2.8 个百分点,检测速度提高了 27.0%,参数量减少了 26.6%,证明了 MD-YOLO 模型在复杂环境下的实用性和效率。

(3) 可视化试验结果表明,MD-YOLO 模型在各种复杂环境下均能有效检测安全帽佩戴情况,综合性能最优,消融试验进一步验证了各增强模块的有效性。

参考文献(References):

[1] HUANG L, FU Q B, HE M L, et al. Detection algorithm of safety helmet wearing based on deep learning [J]. Concurrency and Computation: Practice and Experience, 2021,33(13):e6234.

[2] 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 Xiaohui, YE Xining. 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.

[3] SHRESTHA K, SHRESTHA P P, BAJRACHARYA D, et al. Hard - hat detection for construction safety visualization[J]. Journal of Construction Engineering,2015(1):721380.

[4] 李晓宇,陈伟,杨维,等.基于超像素特征与 SVM 分类的人员安全帽分割方法[J].煤炭学报,2021,46(6):2009-2022.
LI Xiaoyu, CHEN Wei, YANG Wei, et al. Segmentation method for personnel safety helmet based on super-pixel features and SVM classification[J]. Journal of China Coal Society,2021,46(6):2009-2022.

[5] REN H G, FAN A N, ZHAO J, et al. Lightweight safety helmet detection algorithm using improved YOLOv5[J]. Journal of Real-Time Image Processing,2024,21(4):125.

[6] JIANG P Y, ERGU D J, LIU F Y, et al. A review of YOLO algorithm developments[J]. Procedia Computer Science, 2022,199:1066-1073.

[7] 徐守坤,王雅如,顾玉宛,等.基于改进 Faster RCNN 的安全帽佩戴检测研究[J].计算机应用研究,2020,37(3):901-905.
XU Shoukun, WANG Yaru, GU Yuwan, et al. Safety helmet wearing detection study based on improved Faster RCNN[J]. Application Research of Computers, 2020, 37(3):901-905.

[8] 韩忠利.基于卷积神经网络的矿用安全帽佩戴检测[J].工矿自动化,2024,50(增刊 1):82-87.
HAN Zhongli. Mining safety helmet wearing detection based on convolutional neural network [J]. Journal of Mine Automation,2024,50(Suppl.1):82-87.

[9] 李忠飞,冯仕咏,郭骏,等.融合坐标注意力与多尺度特征的轻量级安全帽佩戴检测[J].工矿自动化,2023,49(11):151-159.
LI Zhongfei, FENG Shiyong, GUO Jun, et al. Lightweight safety helmet wearing detection fusing coordinate attention and multiscale feature[J]. Journal of Mine Automation, 2023,49(11):151-159.

[10] 孙光灵,黄磊,吴倩.基于 KD-YOLO 的轻量化安全帽佩戴检测方法[J].淮北师范大学学报(自然科学版),2024,45(2):41-48.
SUN GuangLing, HUANG Lei, WU Qian. Lightweight safety helmet wearing detection method based on KD-YOLO [J]. Journal of Huaibei Normal University (Natural Science),2024,45(2):41-48.

[11] WANG C Y, YEH I H, LIAO H Y M. YOLOv9:Learning what you want to learn using programmable gradient information [C]//Computer Viston-ECCV 2024. Cham: Springer,2024.

[12] 刘一江,樊福景,王通.基于改进 YOLOv5 的安全帽小目标检测算法[J].信息技术与信息化,2024(5):115-119.

LIU Yijiang, FAN Fujing, WANG Tong. Safety helmet small object detection algorithm based on improved YOLOv5[J]. Information Technology and Informatization,2024(5):115-119.

[13] 王贞,邱杭,吴斌,等.基于 CCG-YOLOv8 的施工场景下安全帽佩戴检测[J].武汉理工大学学报,2024,46(6):73-80.
WANG Zhen, QIU Hang, WU Bin, et al. Safety helmet wearing detection based on CCG-YOLOv8 for construction scenarios[J]. Journal of Wuhan University of Technology, 2024,46(6):73-80.

[14] TIAN Y L, ZHANG Q S, REN Z L, et al. Multi-Scale dilated convolution network based depth estimation in intelligent transportation systems[J]. IEEE Access,2019,7:185179-185188.

[15] WANG W H, DAI J F, CHEN Z, et al. InternImage: exploring large-scale vision foundation models with deformable convolutions[C]//IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). New York:IEEE,2023:14408-14419.

[16] DAI X Y, CHEN Y P, XIAO B, et al. Dynamic head: unifying object detection heads with attentions[C]//IEEE/CVF Conference on Computer Vision and Pattern Recognition. New York:IEEE,2021:7373-7382.

[17] JIANG T T, CHENG J Y. Target recognition based on CNN with Leaky ReLU and PReLU activation functions [C]//International Conference on Sensing, Diagnostics, Prognostics, and Control (SDPC). New York:IEEE,2019: 718-722.

[18] YU Z P, HUANG H B, CHEN W J, et al. Yolo-FaceV2:a scale and occlusion aware face detector[J]. Pattern Recognition,2024,155:110714.

[19] CAI Z W, RAVICHANDRAN A, MAJI S, et al. Exponential moving average normalization for self-supervised and semi-supervised learning[C]//IEEE/CVF Conference on Computer Vision and Pattern Recognition. New York:IEEE,2021:194-203.

[20] 郑楚伟,林辉.基于 Swin Transformer 的 YOLOv5 安全帽佩戴检测方法[J].计算机测量与控制,2023,31(3):15-21.
ZHENG Chuwei, LIN Hui. YOLOv5 helmet wearing detection method based on Swin Transformer[J]. Computer Measurement & Control, 2023,31(3):15-21.

[21] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot multi-box detector[C]//European Conference on Computer Vision. New York:IEEE,2016:21-37.