



DOI:10.12404/j.issn.1671-1815.2406421

引用格式:康程洋,张延军,张睿.基于改进YOLOv7的煤矿井下低质影像安全帽检测[J].科学技术与工程,2025,25(20):8595-8603.

Kang Chengyang, Zhang Yanjun, Zhang Rui. Detection of low-quality image safety helmet in coal mine based on improved YOLOv7[J]. Science Technology and Engineering, 2025, 25(20): 8595-8603.

基于改进YOLOv7的煤矿井下低质影像安全帽检测

康程洋¹, 张延军^{1*}, 张睿²

(1. 太原科技大学机械工程学院, 太原 030024; 2. 太原科技大学计算机科学与技术学院, 太原 030024)

摘要 由于井下环境复杂,低光照以及安全帽的物体较小导致的检测效果不理想。针对复杂环境下的低质影像,提出一种改进YOLOv7的煤矿井下低质影像安全帽检测。首先,针对低光照条件下的图像特征易受噪声干扰的局限性,提出多尺度MELAN模块,通过构建多尺度注意力机制来捕捉更广泛的上下文信息,进而强化特征信息的提取,并有效抑制噪声的干扰;其次,在主干网络使用软池化和全维动态卷积构建OD-SMP模块,减少了特征映射中的信息弥散,保留了更多上下文信息,增强对小目标的检测能力;最后,对于煤矿井下不同光照和距离的复杂背景和环境,导致检测样本质量出现参差不齐的情况,使用Wise-IoU作为损失函数。实验结果表明:改进模型的平均精度为94.9%,比原来的YOLOv7模型高出13.5%,具有更好的检测效果。

关键词 低光照; 多尺度; 全维动态卷积; 注意力机制; 损失函数

中图分类号 TP391.41; **文献标志码** A

Detection of Low-quality Image Safety Helmet in Coal Mine Based on Improved YOLOv7

KANG Cheng-yang¹, ZHANG Yan-jun^{1*}, ZHANG Rui²

(1. School of Mechanical Engineering, Taiyuan University of Science and Technology, Taiyuan 030024, China;
2. School of Computer Science and Technology, Taiyuan University of Science and Technology, Taiyuan 030024, China)

[Abstract] Due to the complex underground environment, low lighting conditions, and the small size of hard hats, the detection results are not ideal. To address low-quality images in complex environments, an improved YOLOv7 for hard hat detection in low-quality images from underground coal mines was proposed. Firstly, addressing the limitation that image features were susceptible to noise interference under low-light conditions, a multi-scale MELAN module was introduced. By constructing a multi-scale attention mechanism, broader contextual information was captured, thereby enhancing feature extraction and effectively suppressing noise interference. Secondly, the OD-SMP module was constructed using soft pooling and full-dimensional dynamic convolution in the backbone network, which reduced information diffusion in feature mappings, retained more contextual information, and enhanced the detection capability for small targets. Finally, to address the varying quality of detection samples caused by the complex backgrounds and environments with different lighting and distances in underground coal mines, Wise-IoU was used as the loss function. Experimental results show that the average precision of the improved model is 94.9%, which is 13.5% higher than the original YOLOv7 model, demonstrating better detection performance.

[Keywords] low illumination; multi scale; full dimensional dynamic convolution; attention mechanism; loss function

煤矿工人必须佩戴安全帽,这是采矿现场安全管理的关键,安全帽可以有效减少和分配冲击力。根据最近的矿山安全事故报告,由于缺乏安全防范意识,偶尔会发生不戴头盔造成的伤亡。因此,为了保护煤矿工人的生命安全,煤矿单位必须监督他们正确佩戴安全帽。在早期的人工监测任务中,经常使用人工监管来确定煤矿工人是否佩戴头盔,由

于监管人员的主观性和难以了解矿山情况,这容易产生安全隐患,并增加了高昂的劳动力成本。5G、互联网和大数据等技术的发展和成熟,促进了智能矿山的建设。实时监控人员信息和提供安全保障的主要方法是“视频监控”相结合^[1]。智能检测煤矿工人是否佩戴安全帽,对消除人工监测造成的缺陷具有重要的研究价值,这引起了学者们的广泛

收稿日期:2024-08-27; 修订日期:2025-04-23

基金项目:山西省重点研发计划(202102010101010)

第一作者:康程洋(2000—),男,汉族,河南商丘人,硕士研究生。研究方向:机器视觉及智能控制。E-mail:1654032526@qq.com。

*通信作者:张延军(1982—),男,满族,河北承德人,博士,教授。研究方向:机电液传动及智能控制。E-mail:hebei0735@126.com。

关注。

在当前的研究背景下,深度学习技术已被中外学者广泛应用于安全帽检测领域。徐先峰等^[2]构建 MobileNet-SSD 算法提高检测速度,解决安全帽数据集规模较小、网络难以充分拟合特征的问题。李明山等^[3]在 SSD (single shot multibox detector) 模型中引入特征融合机制,通过构建分支网络的方式实现了特征信息的有效整合。针对默认框的配置进行了优化调整,这些改进措施显著提高了算法在实际应用中对安全帽检测的准确性。李嘉信等^[4]提出一种面向小目标的多空间层次安全帽佩戴检测算法,减少了检测中漏检的现象。丁田等^[5]在主干网络中引入注意力机制并使用 CIoU 损失函数来计算损失,减少检测过程中的干扰,提高对小目标的检测能力。

尽管基于计算机视觉的安全帽检测算法已应用于建筑、工业车间等,但在煤矿井下仍存在一些特定的挑战。对于不同光照和距离的复杂背景和环境,检测算法需要具有更强的目标识别能力和智能处理能力^[6]。鉴于此,提出改进 YOLOv7 的煤矿井下低质影像安全帽检测算法。煤矿井下安全帽佩戴检测做以下改进:在特征提取网络 ELAN 模块加入设计的多尺度注意力机制 MEMA 模块,增强获得的特征信息,减弱井下低光照环境下噪声的影响^[7];在主干网络使用软池化构建 OD-SMP 模块,减少了特征映射中的信息弥散,保留更多上下文信息,增强对小目标的检测能力;对于煤矿井下不同光照和距离的复杂背景和环境,导致检测样本质量出现参差不齐的情况,使用 Wise-IoU 损失函数,平衡样本质量。

1 YOLOv7 算法模型

YOLOv7^[8]的基本架构包括:输入、主干和头部。输入图像通过 Input 进行预处理,并将其设置为 640×640 RGB 图像。主干包括 CBS (卷积-批归一化 Silu 模块)、ELAN (高级层级集合网络) 和 MP (最大池化双路径模块)。输入图像通过主干进行特征提取。然后,在头部采用多尺度特征融合模式,可以保留更多的细节,有助于提高模型对目标的表示能力。YOLOv7 有 3 个检测头,特征融合后的特征将传递到输出层,输出层将预测目标的位置和类别,并生成相应的边界框。检测头输出 3 个特征尺度: 20×20 、 40×40 和 80×80 , 3 个尺度分别以大、中、小目标对应的尺度检测目标。YOLOv7 在预测结果中应用非极大值抑制 (non-maximum suppression, NMS) 算法来去除冗余的边界框,并确保每个

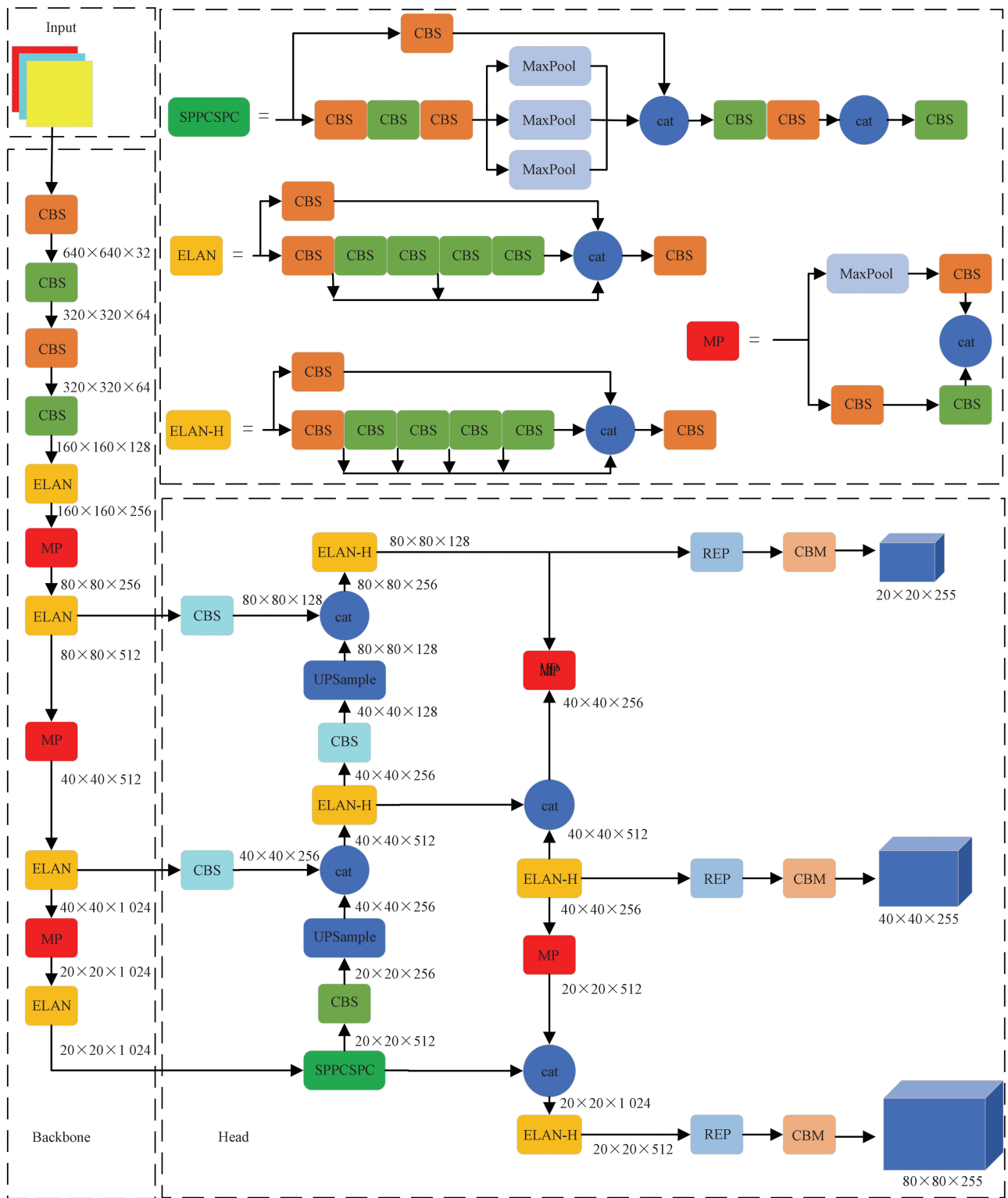
目标只有一个与其对应的边界框。YOLOv7 网络结构如图 1 所示。

2 YOLOv7 算法改进模型

2.1 多尺度 MELAN 模块

为减少低光照噪声信息对检测效果的干扰,在特征提取部分的 ELAN (efficient layer aggregation network) 模块中添加自设计 MEMA 注意力模块,通过多尺度注意力模块对全局通道间信息建模的能力,抑制不重要信息的能力,它能够使空间语义特征在每个特征组中均匀分布,对于煤矿井下低照度环境,产生更清晰的目标特征,有效地增强神经元信息,使网络更加关注数据的重要特征,增强安全帽特征,减弱低照度产生的噪声信息对检测器的影响,多尺度 MELAN 模块如图 2 所示。一种新型的注意力机制 EMA (efficient multi-scale attention) 模块,结构图如图 3 所示,该模块无需进行降维操作即可完整保留每个通道的信息,同时显著降低计算成本^[9]。能够将部分通道巧妙地重构为批量维度,并进一步将这些通道分组为多个子特征集。这一设计确保了每个特征组内部的空间语义特征分布均匀,从而提升了数据处理的效果和效率。EMA 模块不仅通过编码全局信息来动态调整每个并行分支的通道权重,还利用跨维度交互的机制,有效整合了两个并行分支的输出特征,旨在扩展网络的全局感受野、保留精确的位置信息,同时降低计算需求。然而,EMA 的平均池化计算区域内像素平均值,会丢失图像显著特征,而 MEMA 注意力机制采用全局平均池化和最大池化的双重策略代替 EMA 模块的全局平均池化,分别沿水平和垂直方向学习图像的一维特征,更好的保留安全帽的细节信息^[10]。

MEMA 注意力机制结构如图 4 所示,针对输入特征 $X \in R^{C \times H \times W}$ (R 为邻域; C 为特征图的通道数; H 为输出特征图的高度; W 为输出特征图的宽度), MEMA 首先在通道维度将其划分成 G 个子特征 $X = [X_0, X_1, \dots, X_{G-1}]$, $X_i \in R^{C/G \times H \times W}$, 取 $G \ll C$, 每组特征通过学习到的注意力权重进行增强。针对每个子特征, MEMA 架构巧妙地设计 3 条并行处理路径,旨在从不同维度提取分组特征图的注意力强度系数。其中,前两个并行路径均嵌入了 1×1 卷积层,随后分别应用了自适应全局平均池化和最大池化技术,这两种池化方法沿着 X 轴和 Y 轴方向独立地对特征图的通道信息进行压缩编码,然后这两条路径输出的编码特征被巧妙地连接起来,形成一个融合了多方向信息的特征表示,紧接着,这一综合



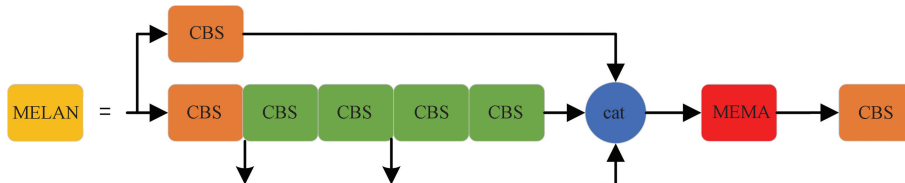
Input 为输入; Backbone 为主干; Head 为头部; CBS 为卷积-批归一化 Silu 模块; cat 为拼接操作; UPSample 为上采样量模块; ELAN 为高级层级集合网络; MP 为最大池化双路径模块; SPPCSPC 为空间金字塔池化-跨阶段部分连接; ELAN-H 为头部增强型 ELAN; CBM 为卷积-批归一化 Mish 模块

图 1 YOLOv7 网络结构

Fig. 1 YOLOv7 network structure

特征通过一个共享的 1×1 卷积层进行进一步处理, 输出两个向量后使用 Sigmoid 激活函数进行线性拟合; 第 3 个分支包含 3×3 卷积, 旨在捕获多尺度特

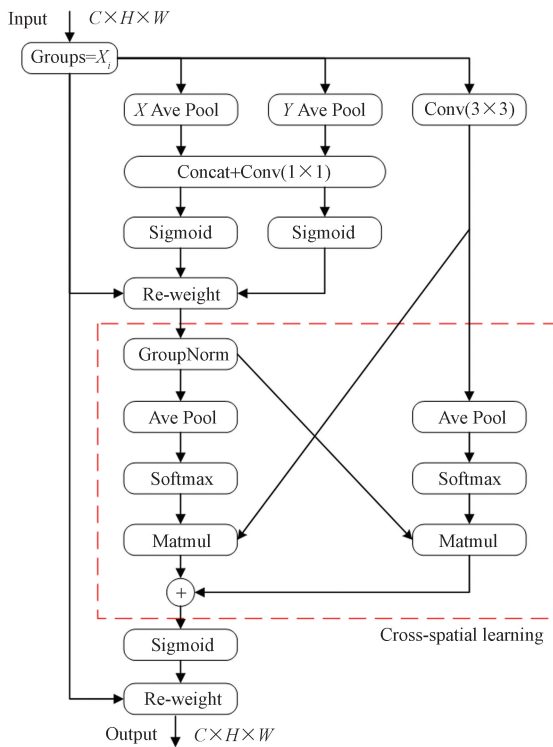
征表示。在跨空间学习 (cross-spatial learning) 部分, MEMA 对不同空间维度方向的跨空间信息进行聚合。首先, 对引入张量为 1×1 分支所输出的全局



CBS 为卷积操作;cat 为拼接操作;MEMA 为注意力机制

图2 聚合周边信息的 MELAN 模块

Fig. 2 MELAN module for aggregating peripheral information

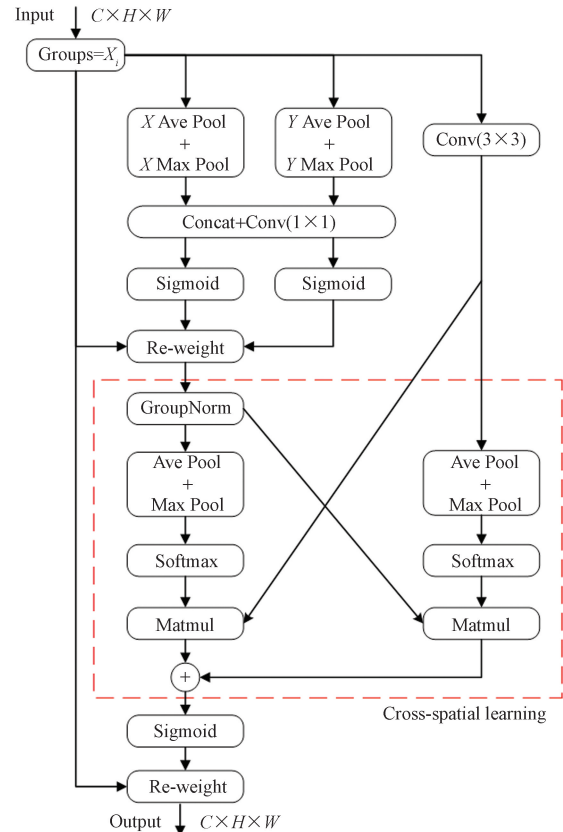


Ave Pool 为平均池化;Softmax 为归一化指数函数;Conv 为卷积;Concat 为拼接操作;Sigmoid 为激活函数;Re-weight 为分配权重;Matmul 为矩阵相乘函数;Input 为输入;Output 为输出;Groups 为分组;GroupNorm 为组归一化;Cross-spatial learning 为跨空间学习

图3 EMA 模块结构

Fig. 3 EMA module structure

信息进行归一化操作;其次,对 1×1 分支与 3×3 分支各自生成的全局空间信息进行独立编码,分别采用 2D 全局平均池化和最大池化技术,再利用 Softmax 函数进行拟合线性变换,将输出中贡献度最小的分支在信道维度上进行特征转换,调整其形状以匹配所需的维度结构,这一步骤巧妙地实现了跨空间信息的聚合,通过强化或抑制特定分支的贡献,提升了整体特征表示的鲁棒性和信息量;最后,通过跨空间交互模块聚合两个通道注意力权重值,捕捉像素级的成对关系,并通过 Sigmoid 激活函数输出特征映射,从而增强原始输入特征,得到最终的输出特征图。



Ave Pool 为平均池化;Max Pool 为最大池化;Softmax 为归一化指数函数;Conv 为卷积;Concat 为拼接操作;Sigmoid 为激活函数;Re-weight 为分配权重;Matmul 为矩阵相乘函数;Input 为输入;Output 为输出;Groups 为分组;GroupNorm 为组归一化;Cross-spatial learning 为跨空间学习

图4 MEMA 模块结构

Fig. 4 MEMA module structure

2.2 构建软池化 OD-SMP 模块

煤矿井下环境的图像含有的特征细节信息不明显,且易与背景信息混淆,主干网络在提取安全帽和头部特征时,MP 模块中的最大池化会丢失一部分本就不丰富的目标细节特征,造成对目标有效特征的提取能力不足。为了减少 MP 模块的信息丢失,增强模型对目标的特征提取能力,本研究利用 SoftPool,增强检测器对受到噪声影响的不明显特征信息的辨识能力,更好地捕捉小目标的细节特征,对于小目标的检测尤其有帮助;引入全维动态卷积

ODConv,替代原有卷积层,从多个维度捕捉关键的上下文信息,然后为每个分支提取的特征图分配可训练的权重值,协助模型关注对输出特征贡献度大的信息,从而构建了一种联合动态卷积的加权采样模块(OD-SMP),提升模型在煤矿井下尺度变化剧烈场景中的识别能力,其结构如图5所示。

2.2.1 SoftPool 池化

SoftPool 采用一种独特的加权池化方法,即利用 Softmax 函数进行操作,这使得它与传统的池化技术有所不同。SoftPool 不仅能保持特征的表达性,而且其操作是可微的,确保在神经网络训练过程中,每一次反向传播的梯度都能得到有效更新。SoftPool 可以让池化内核中的每一个激活因子被充分利用,这一特性仅需要极少的额外内存开销。同时,增大了相近特征信息的区分度,又保留了整体感受野的特征信息,从而显著提高了算法的准确率。

SoftPool 以自然指数 e 作为底数,使得较大的激活值在输出中占据更大的影响力。更重要的是,SoftPool 的可微性使得在反向传播过程中,局部邻域 R 内的每个激活都能至少获得一个最小的梯度分配,从而保证了梯度更新的全面性和有效性。在 SoftPool 的池化过程中,每一个激活 a_i 都会被赋予

一个权重 w_i ,该权重为该激活的自然指数相对于邻域 R 内所有激活的自然指数之和的比值,可表示为

$$w_i = \frac{e^{a_i}}{\sum_{j \in R} e^{a_j}} \tag{1}$$

通过对内核邻域内所有激活的加权求和得到 SoftPool 输出值为

$$\bar{a} = \sum_{i \in R} w_i a_i \tag{2}$$

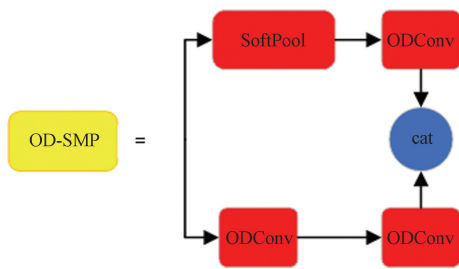
2.2.2 全维动态卷积 ODConv 模块

全维动态卷积^[11]采用多维注意力并行策略,通过在卷积核的4个维度上并行地分配注意力权重,打破传统卷积固有感受野的束缚,通过捕捉和强化显著表型特征,帮助模型适应不同尺寸和形状的特征,更好地捕捉安全帽的细节特征。

如图6所示,ODConv 模块在处理输入特征 X 时,利用全局平均池化层(GAP)对其进行空间维度的压缩,首先将特征图转化为通道数为 C_{in} 的特征向量,其次采用一个全连接层(FC)将特征向量映射成低维空间,然后采用 ReLU 激活函数对特征向量中的负值进行剔除,以筛选出关键信息。构建了4个独立的头部分支,每个分支专注于生成特定类别的注意力标量,4种注意力分支可以帮助模型提取丰富的上下文信息^[12]。全维动态卷积表达式为

$$y = (\alpha_{a1} \odot \alpha_{b1} \odot \alpha_{c1} \odot \alpha_{d1} \odot W_1 + \alpha_{a2} \odot \alpha_{b2} \odot \alpha_{c2} \odot \alpha_{d2} \odot W_2 + \dots + \alpha_{an} \odot \alpha_{bn} \odot \alpha_{cn} \odot \alpha_{dn} \odot W_{1n}) * X \tag{3}$$

式(3)中: $X \in R^{H \times W \times C_{in}}$, $y \in R^{H \times W \times C_{out}}$ 分别为输入和输出特征,其中, H 和 W 分别为特征的长、宽, C_{in} 和 C_{out} 为通道数; $W_i \in R^{k \times k \times C_{in}}$ 为输出滤波器的第 i 个卷积核; $\alpha_{ai} \in R^{K \times K}$ 为沿着卷积核 W_i 核空间的注意力标量; $\alpha_{bi} \in R^{C_{in}}$ 为沿着卷积核 W_i 输入通道维度的注意力标量; $\alpha_{ci} \in R^{C_{out}}$ 为沿着卷积核 W_i 输出通道维



红色模块为改进部分;SoftPool 为软池化;ODConv 为全维动态卷积;cat 为拼接操作

图5 OD-SMP 模块结构图

Fig.5 OD-SMP module structure

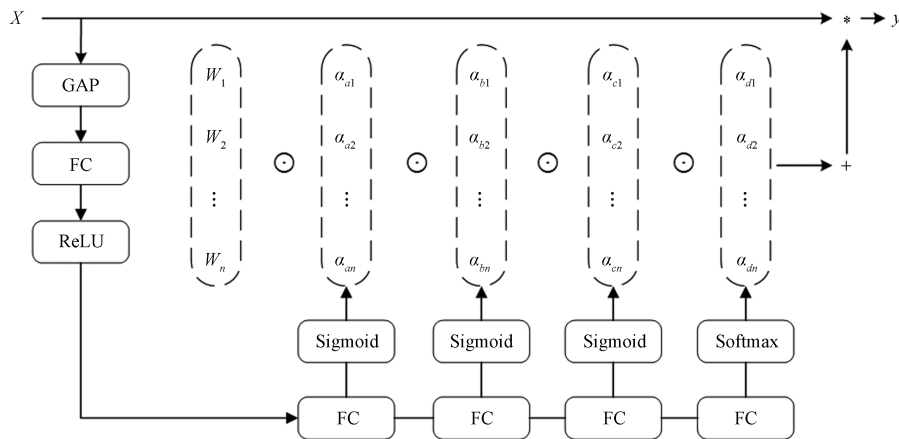


图6 全维动态卷积计算流程图

Fig.6 Full dimensional dynamic convolution calculation flowchart

数的注意标量; $\alpha_{di} \in R$ 为沿卷积核数的注意标量; \odot 为沿核空间不同维数的乘法运算; $*$ 为卷积运算。

全维动态卷积通过并行分支计算为卷积核沿核空间 4 个维度捕捉有效的上下文信息, 4 种注意力权重按照特定的顺序逐步累加到卷积核上, 实现对不同大小的物体的动态调整^[13]。

2.3 Wise-IoU 损失函数

在 YOLOv7 中, 其损失函数 CIoU 是通过综合边界框回归损失、置信度损失以及分类概率损失来计算得出的。CIoU^[14] 的设计在特定情境下存在局限性, 即在设定边框长宽比例时, 若使长宽值相等但方向各异, 会导致其兼顾的长宽比部分失效 ($v = 0$)。此外, CIoU 在处理长宽变化时通常呈现负相关性, 即一边增加时另一边减少, 这种趋势在边界框预测中易引发目标预测框与真实框的不匹配问题。上述两方面因素共同导致了模型预测框与真实目标框之间的不匹配, 进而严重影响了 CIoU 的收敛性能, 最终导致在安全帽佩戴检测任务中检测效果的不理想。引入损失函数 Wise-IoU 代替原损失函数进行模型优化, 由于煤矿井下不同光照和距离的复杂背景和环境, 导致检测样本质量出现参差不齐的情况, 而 Wise-IoU v3 版损失函数能平衡这种情况^[15]。Wise-IoU 作为一种改进型的损失函数, 在传统 IoU 基础上融入了类别权重的概念。其主要目标在于通过精细化的权重调整来缩小不同类别之间的性能差异, 从而优化模型对测试结果的预测准确性。具体实现上, Wise-IoU 为每个类别分配特定的权重值, 并在计算 IoU 重叠度时应用这些权重, 以实现不同类别间重叠区域的差异化评估, 最终得到更为精确和可靠的模型性能评价结果。Wise-IoU 共有 3 个版本, 即 Wise-IoU v1、Wise-IoU v2、Wise-IoU v3。

由于训练数据中不可避免地存在低质量样本, 传统的几何度量 (如距离、纵横比) 在评估这些样本时可能会过度惩罚, 进而造成模型的泛化性能下降。一个理想的损失函数设计应旨在当锚框与目标框达到良好匹配时, 适度减少对几何度量误差的惩罚力度, 避免对训练过程造成不必要的干扰, 进一步提升模型的性能, 强化了模型的泛化能力。基于这一改进, 引入一种基于距离度量的注意力机制, 即距离注意力, 进而构建了具备双层注意力结构的 Wise-IoU v1, 以增强模型在煤矿井下复杂场景下的适应性。

$R_{WIoU} \in [1, e]$, R_{WIoU} 为记锚框和目标框中心点连接的归一化长度, 这将显著放大普通质量锚框的 L_{IoU} 。

$L_{IoU} \in [0, 1]$, L_{IoU} 为度量目标检测任务中预测框与真实框的重叠程度, 这将大幅度减少高质量锚框 R_{WIoU} 的冗余性, 尤其是在锚框与目标框实现高度匹配时, 更能有效减轻对中心点距离这一单一维度的过度依赖, 进而优化模型的效能与精确度。

$$L_{WIoUv1} = R_{WIoU} L_{WIoU} \quad (4)$$

$$R_{WIoU} = \exp \left[\frac{(x - x_{gt})^2 + (y - y_{gt})^2}{(W_g^2 + H_g^2)^*} \right] \quad (5)$$

式中: L_{WIoU} 为 Wise-IoU 度量目标检测任务中预测框与真实框的重叠程度; L_{WIoUv1} 为 Wise-IoU v1 度量目标检测任务中预测框与真实框的重叠程度; W_g 、 H_g 为最小的封闭框的宽度和长度; x 、 y 和 x_{gt} 、 y_{gt} 分别为最小的封闭框内记锚框和目标框大小; 为了防止放大普通质量锚框参数 R_{WIoU} 对阻碍收敛的影响, 将 W_g 、 H_g 从计算中分离 (上标 * 表示此操作)。

未引入新的度量指标的原因在于: 它能够有效消除了阻碍收敛的因素。

在 Wise-IoU v2 中, 受到 Focal Loss 设计思路的启发, 对 Wise-IoU v1 进行优化, 形成一个独特的单调聚焦系数 L_{IoU}^* 。这一改进显著降低了简单样本在损失函数中的权重, 从而引导模型更加专注于那些难以处理的复杂样本。这一策略显著增强了模型对困难示例的识别能力, 进而实现了分类性能的显著提升, L_{WIoUv2} 为 Wise-IoU v2 度量目标检测任务中预测框与真实框的重叠程度, 可表示为

$$L_{WIoUv2} = L_{IoU}^* L_{WIoUv1} \quad (6)$$

在模型训练过程中, 随着 L_{IoU} 的递减, L_{IoU}^* 也呈现下降趋势, 这一现象导致训练后期收敛进程趋于迟缓。鉴于此, 引入的均值 L_{IoU} 作为归一化基准, 可表示为

$$L_{WIoUv2} = \left(\frac{L_{IoU}^*}{L_{IoU}} \right)^\gamma L_{WIoUv1} \quad (7)$$

式(7)中: γ 为调节因子, 控制权重随 IoU 变化的敏感度。

在锚框的评估中, 若其离群度偏低, 这往往标志着其质量较高。鉴于此, 为其设定一个较低的梯度增益权重, 旨在确保边界框回归过程能够优先优化这些普通质量锚框的精确性, 从而更有效地提升整体性能。而针对那些离群度显著较高的锚框, 赋予其较小的梯度增益是一种合理的策略, 这有助于防止这些低质量样本在训练过程中产生过大的、潜在的负面影响的梯度。

Wise-IoU v3 利用离群度 β 构造一个非单调聚焦系数 r , 并将其应用于 Wise-IoU v1, 可表示为

$$r = \frac{\beta}{\delta \alpha^{\beta-\delta}} \quad (8)$$

式(8)中: α 和 δ 为超参数。

$$\beta = \frac{L_{IoU}^*}{L_{IoU}} \in [0, +\infty) \quad (9)$$

$$L_{WIoUv3} = rL_{WIoUv1} \quad (10)$$

式(9)定义了离群度 β 作为衡量回归框品质的一个指标,通过引入动量为 m 的滑动平均值 \bar{L}_{IoU} ,该值被动态地维护与更新,旨在维持整体性能 β 于较高水平,从而有效缓解训练末期收敛速率减缓的问题。在此框架下,超参数 α 与 δ 共同作用于决策过程:当某一回归框的离群度 β 达到预设阈值 C 时,该框将享有最大的梯度增益优势。值得注意的是,由于滑动平均值 \bar{L}_{IoU} 的动态特性,回归框品质的评判标准也随之灵活调整,确保了Wise-IoU v3能够实时制定最优的梯度增益分配策略,以适应训练过程中的各种情况。此外,Wise-IoU v3还特别关注于中等质量锚框的优化,进一步提升模型的定位精确度与性能。

3 实验方法及结果分析

3.1 环境与数据集

使用从某煤矿上收集的图片,有两种图片类型,分别为安全帽和人,共2000张。使用Labelimg对数据集中的图像进行了标注。随机以8:2的比例分配为1600张训练集和400张验证集,实验环境如表1所示。

3.2 训练网络模型

在模型训练的过程中,输入图片尺寸设置为 640×640 ,batch size设置为8,设定了动量因子为

0.9,训练的学习率为0.001。通过这样的设置,进行300轮的模型迭代训练。

3.3 消融实验

为了证实本研究所提出的MELAN、OD-SMP和WIoU的有效性,使用相同的数据集进行训练和验证,分别进行8组消融实验,结果如表2所示。

表1 实验环境

Table 1 Experimental environment

配置名称	版本参数
操作系统	Ubuntu 18.04
GPU	NVIDIA GeForce RTX 3060
CPU	Intel Core i9-12900H
RAM/G	16
深度学习框架	PyTorch
CUDA 版本	11.6
Python 版本	3.8

表2 消融实验结果

Table 2 Results of ablation experiment

YOLOv7	MELAN	OD-SMP	Wise-IoU	mAP/ %	mAP@0.5/ %	GFLOPs
✓	—	—	—	81.4	82.6	26.7
✓	✓	—	—	83.9	82.9	30.0
✓	—	✓	—	83.1	83.2	23.8
✓	—	—	✓	86.3	85.2	26.7
✓	✓	—	✓	89.5	88.3	30.0
✓	—	✓	✓	90.9	87.6	23.8
✓	✓	✓	—	91.2	90.1	27.2
✓	✓	✓	✓	94.9	93.5	27.2

注:✓表示模型包含该模块;“—”表示模型不包含该模块;mAP为平均精度均值;mAP@0.5为IoU阈值为0.5的平均精度;GFLOPs为十亿次浮点预算量。



图7 检测效果图

Fig.7 Detection effect diagram

由表 2 可知,原始网络对安全帽佩戴检测的平均准确率均值为 81.4%;在不改变模型原始损失函数的情况下,将 Backbone 部分加入 MELAN 模块,检测模型的平均准确率均值提升 2.5 个百分点,参数量增加 1.1%,原因在于:MEMA 注意力模块增强有用的特征并抑制无用的特征;将 Backbone 部分加入 OD-SMP 模块后平均准确率均值提升 1.7 个百分点,表明 SoftMax 池化可以减少 MP 模块中的最大池化会丢失一部分本就不丰富的目标细节特征,造成对小目标有效特征的提取能力不足;使用 Wise-IoU 作为回归损失函数后,由于 Wise-IoU 能平衡检测样本质量出现参差不齐的情况,模型检测性能达到最佳,平均准确率均值提升至 94.9%。检测效果如图 7 所示。

3.4 对比实验

为了全面评估改进模型在检测任务中的性能,选取当前主流的几种模型进行对比分析。在确保实验条件与数据集保持一致的前提下,进行了全面的对比。实验结果如表 3 所示。结果表明,改进后的模型具有良好的检测效果。改进前后 mAP 曲线如图 8 所示。

表 3 不同算法对比效果结果

Table 3 Comparison results of different algorithms

不同算法	mAP/%	mAP@0.5/%	GFLOPs
SSD	58.4	49.2	343.37
Fast-RCNN	66.5	51.9	193.78
YOLOv5	75.0	38.0	15.80
YOLOv7	81.4	82.6	26.70
YOLOv7-A	93.5	91.7	28.30
YOLOv8	88.8	86.8	8.20
本文算法	94.9	93.5	27.20

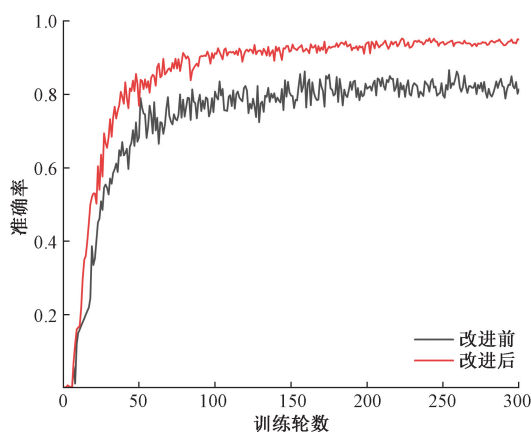


图 8 改进前后 mAP 曲线

Fig. 8 mAP curves before and after improvement

4 结论

(1) 针对煤矿工人安全帽佩戴检测环境面临的

低质影像存在低光照、小目标等痛点,提出一种基于改进 YOLOv7 的煤矿井下低质影像安全帽检测,提高了煤矿工人安全帽佩戴检测的精度,与 YOLOv7 相比具有较好的检测效果。本文算法加入 MELAN 模块,增强特征采样的感受野,提升浅层细节特征的利用率。改进原有 MP 模块的池化方法,全维动态卷积替代原有卷积层,进一步增强模型对小目标的特征提取能力。使用 Wise-IoU 损失函数替换原有的损失函数,应对井下恶劣多变的环境导致数据集质量参差不齐的情况。为验证方法的有效性,实验在自制数据集上完成。

(2) 本文算法相较于原算法准确率提升 13.5%。与 SSD、Fast-RCNN、YOLOv5、YOLOv7-A 和 YOLOv8 相比准确率分别提升 36.5%、28.4%、19.9%、1.4% 和 6.1%;在模型参数计算量上,与 SSD 和 Fast-RCNN 均大幅度减少;与 YOLOv5、YOLOv7 和 YOLOv8 相比较,参数计算量略有增加,但 mAP 和 mAP@0.5 效果较好。因此,改进 YOLOv7 煤矿井下安全帽佩戴检测算法具有良好性能,能满足对煤矿井下安全帽佩戴的检测。

参 考 文 献

- [1] Li X, Wang S, Liu B, et al. Improved YOLOv4 network using infrared images for personnel detection in coal mines[J]. Journal of Electronic Imaging, 2022, 31(1): 013017.
- [2] 徐先峰, 赵万福, 邹浩泉, 等. 基于 MobileNet-SSD 的安全帽佩戴检测算法[J]. 计算机工程, 2021, 47(10): 298-305, 313. Xu Xianfeng, Zhao Wanfu, Zou Haoquan, et al. Safety helmet wearing detection algorithm based on MobileNet-SSD[J]. Computer Engineering, 2021, 47(10): 298-305, 313.
- [3] 李明山, 韩清鹏, 张天宇, 等. 改进 SSD 的安全帽检测方法[J]. 计算机工程与应用, 2021, 57(8): 192-197. Li Mingshan, Han Qingpeng, Zhang Tianyu, et al. Improving the safety helmet detection method for SSDs[J]. Computer Engineering and Applications, 2021, 57(8): 192-197.
- [4] 李嘉信, 胡杨, 黄协舟, 等. 面向小目标的多空间层次安全帽检测[J]. 计算机工程与应用, 2024, 60(6): 230-237. Li Jiaxin, Hu Yang, Huang Xiezhou, et al. Multi spatial hierarchical safety helmet detection for small targets[J]. Computer Engineering and Applications, 2024, 60(6): 230-237.
- [5] 丁田, 陈向阳, 周强, 等. 基于改进 YOLOX 的安全帽佩戴实时检测[J]. 电子测量技术, 2022, 45(17): 72-78. Ding Tian, Chen Xiangyang, Zhou Qiang, et al. Real time detection of helmet wearing based on improved YOLOX[J]. Electronic Measurement Technology, 2022, 45(17): 72-78.
- [6] Shao X, Liu S, Li X, et al. FM-YOLOv7: an improved detection method for mine personnel helmet[J]. Journal of Electronic Imaging, 2023, 32(3): 033013.
- [7] 高凯, 董立红, 邓凡. 基于递归门控卷积和上下文注意力的煤块检测算法[J]. 矿业研究与开发, 2023, 43(6): 183-190. Gao Kai, Dong Lihong, Deng Fan. Coal block detection algorithm

- based on recursive gated convolution and contextual attention[J]. Mining Research and Development, 2023, 43(6): 183-190.
- [8] 宋立业, 赵小萱, 崔昊. 基于改进 YOLOv7 的矿用电铲检测算法[J]. 工矿自动化, 2023, 49(12): 18-24, 32.
Song Liye, Zhao Xiaoxuan, Cui Hao. Mining electric shovel detection algorithm based on improved YOLOv7[J]. Industrial and Mining Automation, 2023, 49(12): 18-24, 32.
- [9] Ouyang D L, He S, Zhang G Z, et al. Efficient multi-scale attention module with cross-spatial learning [C]//IEEE International Conference on Acoustics, Speech and Signal Processing. New York: IEEE, 2023: 1-5.
- [10] 顾清华, 周宇静, 王丹, 等. 基于多尺度特征提取的露天矿区道路坑洼检测模型[J/OL]. (2025-04-24) [2024-06-18]. <http://kns.cnki.net/kcms/detail/34.1055.td.20240516.1134.006.html>.
Gu Qinghua, Zhou Yujing, Wang Dan, et al. A detection model for road potholes in open-pit mining areas based on multi-scale feature extraction[J/OL]. (2025-04-24) [2024-06-18]. <http://kns.cnki.net/kcms/detail/34.1055.td.20240516.1134.006.html>.
- [11] Chao L, Aojun Z, Anbang Y. Omni-dimensional dynamic convolution[J]. arXiv Preprint, 2019: arXiv: 2209.07947.
- [12] 崔丽群, 曹华维. 改进 YOLOv7 的航拍图像目标检测[J]. 计算机工程与应用, 2024, 60(20): 189-197.
Cui Liqun, Cao Huawei. Improving YOLOv7 for aerial image object detection[J]. Computer Engineering and Applications, 2024, 60(20): 189-197.
- [13] 王俊岭, 王晨晨, 熊玉华. 基于 YOLOv5l-Im 的排水管道缺陷检测方法的效果分析[J]. 科学技术与工程, 2024, 24(18): 7833-7842.
Wang Junling, Wang Chenchen, Xiong Yuhua. Drainage pipeline defect detection method and effect analysis based on YOLOv5l-Im [J]. Science Technology and Engineering, 2024, 24(18): 7833-7842.
- [14] 秦子豪, 雷鸣, 宋文广, 等. 基于轻量化深度学习模型的安全帽检测方法[J]. 科学技术与工程, 2022, 22(14): 5659-5665.
Qin Zihao, Lei Ming, Song Wenguang, et al. Safety helmet detection method based on lightweight deep learning model[J]. Science Technology and Engineering, 2022, 22(14): 5659-5665.
- [15] 桑雨, 李立权, 李铁. 轻量化 YOLOv7-tiny 的遥感图像小目标检测[J]. 科学技术与工程, 2024, 24(18): 7726-7732.
Sang Yu, Li Liqun, Li Tie. Small target detection in remote sensing images of lightweight YOLOv7-tiny[J]. Science Technology and Engineering, 2024, 24(18): 7726-7732.