

基于YOLO v11的轻量化安全帽佩戴检测算法

向德怀

上海理工大学光电信息与计算机工程学院, 上海

收稿日期: 2025年9月18日; 录用日期: 2025年9月28日; 发布日期: 2025年12月19日

摘要

在建筑、采矿、勘探等工地施工场景中, 安全帽佩戴检测算法是预防工人生产管理安全的重要手段, 但现有目标检测模型存在检测精度低、参数量多等问题。本研究提出了一种基于YOLO v11框架LH-YOLO模型, 旨在实现精度与效率的高度平衡。首先使用StarNet作为骨干网络, 大幅降低计算复杂度; 其次, 设计了C3k2_SN模块特征融合颈部, 模块采用星形跨阶段部分卷积, 高效利用参数的同时, 增强了特征表达的能力; 最后构建了LSCD轻量化共享卷积头, 并引入PIoUv2损失函数对边界框定位进行精细优化, 进一步压缩检测头参数并优化预测。实验结果表明, 在公开数据集SHWD (Safety Helmet Wearing Dataset)上进行实验, 改进模型精确率达到93.4%, 与原模型参数量降低了32.8%、十亿次的浮点计算量(GFLOPs)下降了34.8%、平均精度提升1.4%, 兼顾高检测性能和实时性要求, 更适用于实际生产环境的部署与应用。

关键词

YOLO 11n, 安全帽佩戴检测, 目标检测, 深度学习

A Lightweight Safety Helmet Wearing Detection Algorithm Based on YOLO v11

Dehuai Xiang

School of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology, Shanghai

Received: September 18, 2025; accepted: September 28, 2025; published: December 19, 2025

Abstract

In construction, mining, exploration, and other construction site scenarios, the safety helmet wearing detection algorithm is an important means to prevent safety incidents in worker production management. However, existing object detection models suffer from issues such as low detection accuracy and a large number of parameters. This study proposes an LH-YOLO model based on the

YOLO v11 framework, aiming to achieve a high balance between accuracy and efficiency. Firstly, StarNet is adopted as the backbone network, which significantly reduces computational complexity. Second, a feature fusion neck with the C3k2_SN module is designed. The module adopts star-shaped cross-stage partial convolution, which efficiently utilizes parameters while enhancing the capability of feature expression. Finally, an LSCD lightweight shared convolutional head is constructed, and the PloUv2 loss function is introduced to finely optimize bounding box localization, further compressing the parameters of the detection head and optimizing predictions. Experimental results show that when tested on the public dataset SHWD (Safety Helmet Wearing Dataset), the precision of the improved model reaches 93.4%. Compared with the original model, its parameter count is reduced by 32.8%, Giga Floating-Point Operations (GFLOPs) decrease by 34.8%, and average precision (AP) increases by 1.4%. The model balances high detection performance and real-time requirements, making it more suitable for deployment and application in actual production environments.

Keywords

YOLO 11n, Safety Helmet Wearing Detection, Object Detection, Deep Learning

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

在复杂多变的建筑工地环境中, 工地存在高空坠物等风险[1]。因此, 安全帽作为关键的工人防护设备, 在各类行业中对减轻头部伤害、保障工人健康发挥着至关重要的作用[2]。

通过人工或监控检查工人遵守安全帽佩戴要求, 这种传统方式存在效率低、速度慢、成本高且误检率高的问题, 难以对大面积区域有效监管; 而基于视觉系统与深度学习技术相互结合的现代化方式以低成本、易部署、高效率的优势成为主流方案。因此, 研发高性能的基于深度学习安全帽检测算法具有很高社会、学术价值。目前已有众多学者针对安全帽检测展开研究, 当前的安全帽识别算法主要分为双阶段检测方法和单阶段检测方法。双阶段检测方法以 Fast R-CNN、Mask R-CNN、U-Net 为代表, 通过两个独立的网络模型分阶段检测, 在检测大目标和复杂场景时有精度优势, 但其计算量大、速度慢, 无法满足实时应用需求; 单阶段目标检测方法以 You Only Look Once (YOLO) 系列和 Single Shot Multi Box Detector (SSD) 为代表, 通过单次计算完成目标位置估计与类别分类, 检测速度较快, 简化了检测流程, 提供了更快速的解决方案。

在单阶段检测方法中, YOLO (You Only Look Once) 系列算法兼顾精度与处理速度, 基于 YOLO 的改进算法在安全帽佩戴检测领域广泛研究[3]。例如, 董晨航等[4]提出基于 YOLOv5 的 PG-YOLO 改进模型, 通过全网络 Ghost 模块替换标准卷积模块, 引热议 LightC3 消除 Backbone 冗余, 避免过度压缩导致精度骤降, 创新 R-pruning 剪枝后通道数恢复机制, 并设计知识蒸馏(KD), 抵消轻量化导致的精度损失, 模型在精度仅损失 0.1% 情况下大幅减少模型权重、提升推理速度。王丽丽等[5]提出基于 YOLO-M, 以 YOLOv5s 为基础, 替换骨干网络为 MobileNetv3, 实现网络高效特征提取及轻量化, 设计 Res-FPN 特征融合模块, 解决小目标特征丢失问题, 提出 BiCAM 注意力模块, 增强密集遮挡目标特征表达, 精度达到 94.63%。韩大光等[6]提出 YOLOv8s-SNC 改进模型, 新增专用小目标检测层, 引入 SPD-Conv 模块优化特征提取过程中信息损, 设计 SEResNeXt 检测头增强特征表达能力, 嵌入 C2f-CA 模块强化关键信息捕

捉与背景干扰抑制能力,提升复杂环境中工业场景下安全帽检测能力;范志鹏等[7]提出轻量化 LG-YOLOv8 改进模型,提出 C2f-GhostDynamicConv 模块,创新融合 Ghost 模块的轻量化特性与动态卷积的自适应特征提取能力,实现“减参提效 + 特征增强”双重目标;引入双向特征金字塔(BiFPN),提升多尺度安全帽检测的特征融合能力;设计轻量化不对称检测头(LADH-Head),减少参数冗余问题;林炳炎等[8]提出 YOLOv8n-ASF-DH 改进模型,骨干网络集成 Triplet Attention 三并行分支注意力机制,强化小目标特征聚焦,颈部引入 ASF 结构,优化多尺度特征融合效率,检测头替换为 DyHead,动态提升多场景适配能力,采用 Focal-EIoU 损失函数,平衡高低质量样本损失贡献,通过多模块协同优化,实现复杂场景下小目标检测性能跃升。张世豪等[9]针对低光环境,构建 MHWD 矿山专属安全帽数据集,提出 BLP-YOLOv10 改进模型,骨干网络通过“通道缩减 + Biformer”实现轻量化与精度平衡,颈部集成低频率增强滤波器(LEF),突破低光场景检测瓶颈,采用 Power-IoU 损失函数,优化锚框回归与训练效率提升,实现“高精度 - 轻量级 - 低光稳健”的检测目标。

尽管此类算法已取得诸多进展,但基于 YOLO 的算法在安全帽检测场景中,面对复杂背景下的小目标检测仍面临挑战。例如,包含建筑物、植被的背景可能会干扰安全帽的检测与识别;同时,安全帽检测算法还存在参数冗余、低算力平台部署困难的问题,模型轻量化迫在眉睫。

YOLOv11 算法在 YOLOv8 的基础上进行了创新性改进,在目标检测任务中表现出卓越性能,实现了精度与速度的突破性平衡[10] [11]。尽管针对安全帽检测的 YOLO 算法已有所改进,但单阶段算法在检测小目标或面临复杂背景干扰时,精度仍有提升空间。因此,迫切需要对算法进行改进,以实现复杂环境下的更优检测性能。

本研究在 YOLOv11 提出了 LH-YOLO 模型,旨在实现精度与效率的高度平衡,主要工作如下:

- (1) 使用 Starnet 网络作为骨干网络,大幅降低计算复杂度;
- (2) 设计了 C3k2_SN 卷积模块特征融合颈部,增强了参数利用率和特征表达能力;
- (3) 构建了 LSCD 轻量化共享卷积头,并引入 PIoUv2 损失函数对边界框定位进行精细优化,进一步压缩检测头参数并优化预测;
- (4) 在实验中对改进模型的性能进行了评估,包括检测精度、推理速度和计算效率等方面;
- (5) 与主流方法进行了对比实验,结果表明,改进方法在保持高效的同时显著提高了检测性能。

2. 改进的 LH-YOLO 算法

2.1. YOLO11n 模型

YOLOv11 是 Ultralytics 于 2024 年发布的最新一代实时目标检测模型,延续了 YOLO 系列高效推理的核心理念,重点优化了特征提取效率和多尺度目标感知能力,考虑到轻量化需求使用去中 YOLO11n 作为基准模型,具体的网络结构如图 1 所示。

YOLOv11n 框架上延续了经典的“Backbone-Neck-Head”三段式架构,数据预处理后输入网络先后执行特征提取 - 特征融合 - 检测头输出任务,并通过加权 NMS (非极大值抑制)生成最终检测结果,该架构保持了 YOLO 系列端到端推理的优势,同时通过模块创新提升了特征表征能力。

在骨干网络部分,框架首次引入 C3K2 (Cross Stage Partial with kernel size 2)模块替代 YOLOv8 的 C2f 模块。C3K2 结合了 C2f 的速度优势和 C3k 的灵活性,外层采用 C2f 结构,内层嵌入 C3 的框架,如下图所示。外层上继承 C2f 模块,通过减少卷积层的数量和采用更高效的特征合并策略来提高速度;内层上 C3K 模块,其通过 `c3k=True` 参数动态调整 shortcut 连接,实现特征复用与梯度传播优化。并且在处理需要不同感受野的场景时,可以通过调整 C3k 中 `kenral` 参数来适应特定的特征提取需求,提供了很高的可配置性。具体到安全帽检测任务,考虑计算复杂度与性能的平衡,这里 C3K2 使用 `k=3`、`c3k=True` 参数

与的默认版本,此时 C3K2 的组合结构如下图所示。综合来看,相较于 C2F 模块, C3K2 模块具备以下优势:减少 15~20%的参数数量,降低计算复杂度;增大有效感受野,提升小目标检测能力;优化梯度流,缓解深层网络训练困难问题。

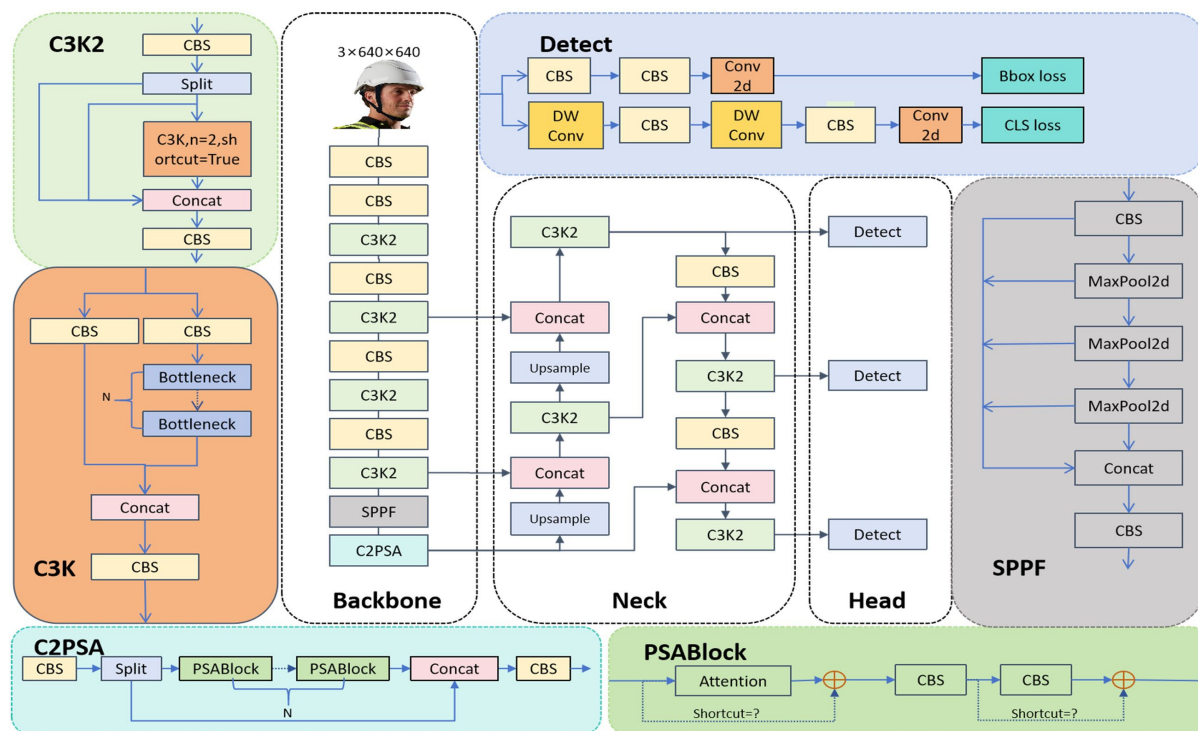


Figure 1. YOLO11 network structure

图 1. YOLO11 网络结构

在颈部网络部分,框架引入 C2PSA (Convolutional block with Parallel Spatial Attention)模块代替 YOLOv10 中的 PSA (Partial Self Attention)模块。C2PSA 结合了 CSP 结构和部分自注意力机制,提升了多尺度特征提取能力,结构如下图所示。外层结构上继承了 CSP 的分段特征处理思想,初始特征经过一次 1×1 卷积后进行分割,需要关注的部分通过多层 PSA 模块迭代捕捉深层特征依赖,与另一部分进行特征拼接,后经过一次 1×1 卷积来恢复原始通道数。不同于 C2F 的全过程输出, C2PSA 仅保留最后一层的输出,在降低计算冗余的同时,也让多层多头注意力机制得到的特征输出更凝练,避免特征混淆。内层 PSA 模块采用十分简洁的串联结构,包含注意力层和前馈层,通过多种卷积核(如 3×3 、 5×5 、 7×7 等)来提取多尺度特征。不同卷积核的卷积操作并行进行,之后将特征图拼接,并使用 SE (Squeeze-and-Excitation)模块为特征通道加权。最后,通过 Softmax 生成的注意力权重应用到各个特征图上,从而实现通道逐点加权(Channel-wise multiplication),提升对重要特征的关注度。总体上 C2PSA 模块借助初始分割后的卷积操作,将输入特征分流为两条路径,一条是直接传递特征的快捷路径(shortcut),另一条是进行注意力机制复杂特征转换的深度处理路径;通过这种双路径设计,能够在不额外增加时间开销的前提下,实现对深层特征的高效提取。

在检测头部分, YOLO11 采用分类分支轻量化设计,引入深度可分离卷积(DWConv)以减少冗余计算,减少 40%参数数量,从而提高计算效率。同时支持多任务输出,通过共享特征图实现检测(边界框 + 类别)、分割(掩码)、姿态(关键点)的统一输出,进一步提升了模型的运行速度和性能。

整体而言, YOLO11 的改进在多层次提升了模型的灵活性、特征提取能力和计算效率。

2.2. 轻量级 StarNet 骨干网络

YOLO11 主干网络参数量大、计算开销高, 于实际工地应用场景中难以适配。针对上述问题, 本文引入 StarNet 作为轻量化主干网络, 其核心创新在于用元素级乘法替代传统特征融合中的加法操作, 在不增加显式参数与计算量的前提下, 大幅提升网络对非线性特征的表达能力。从数学原理来看, 乘法操作 $(W_1 X) * (W_2 X)$ 蕴含特征二次交互机制——无需额外增加卷积核或拓宽通道, 即可将 d 维输入特征隐式映射至 $O(d^2)$ 维的高维非线性空间, 若通过多层级联, 还能实现隐式维度的指数级扩展。这种类似多项式核函数的数学特性, 让 StarNet 能以极低的参数量和计算复杂度, 处理工地场景中目标与背景的复杂特征关系, 为资源受限环境下的检测任务提供了适配性更强的解决方案。

StarNet 结构上通过卷积和星型乘法的堆叠实现函数非线性表达能力提升, 简节的模型设计可直接与 YOLO11 集成, StarNet 结构如图 2 所示: 首先通过 1 个 3×2 卷积层降低特征图分辨率, 减少后续计算压力; 接着通过 5 个功能块实现深度特征提取, 每个功能块均由 3×2 卷积与 Star Block 组成, 且 Star Block 的深度(即重复次数)可根据实际任务需求单独设置, 大幅提升了网络在不同场景下的适配灵活性。作为 StarNet 的核心组件, Star Block 的特征处理流程围绕轻量化进行(结构如图 2 所示): 首先采用 7×1 的深度卷积(DWConv)对输入特征进行初步提取, 深度卷积通过将空间卷积与通道卷积分离, 显著减少了计算量; 随后通过批归一化(BN)对特征进行标准化处理, 避免特征分布偏移导致的训练不稳定问题; 接着通过两路扩展因子为 4 的 1×1 卷积完成特征升维, 其中一路特征经 ReLU6 激活函数引入非线性, 另一路保持线性特征, 同时可兼顾特征的线性表达与非线性变换; 之后对两路输出执行元素级乘法(即星形操作), 生成 $O((4d)^2)$ 使得低维特征中蕴含更丰富的语义信息; 最后依次通过 1×1 卷积将通道数恢复至输入维度、 7×1 深度卷积进一步提取关键特征, 进行残差连接, 缓解梯度消失, 增强浅层特征传递。

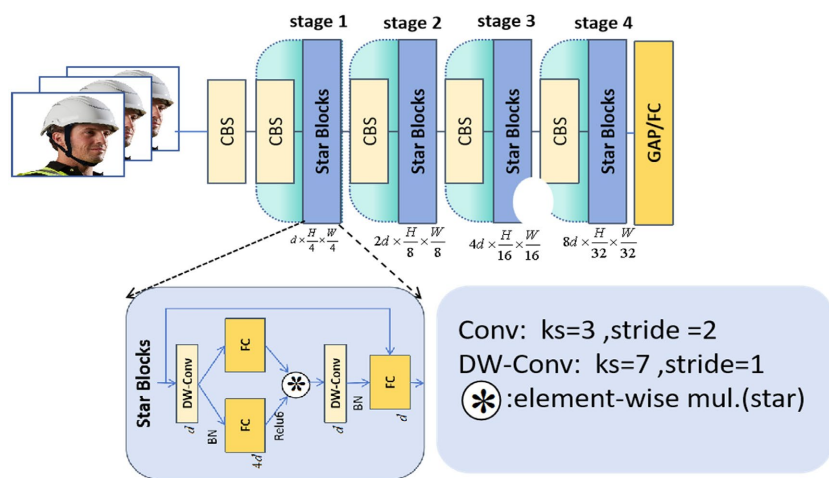


Figure 2. StarNet network structure

图 2. StarNet 网络结构

从理论层面可进一步拆解 StarNet 的特征扩展机制, 首先看单层级星形操作对隐式维度的提升: 记网络输入特征为 z , 初始特征维数为 d , 星形操作的基本形式是对两路线性变换后的特征执行元素级乘法, 即:

$$\text{Star}(X) = (W_1^T X) * (W_2^T X) \quad (1)$$

其中 $X \in R^{(d+1) \times n}$ 、 d 为输入通道数, n 为特征元素数, 含偏置项时 X 扩展为 $([X; 1])$, $W_1, W_2 \in R^{(d+1) \times (d'+1)}$ 为线性变换权重。

在单通道输出、单元素输入的场景下, 设 $w_1, w_2, x \in R^{(d+1) \times 1}$, 则星形操作可展开为:
 $w_1^T x * w_2^T x = (\sum_{i=1}^{d+1} w_1^i x^i) * (\sum_{j=1}^{d+1} w_2^j x^j)$, 进一步展开后得到二次项求和公式 $\sum_{i=1}^{d+1} \sum_{j=1}^{d+1} w_1^i w_2^j x^i x^j$ 。合并同类项后($i=j$ 时为平方项, $i \neq j$ 时为交叉项), 可得到 $\frac{(d+2)(d+1)}{2}$ 个独立项, 当 $d \gg 2$ 时, 独立项数量近似为 $O(d^2)$, 意味着单层级星形操作即可将特征隐式映射至 $O(d^2)$ 维空间。对比传统加法融合仅能实现线性特征叠加、无维度扩展的局限, 星形操作在相同参数量下实现了维度大幅度增长, 为捕捉工地场景中目标与背景的复杂非线性关系提供了基础。

当星形操作跨层堆叠时, 特征的隐式维度会呈现指数级增长, 设第 l 层星形操作的输出为 O_l , 则其维度满足公式:

$$O_l = W_{l,1}^T O_{l-1} * W_{l,2}^T O_{l-1} \in R^{\left(\frac{d}{\sqrt{2}}\right)^{2^l}} \quad (2)$$

以输入通道 $d = 128$ 为例, 经过 5 层星形操作后, 特征的隐式维度约为 90^{32} (近似无穷维), 这一维度规模远超传统网络通过显式加宽通道(如将通道数从 128 翻倍至 256)所能达到的水平。这种“低显式维度 + 高隐式维度”的特性, 正是 StarNet 的关键优势——既能以轻量化的网络结构适配资源受限的硬件环境, 又能通过指数级扩展的隐式维度, 实现对复杂场景目标的精准特征刻画, 完美解决了 YOLOv11 主干网络计算量大与传统轻量化方案表达能力弱的双重痛点。

2.3. C3K2_SN 模块

为突破传统 C3k2 卷积模块在感受野局限、非线性特征交互不足及复杂场景细节捕获效率低的瓶颈, 本文提出一种改进型 C3k2-SN 卷积模块。该模块结合 StarNet 思想使用星形拓扑结构, 嵌入多组 7×7 深度可分离卷积(DWConv)以充分捕获广域空间上下文信息, 实现大核特征提取的轻量化。模型如下图 3 所示。

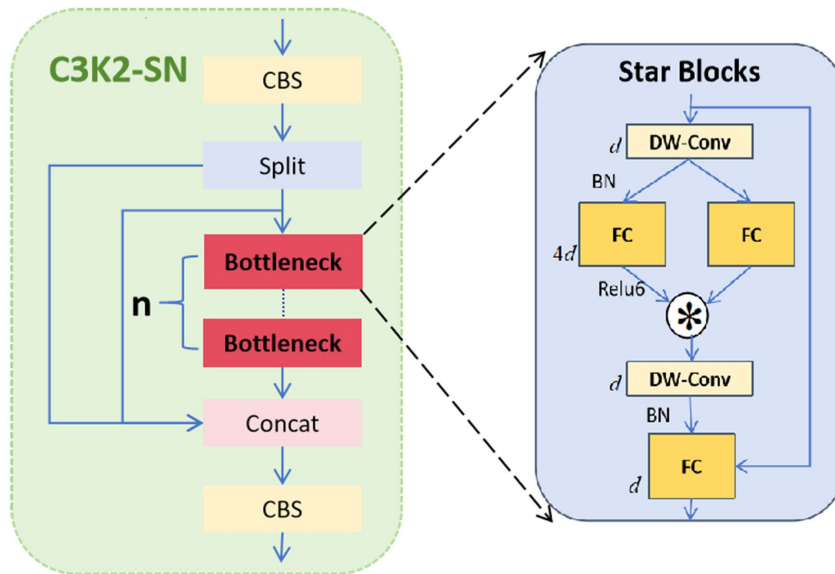


Figure 3. Structure of the C3K2_SN module
 图 3. C3K2_SN 模块

在特征融合机制上, C3k2-SN 模块构建双分支转换融合架构, 将输入特征流划分为两条并行路径。其中一条路径经激活函数(ReLU6)处理后, 与另一条路径进行逐元素乘法运算, 通过门控调节机制实现动态权重分配抑制冗余信息, 显著增强了特征表示的非线性能力。此外, 模块进一步引入残差连接与 DropPath 正则化策略, 残差结构通过跨层梯度传播, 有效缓解深层网络训练中的梯度消失问题; DropPath 通过随机丢弃部分分支路径, 抑制过拟合现象, 增强模型对噪声样本、遮挡场景的鲁棒性。

综上, C3k2-SN 模块星形拓扑结构创新有效解决了传统模块在大范围信息捕获与复杂特征交互中的不足, 结合残差连接与正则化策略的协同作用, 实现了模型综合性能的显著提升。

2.4. 轻量化共享卷积检测头(LSCD)

检测头通常负责将网络的特征映射转换为目标检测框及其相关的类别置信度和位置信息。为了解决 YOLOv11 解耦头参数量过大的问题, 本文设计轻量级共享卷积检测头(light weight shared convolutional detection head, LSCD), 模型使用 Group Norm (GN)把通道分为组, 并计算每一组之内的均值和方差, 以进行归一化处理可以提升检测头定位和分类的性能。GN 的计算与批量大小无关, 其精度也在各种批量大小下保持相对稳定。将 GN 用于卷积网络中, 归一化做特征融合, 获得更多的感受野, 可以提升检测头定位和分类的性能。LSCD 网络结构如图 4 所示, 首先进行一个 Detect 的输入, 经过一个 1×1 的 Conv_GN 的卷积之后, 进行一个合并操作, 通过使用两个 3×3 的共享卷积进行参数共享, 再转换到输出尺度的卷积网络。最后, 运用 Scale 层增加网络的表征能力, 允许网络学习对输入数据进行缩放和平移的操作, 从而提高模型的灵活性和拟合能力。通过使用共享卷积, 大幅减少了参数的数量, 当模型在资源受到限制的设备上使用时, 可以提升模型的检测速度, 使模型更轻便。在使用共享卷积的时, 为了解决模型在检测中每个检测头所检测的目标尺度不一致问题, LSCD 使用 Scale 层对特征进行缩放。在缩放的同时, 让检测头做到在参数量更少、计算量更少的情况下, 尽可能减少精度的损失。

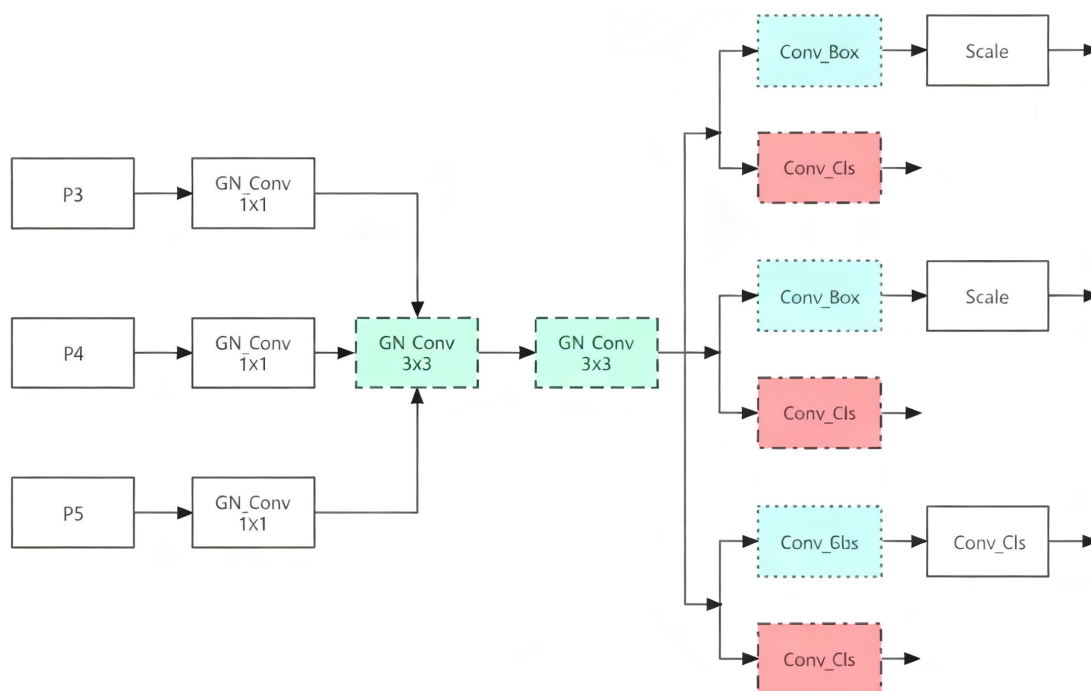


Figure 4. Structure of the LSCD module

图 4. LSCD 模块

在 LSCD 的设计中, 引入 Group Norm (GN) 技术提升检测性能。GN 通过将特征通道划分为若干独立分组, 对每组内的特征分别计算均值与方差并执行归一化处理, 这种方式既能有效抑制特征分布偏移, 强化检测头在定位与分类任务中的稳定性, 又具备与批量大小无关的显著优势——无论训练或推理时的批量大小如何变化, GN 的精度都能保持相对稳定, 避免了传统批量归一化(BN)在小批量场景下精度波动的问题。同时, 将 GN 融入卷积网络的特征融合过程, 还能帮助模型更高效地整合多尺度特征信息, 扩大有效感受野, 进一步强化检测头对目标细节(如安全帽边缘、反光条)与全局位置的捕捉能力, 为后续的检测任务提供更可靠的特征基础。LSCD 的具体网络结构如图 4 所示。

2.5. PIoU v2 回归损失优化函数

YOLO11 默认使用 CIoU 作为边界框损失函数, CIoU 在 IOU 仅关注重叠面积的基础上, 设计“锚框与目标框中心距离”与“宽高比差异”的惩罚项, 整体公式:

$$L_{CIoU} = 1 - IoU + \frac{d^2}{c^2} + \alpha v \quad (3)$$

$$v = \frac{4}{\pi^2} \times \left(\arctan\left(\frac{w_{gt}}{h_{gt}}\right) - \arctan\left(\frac{w}{h}\right) \right)^2 \quad (4)$$

(d 为中心距离, c 为能够同时完全包含锚框和目标框的最小矩形的对角线长度, α 为权重系数。 v 为宽高比差异项, 反映锚框与目标框形状上的匹配程度, 式中 w_{gt} 、 h_{gt} 为目标框的长度和宽度, w 、 h 为锚框的长度和宽度, $\arctan(\bullet)$ 函数将宽高比映射到角度)。

CIoU 的中心距离项分母为 c ——当锚框扩大时, c 同步增大, 导致中心距离项 $\frac{d^2}{c^2}$ 减小、损失降低, 模型会误将“锚框扩大”视为优化方向, 最终引发锚框扩大问题, 增加冗余计算且影响检测精度; 同样的, v 宽高惩罚项中反正切角度差平方的间接量化方式可能导致“宽高比差异小但实际尺寸偏差大”时, 损失无法有效约束, 最终影响回归精度; 并且 v 宽高惩罚项未考虑实际目标框的尺寸, 在包含小目标、模糊目标等低质量样本的安全帽场景下, 该项会导致该过度放大该类型目标惩罚, 从而降低模型的泛化能力。

针对锚框扩大以及低质量样本锚框问题, 我们引入 Power-IoU 损失函数, 其通过目标尺寸自适应惩罚因子和梯度调整函数, 引导锚框沿直线路径回归, 整体公式:

$$L_{PIoU} = 1 - (IoU - f(P)) = L_{IoU} + f(P) \quad (0 \leq L_{PIoU} \leq 2) \quad (5)$$

$$P = \frac{1}{4} \left(\frac{dw_1}{w_{gt}} + \frac{dw_2}{w_{gt}} + \frac{dh_1}{h_{gt}} + \frac{dh_2}{h_{gt}} \right) \quad (6)$$

$$f(x) = 1 - e^{-x^2} \quad (7)$$

$$f'(x) = 2xe^{-x^2} \quad (8)$$

其中, P 为目标尺寸自适应惩罚因子, 其中 dw_1 、 dw_2 、 dh_1 、 dh_2 表示预测框与目标框对应边的绝对距离, 用于量化预测框与目标框的边缘距离, 与 CIoU 中心距离项 $\frac{d^2}{c^2}$ 相比分母仅依赖目标框参数宽高 w_{gt} 、 h_{gt} , 有效的避免锚框扩大影响。 $f(x)$ 为梯度调整函数, 根据锚框的质量自适应调整损失梯度的幅值, 具体来

说低质量锚框(与目标框差异大, 对应惩罚因子 $P > 2$, 数值较大)会向模型引入有害梯度, 导致模型优化方向偏离正确路径。当 $x = P$ 带入 $f(x)$ 其导数 $f'(x)$ 会因指数项 e^{-x^2} 快速衰减而取值较小, 这种小梯度能弱化低质量锚框对损失的贡献。同理, 函数能放大中等质量($P \approx 1$)锚框的梯度, 加速锚框回归效率; 降低高质量锚框($P \approx 0$)的梯度, 稳定模型收敛过程, 实现锚框质量的自适应分配。

之后为了模型更精准地聚焦关键(中高质量)锚框, 进一步设计非单调注意力功能实现聚焦。具体实现如下:

$$L_{\text{PIoU v2}} = u(\lambda q) \cdot L_{\text{PIoU}} = 3\lambda q \cdot e^{-(\lambda q)^2} \cdot (1 - \text{IoU} + f(P)) \quad (9)$$

$$q = e^{-P} \in (0, 1] \quad (10)$$

$$u(\lambda q) = 3\lambda q e^{-(\lambda q)^2} \quad (11)$$

式中, 注意力函数 $u(\lambda q)$ 为一条单峰曲线, 锚框质量由 $q = e^{-P}$ 量化, 在不同场景下中等质量锚框 q 值范围不同, λ 作为超参数调整函数峰值位置, 使得注意力函数峰值聚焦中等质量锚框, 避免模型一刀切, 提升模型泛化能力。CIOU 和 PIoU 示意图如图 5 所示。

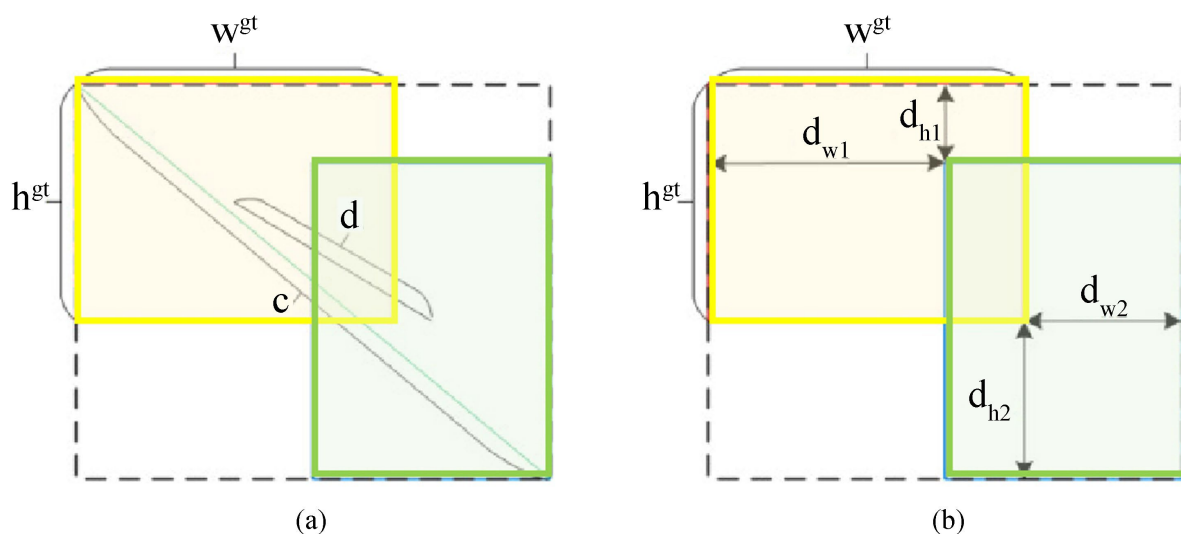


Figure 5. Schematic diagram of CIOU and PIoU

图 5. CIOU 和 PIoU 示意图

2.6. 改进的 LH-YOLO 算法

针对现有安全帽佩戴检测模型参数量大导致的计算资源消耗问题, 本文在 YOLO11n 为基初上, 提出改进的轻量化模型 LH-YOLO。首先, 使用轻量级 StarNet 网络对于 YOLO11n 中骨干网络进行替换, 通过星操作实现高维和非线性特征空间的映射, 在提升了网络的性能同时大幅降低计算复杂度; 其次, 设计了 C3k2_SN 卷积模块特征融合颈部, 增强了参数利用率和特征表达能力; 最后, 使用 LSCD 轻量化共享卷积头, 通过共享卷积与特征进行缩放技术, 优化检测头在参数量、计算量大的问题同时提升精度。并引入 PIoU v2 损失函数对边界框定位进行精细优化, 进一步压缩检测头参数并优化预测。改进后的模型结构图如图 6 所示。

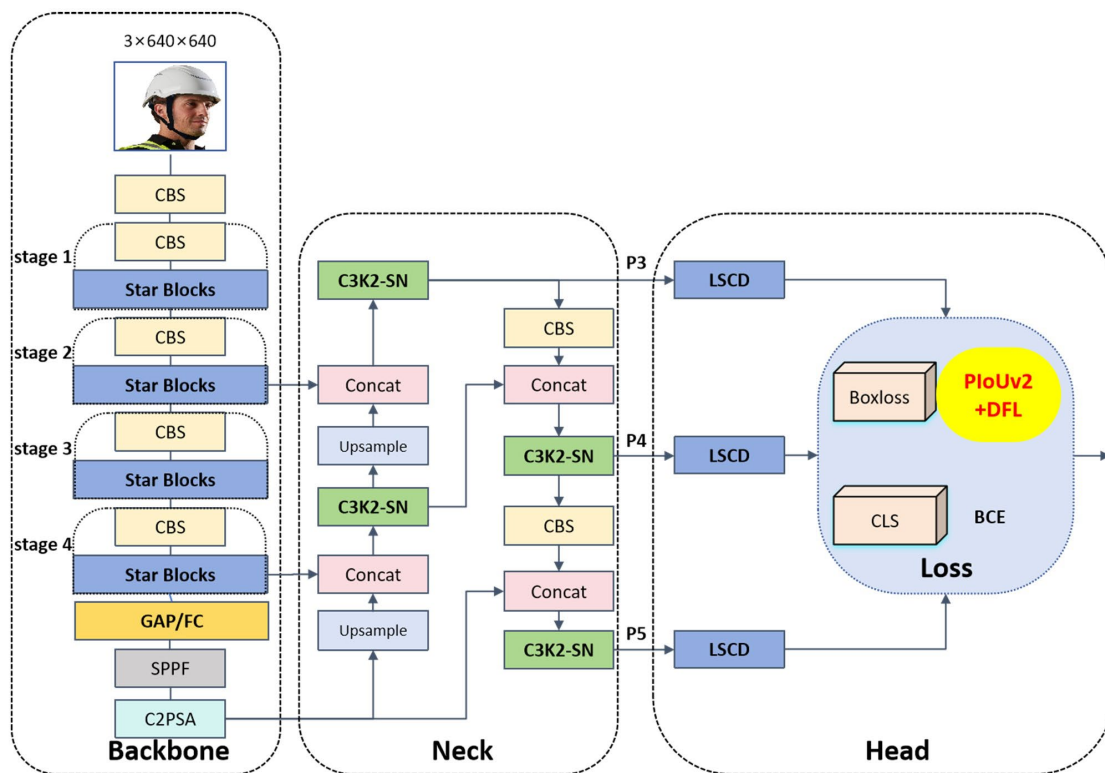


Figure 6. LH-YOLO network structure

图 6. LH-YOLO 网络结构

3. 实验与分析

3.1. 数据集与实验环境

3.1.1. 数据集



Figure 7. Typical images of the experimental dataset

图 7. 实验数据集典型图片

本文实验采用公开的安全帽数据集(SHWD)评估 LH-YOLO 模型的性能。该数据集为安全帽佩戴检测和头部目标识别设计, 提供两类(是否佩戴安全帽)目标定位和边界框坐标。数据集包含 7581 张图像, 其中 9044 个对象正确佩戴了安全帽(正样本), 而 111514 个对象未佩戴安全帽(负样本)。SHWD 数据集可通过链接(<https://github.com/njvisionpower/SafetyHelmet-Wearing-Dataset>)获取。本文将数据集转换为 YOLO 格式的数据集进行训练。同时将数据集按照 7:2:1 的比例随机划分为训练集、测试集、验证集, 训练集包含 5457 张图片, 测试集包含 1517 张图片, 验证集包含 607 张图片。训练过程中同步使用 YOLO11 中 Mosaic 增强、Mixup 增强对数据集进行数据增强处理, 提升模型鲁棒性。实验数据集典型图片如图 7 所示。

3.1.2. 实验环境

在本文实验在 Autodl 平台进行, 使用 Pytorch 框架, 深度学习的框架为 torch-2.5.1 + cu124, 语言版本为 Python3.10.15, GPU 为 NVIDIA GeForce RTX 3090, 24253 MiB, CPU 为 Intel (R) Xeon (R) Platinum 8362 CPU @ 2.80GHz, Cuda 版本为 12.4。实验环境的配置参数如表 1 所示。

Table 1. Configuration parameters of the experimental environment

表 1. 实验环境的配置参数

配置	参数
操作平台	Autodl
CPU	Intel (R) Xeon (R) Platinum 8362 CPU@2.80 GHz
GPU	NVIDIA GeForce RTX 3090, 24253MiB
CUDA	12.4
Pytorch	torch-2.5.1 + cu124
Python	3.10.15

实验主要训练参数如表二所示, 为加速网络训练速度, 初始权重设计为在 COCO 数据集上预训练权重(yolo11n.pt)。输入图像分辨率设置为 640×640 像素, 将批次大小设置(batch size)为 32, 训练轮数(epochs)设为 120 次, 耐心值(patience)设为 100 次, 采用随机梯度下降(SGD)优化器作为优化算法, 初始学习率(initial learning rate)设为 0.01, 动量(momentum)设为 0.937, 权重衰减系数(weight-decay)设为 0.005。实验主要训练参数如表 2 所示。

Table 2. Main training parameters of the experiment

表 2. 实验主要训练参数

配置	参数
Batch size	32
Epochs	120
Patience	100
优化器	SGD
Initial learning rate	0.01
Momentum	0.937
Weight-decay	0.005

3.2. 评价指标

性能指标用于衡量目标检测模型在准确性和效率上的核心表现。为了量化不同模型之间的性能差异，本文引入以下目标检测评价指标。

3.2.1. 混淆矩阵(Confusion Matrix)

混淆矩阵用于统计模型预测结果与样本实际标签的匹配情况：在评估中，一般使用 **T (True)**表示预测框与对应真实框的 IoU 达到或超过阈值的结果(即定位与分类均正确)，使用 **F (False)**表示预测框与对应真实框的 IoU 低于阈值或分类错误的无效检测，使用 **P (Positive)**表示模型判定属于当前目标类别的预测样本，使用 **N (Negative)**表示被模型排除在该类别之外的预测结果。则有：TP 表示真正类，FP 表示假正类，TN 表示真负类，FN 表示假负类。

3.2.2. 精度(Precision, P)

精度 P 衡量模型预测为正类的样本中实际为正类的比例，反映正类预测的“准确性”：

$$\text{Precision} = \frac{TP}{TP + FP} \quad (12)$$

3.2.3. 召回率(Recall, R)

召回率衡量实际为正类的样本中被模型正确预测为正类的比例，反映模型对正类样本的“查全率”：

$$\text{Recall} = \frac{TP}{TP + FN} \quad (13)$$

3.2.4. 平均精度(AP)

平均精度是综合反映精度与召回率权衡关系的指标，通过计算召回率在[0,1]区间内的精度平均值得到：

$$AP = \int_0^1 P(R) dR \quad (14)$$

式中， $P(R)$ 表示在召回率为 R 时的精度。

3.2.5. 均值平均精度(mAP)

均值平均精度是对所有类别平均精度的平均值，综合衡量模型在多类别目标检测中的整体精度：

$$mAP = \frac{\sum_{i=1}^N AP_i}{N} \quad (15)$$

式中， AP_i 表示第 i 类的平均精度， N 表示类别总数。

3.2.6. 计算量(GFLOPs)

浮点运算次数(FLOPs)是衡量算法或模型计算复杂度的标准指标，在神经网络中通常指网络单次前向传播所需的浮点运算数量，再除以 10^9 得到计算量 GFLOPs。其公式如下：

$$GFLOPs = \frac{c_{in} \times c_{out} \times k_w \times k_h \times f_h \times f_s}{10^9} \quad (16)$$

式中， $f_h \times f_s$ 表示的是特征图的大小。

3.2.7. 参数数量(Parameters)

神经网络中的参数指训练过程中可学习的权重与偏置，参数总数直接反映模型的容量、复杂度与内

存需求。其公式如下：

$$\text{Params} = c_{in} \times c_{out} \times k_w \times k_h \quad (17)$$

式中， c_{in} 表示的是输入通道数， c_{out} 表示的是输出通道数， $k_w \times k_h$ 表示的是卷积核的大小。

3.3. 消融实验

为衡量不同模块对安全帽检测模型性能的影响，验证网络中各模块的优化效果，本文基于所提方法开展消融实验。实验结果如表 3 所示：以 YOLOv11n 为基准模型，考察四个模块(StarNet 主干网络、C3K2-SN 模块、LSCD 检测头模块、PIoU v2 损失函数)的作用，表格中每一行代表不同模块的组合实验，勾选标记表示该模块被选中。消融实验结果如下表 3 所示。

Table 3. Ablation experiment results

表 3. 消融实验结果

G	S	C	L	P	P/%	R/%	mAP@0.5%	Params/M	GFLOPs/G
1	-	-	-	-	91.1	86.5	92.0	2.58	6.4
2	√	-	-	-	91.2	84.1	91.0	1.94	5
3	-	√	-	-	92.7	84.7	92.3	2.47	6.4
4	-	-	√	-	92.1	86.1	92.4	2.42	5.6
5	-	-	-	√	90.5	87.6	92.8	2.58	6.4
6	√	√	√	√	92.3	88.2	93.4	1.73	4.3

从实验结果可得出以下结论。

单独使用任一模块均可提升模型性能，但模块组合使用时效果更优。

最后一行(使用所有四个模块，即本文提出的 LH-YOLO 模型)的性能指标最高：与 YOLOv11n 模型相比，精度、召回率、mAP@0.5、分别提升 1.2%、1.7%、1.4%，表明这些模块协同作用时检测性能最佳；与原模型参数量降低了 32.9%、十亿次的浮点计算量(GFLOPs)下降了 32.8%，表明轻量化指标上表现出色；集成所有模块后，mAP@0.5 达到 93.4%，证明模型在不同交并比阈值下均具有稳定性能。

未使用任何所提模块时，模型性能指标最低，表明每个模块均对提升模型检测能力具有积极作用。

在所有实验设置中，召回率(R)存在波动，但在所有模块组合使用时达到最高值，表明模块间存在互补效应。

实验结果表明，LH-YOLO 模型在头部与安全帽佩戴类别的检测中均具有较高的准确性与可靠性。

3.4. 对比实验

为了进一步验证所提出的模型的检测性能，本文选择了与 6 个优秀的目标检测模型进行比较，包括 Faster-RCNN、YOLOv3、YOLOv5m、YOLOv8n、YOLOv10n、YOLOv11n。LH-YOLO 与目前主流的目标检测模型相比，在检测精度、模型参数量、模型计算量方面都优于其他的模型，对比实验结果如表 4 所示。

Table 4. Ablation experiment results

表 4. 对比实验结果

模型	mAP@0.5%	Params/M	GFLOPs/G
Yolov11n	92.0	2.58	6.4

续表

Faster-RCNN	71.3	15.8	28.3
YOLOv3	86.11	61.5	65
YOLOv5m	88.04	21.1	51.4
YOLOv8n	91.4	3.01	8.2
YOLOv10n	90.6	2.27	6.5
LH-YOLO	93.4	1.73	4.3

3.5. 实验结果及展示

为了更直观地展示本研究的优化效果,采用 YOLOv11n 模型与改进型模型 LH-YOLO 对 SHWD 数据集进行检测与评估,结果如图所示(每行第一列为原图,第二列为 YOLOv11n 模型检测结果,第三列为 LH-YOLO 模型检测结果)。

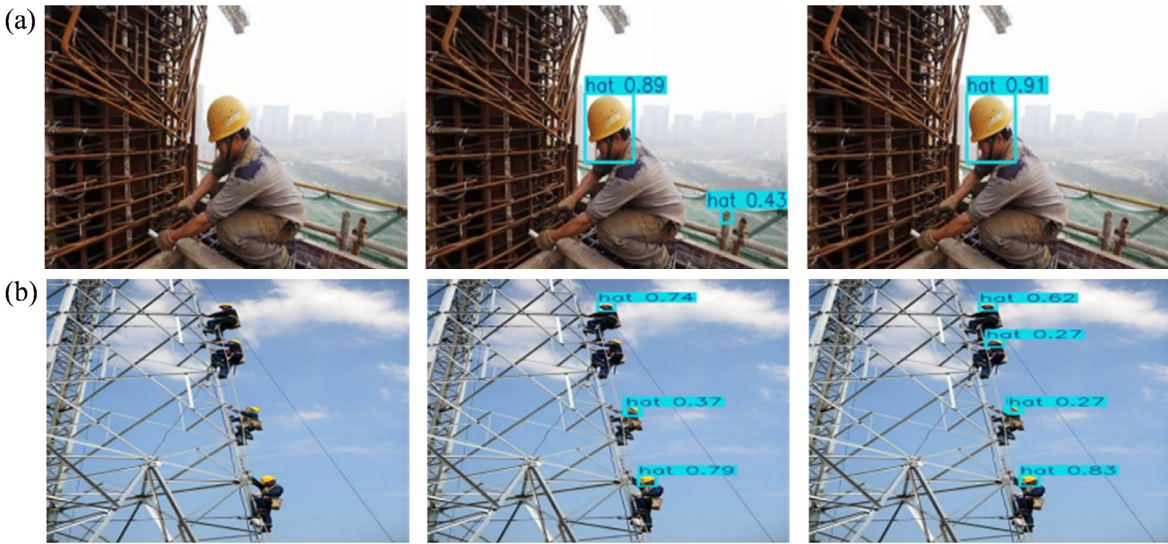


Figure 8. Comparison of detection results
图 8. 检测结果对比图

对于检测结果对比图 8(a),对于建筑外部场景下脚手架搭建工人,YOLOv11n 模型出现误检行为,对将后方栏杆小目标断点是为安全帽目标,可能原因是视野后方环境复杂,而优化后的模型高置信度识别了目标对象同时避免误检。对于检测结果对比图 8(b),对于户外高处作业场景下电塔攀爬作业的工人,YOLOv11n 模型出现漏检行为,对由高到低第二位佩戴安全帽小目标的工人呈现未识别,可能原因是目标与与上方目标位置接近且目标与背景相似,而优化后的模型识别了目标对象同时避免漏检。用过上述对比,证实了模型的优化效果的有效性,提高了模型对各类目标的整体检测正确率,本章优化后的模型对安全帽现场图的检测效果更加优秀。

4. 总结

针对安全帽佩戴检测任务中存在的模型参数量计算量大计算效率低的问题,本研究提出 LH-YOLO 安全帽检测模型。通过系列对比实验与消融实验,得出以下结论。

提出的 LH-YOLO 模型在 YOLOv11n 基础上, 通过融合轻量级 StarNet 网络、设计 C3k2-SN 卷积模块与轻量化共享卷积检测头(LSCD)、采用 PIoU v2 损失函数, 优化了特征提取与融合过程。该方案提升参数利用效率, 增强了模型的尺度感知与空间感知能力, 损失函数通过设计惩罚因子避免锚框扩大并增加单超参数控制注意力权重, 聚焦中等质量锚框。这些改进措施共同提升了模型性能, 降低模型大小同时显著增强了模型在复杂场景下对各类目标的检测能力。

相较于基准 YOLOv11n 模型, LH-YOLO 在平均精度增加 1.4%达到 93.4%的情况下, 参数量降低了 32.8%、十亿次的浮点计算量(GFLOPs)下降了 34.8%, 表明改进模型显著提升了安全帽检测任务的性能。

但该模型在背景复杂且存在噪声干扰时仍可能出现误检与漏检。后续工作将考虑引入图像分割算法作为预处理步骤, 以降低复杂背景与噪声的影响, 从而进一步提升模型检测能力。

参考文献

- [1] 李卉, 张云波, 祁神军. 建筑施工坍塌事故致因分析及对策[J]. 建筑经济, 2018, 39(8): 53-57.
- [2] 叶贵, 李学征, 杨丽萍, 等. 建筑工人不安全行为量化分类研究[J]. 安全与环境学报, 2021, 21(6): 2617-2627.
- [3] 高腾, 张先武, 李柏. 深度学习在安全帽佩戴检测中的应用研究综述[J]. 计算机工程与应用, 2023, 59(6): 13-29.
- [4] Dong, C., Pang, C., Li, Z., Zeng, X. and Hu, X. (2022) PG-YOLO: A Novel Lightweight Object Detection Method for Edge Devices in Industrial Internet of Things. *IEEE Access*, **10**, 123736-123745. <https://doi.org/10.1109/access.2022.3223997>
- [5] Wang, L., Zhang, X. and Yang, H. (2023) Safety Helmet Wearing Detection Model Based on Improved YOLO-M. *IEEE Access*, **11**, 26247-26257. <https://doi.org/10.1109/access.2023.3257183>
- [6] Han, D., Ying, C., Tian, Z., Dong, Y., Chen, L., Wu, X., et al. (2024) YOLOv8s-SNC: An Improved Safety-Helmet-Wearing Detection Algorithm Based on YOLOv8. *Buildings*, **14**, Article No. 3883. <https://doi.org/10.3390/buildings14123883>
- [7] Fan, Z., Wu, Y., Liu, W., Chen, M. and Qiu, Z. (2024) Lg-yolov8: A Lightweight Safety Helmet Detection Algorithm Combined with Feature Enhancement. *Applied Sciences*, **14**, Article No. 10141. <https://doi.org/10.3390/app142210141>
- [8] Lin, B. (2024) Yolov8n-asf-dh: An Enhanced Safety Helmet Detection Method. *IEEE Access*, **12**, 126313-126328. <https://doi.org/10.1109/access.2024.3435453>
- [9] Du, Q., Zhang, S. and Yang, S. (2024) BLP-YOLOv10: Efficient Safety Helmet Detection for Low-Light Mining. *Journal of Real-Time Image Processing*, **22**, Article No. 10. <https://doi.org/10.1007/s11554-024-01587-6>
- [10] Hidayatullah, P., Syakrani, N., Sholahuddin, M.R., Gelar, T. and Tubagus, R. (2025) YOLOv8 to YOLO11: A Comprehensive Architecture In-Depth Comparative Review.
- [11] Khanam, R. and Hussain, M. (2024) Yolov11: An Overview of the Key Architectural Enhancements.