

## YOLOv8-LPE——面向工业场景的防毒面具佩戴检测方法

朱 毅, 汪邦荣, 汪 军\*, 徐晓峰

(安徽工程大学 计算机与信息学院, 芜湖 241000)

**摘要:** 防毒面具作为一种个体防护装置, 它的正确佩戴对于工业安全、公共卫生以及军事领域的个人防护具有重要作用。鉴于近年来工业污染问题的凸显以及公共健康事件的频发, 确保防毒面具已经被正确佩戴成为了一个紧迫的研究议题。为了更高效、准确地检测工作人员的防毒面具佩戴情况, 提出了基于 YOLOv8 的改进模型 YOLOv8-LPE。将 LSKA 融合到 SPPF 中以提升不同特征层间的语义融合效果, 然后在原来的 3 个检测层的基础上增加了 P2 层辅助检测小目标, 最后增加了 EMA 模块, 增强模型对小目标的特征提取能力。YOLOv8-LPE 在没有明显增加参数量和计算量的情况下, 目标检测平均精度  $mAP@0.5 : 0.95$  相较于基线模型提高了 3.5%, 达到了 60.5%, 高于目前目标检测领域的常见算法, 具有实际应用意义。

**关键词:** 防毒面具佩戴检测; YOLO; 小目标; 目标检测

**中图分类号:** TP391.4 **文献标志码:** A **文章编号:** 1673-4602(2025)06-0149-08

## YOLOv8-LPE—a gas mask wearing detection method for industrial scenarios

ZHU Yi, WANG Bangrong, WANG Jun\*, XU Xiaofeng

(School of Computer and Information, Anhui Polytechnic University, Wuhu 241000, China)

**Abstract:** Gas masks are a type of individual protective device, and the correct wearing of gas masks plays an important role in industrial safety, public health and personal protection in military field. Due to the pervasive issue of industrial pollution in recent years and the frequent occurrence of public health incidents, to ensure that gas masks can be correctly worn has become a pressing research topic. To more efficiently and accurately detect the wearing condition of gas masks, an improved model based on YOLOv8, named YOLOv8-LPE is proposed. LSKA is integrated into SPPF to enhance the semantic fusion effect between different feature layers, then P2 layer is added to assist in detecting small targets on the basis of the original three detection layers, and finally EMA module is added to enhance the model's feature extraction ability for small targets. Without significantly increasing the parameters and Flops, YOLOv8-LPE achieves a 3.5% increase in  $mAP@0.5 : 0.95$  compared to the baseline model, reaching 60.5%, and surpasses the current common algorithms in the field of object detection, which has practical application significance.

**Key words:** gas mask wearing detection; YOLO; small targets; object detection

收稿日期: 2024-04-28

基金项目: 安徽省自然科学基金(2108085QF268)

作者简介: 朱 毅(2000—), 男, 安徽淮南人。硕士, 研究方向为深度学习与目标检测。E-mail: 1849767621@qq.com。

\* 通信作者: 汪 军(1975—), 男, 安徽桐城人。硕士, 教授, 主要从事机器视觉与控制方面的研究。E-mail: 546275490@qq.com。

近年来,政府和企业对员工在工业环境中健康的关注不断增加。尤其在工业生产领域,工人持续接触生产性矿物性粉尘会显著提高患尘肺病的可能性<sup>[1]</sup>,因此,呼吸防护设备,尤其是防毒面具的正确佩戴至关重要。防毒面具若佩戴不当可能引发各种呼吸系统疾病<sup>[2]</sup>。对此,需要提出可靠而精确的防毒面具佩戴检测方法,以确保工人健康。

为提高目标检测精度,一阶段和二阶段的目标检测模型都有较快的发展。经典的一阶段目标检测模型包括 YOLO(You Only Look Once)系列<sup>[3-5]</sup>、SSD(Single Shot MultiBox Detector)<sup>[6]</sup>、RetinaNet 等。具有代表性的二阶段模型,包括 R-CNN、Fast R-CNN、Faster R-CNN<sup>[7]</sup>和 Cascade R-CNN<sup>[8]</sup>。

目前,面向工业领域的检测模型多使用 YOLO 系列模型。关于防毒面具的佩戴情况检测工作较少,只有 WANG 等<sup>[9]</sup>采用的基于 Faster R-CNN 方法的相关研究,但是实时性较差。因此,本文还探索了类似应用的研究,包括面部检测、口罩检测和头盔检测。YU 等<sup>[10]</sup>开发了一种基于 YOLOv5 的实时人脸检测器 YOLO-FaceV2,创建了 RFE 模块来增强小面孔的感知领域,并使用 NWD(Normalized Wasserstein Distance)Loss 来解释 IoU(Intersection over Union)对小物体位置偏差的敏感性。对于面部遮挡的问题,提出了 SEAM 注意模块,并引入了拒绝损失来缓解这个问题。魏明军等<sup>[11]</sup>提出了 Mask-YOLO 模型,针对复杂场景下的口罩佩戴检测情景,在原模型中引入通道注意力机制以突出重要特征。LI 等<sup>[12]</sup>提出了基于 YOLOv5 的安全帽检测模型,在训练过程中提出了一种分级正样本选择机制,提高了模型的拟合能力。KIM 等<sup>[13]</sup>提出了一种基于 YOLOv8 的改进模型用于检测无人机视角下的小目标,采用在 Neck 层中添加一个丰富小目标语义信息的特征层 SODH(Small Object Detection Head),以增强对小目标的特征提取。

受上述方法的启发,本文提出了基于 YOLOv8 的防毒面具佩戴稳定检测模型。为了满足工业需求,采集了一个防毒面具数据集。在对数据集进行分析后,得出真实场景有复杂背景和众多小目标的结论。本文将大核分离卷积注意力(Large Separable Kernel Attention,LSKA)模块<sup>[14]</sup>添加到空间金字塔池化融合(Spatial Pyramid Pooling Fusion,SPPF)模块中,提升不同特征层间的语义融合;然后将主干的 P2 层的输出集成到颈部,增加了一个小目标检测层,辅助进行目标检测任务;最后在头部添加了高效多尺度注意力(Efficient Multi-Scale Attention,EMA)<sup>[15]</sup>模块,在不进行通道降维的情况下学习有效的待检测目标语义信息。

## 1 防毒面具佩戴检测数据集

防毒面具是一种用于保护佩戴者呼吸系统的个人防护装备,主要用于过滤空气中的有害物质,如有毒气体、化学污染物、生物病原体以及颗粒物,在军事、工业、医疗等领域具有广泛的应用,根据工作原理的不同,可分为过滤式和隔离式两种类型。工业生产中常用的是过滤式防毒面具而不是隔离式防毒面具。本文的防毒面具检测数据集特别关注于工业环境中常用的 3M 过滤防毒面具,如图 1 所示。

本文中使用的数据集包括来自真实的工业场景和模拟场景的数据。该数据集包括 7992 张图像,其中包括 6058 个训练样本和 1934 个测试样本。根据头部对摄像头的方向将其分为 3 个不同的类别:正对摄像头、侧对摄像头和背对摄像头。考虑到实际场景下使用防毒面具的情况,将这 3 个类别进一步细分为 7 个标签。表 1 和图 2 提供了标签名称和含义。例如,“side\_head\_wear\_wrong”这样的标签指的是戴口罩。在工业环境中,只佩戴口罩的工人进入生产车间较常见,不符合安全防护要求,也应列为被检测的目标。

## 2 构建防毒面具佩戴检测模型

### 2.1 YOLOv8 模型

YOLOv8 的网络结构主要包括主干网络(Backbone)、颈部(Neck)和头部(Head)3 个部分。YOLOv8 的 Backbone 部分主要用来进行特征提取,采用了 Darknet-53 结构,包含的主要模块是 Conv、C2f(CS-PLayer\_2Conv)和 SPPF。Neck 部分主要用来进行特征融合,采用了 PAN-FPN(Path Aggregation Networks-Feature Pyramid Networks)结构,引入了自上向下和自下向上的路径,使得底层信息更容易传递到顶层。Head 部分相比 YOLOv5 改动较大,直接将耦合头改为类似 YOLOx 的解耦头结构(Decoupled-

Head), 仅有分类分支和回归分支。



图 1 3M 防毒面具

表 1 数据集标签

| 序号 | 标签                    | 标签含义         |
|----|-----------------------|--------------|
| 1  | front_head_wear_mask  | 正确佩戴防毒面具正对镜头 |
| 2  | side_head_wear_mask   | 正确佩戴防毒面具侧对镜头 |
| 3  | front_head_no_wear    | 未佩戴防毒面具正对镜头  |
| 4  | side_head_no_wear     | 未佩戴防毒面具侧对镜头  |
| 5  | front_head_wear_wrong | 错误佩戴防毒面具正对镜头 |
| 6  | side_head_wear_wrong  | 错误佩戴防毒面具侧对镜头 |
| 7  | back_head             | 背对镜头         |



图 2 数据集可视化

## 2.2 YOLOv8-LPE 模型

针对防毒面具佩戴数据集中目标所占像素小(小于  $32 \times 32$ )、背景复杂和类别不均衡问题,原有的 YOLOv8n 模型存在着漏检、误检、精度低等问题。因此,本文基于 YOLOv8n 模型做出以下改进,提出 YOLOv8-LPE(YOLOv8-LSKA P2 EMA)。

1) 将 LSKA 添加到 SPPF 模块中,以提升不同特征层间的语义融合效果。

2) 在 YOLOv8 原有的 P3、P4、P5 检测层外,增加了丰富小目标语义信息的 P2 检测层,通过 4 个检测层处理检测工作。

3) 在 Neck 层中使用 EMA,由于 Neck 层的输出包含目标的特征信息,EMA 模块可以增强模型对目标特征信息的学习能力。改进后的模型结构如图 3 所示。

### 2.2.1 改进的 SPPF 模块

将 LSKA 融合进 SPPF 如图 4 所示,SPPF 模块将输入降维传入串行的 3 个  $5 \times 5$  的最大池化层,然后将每层的输出进行拼接,再送入 LSKA 模块,大型可分离卷积注意力通过更大的感受野进行特征提取,最后通过普通卷积进行特征融合,调整模型 Backbone 部分最终输出的特征向量大小。

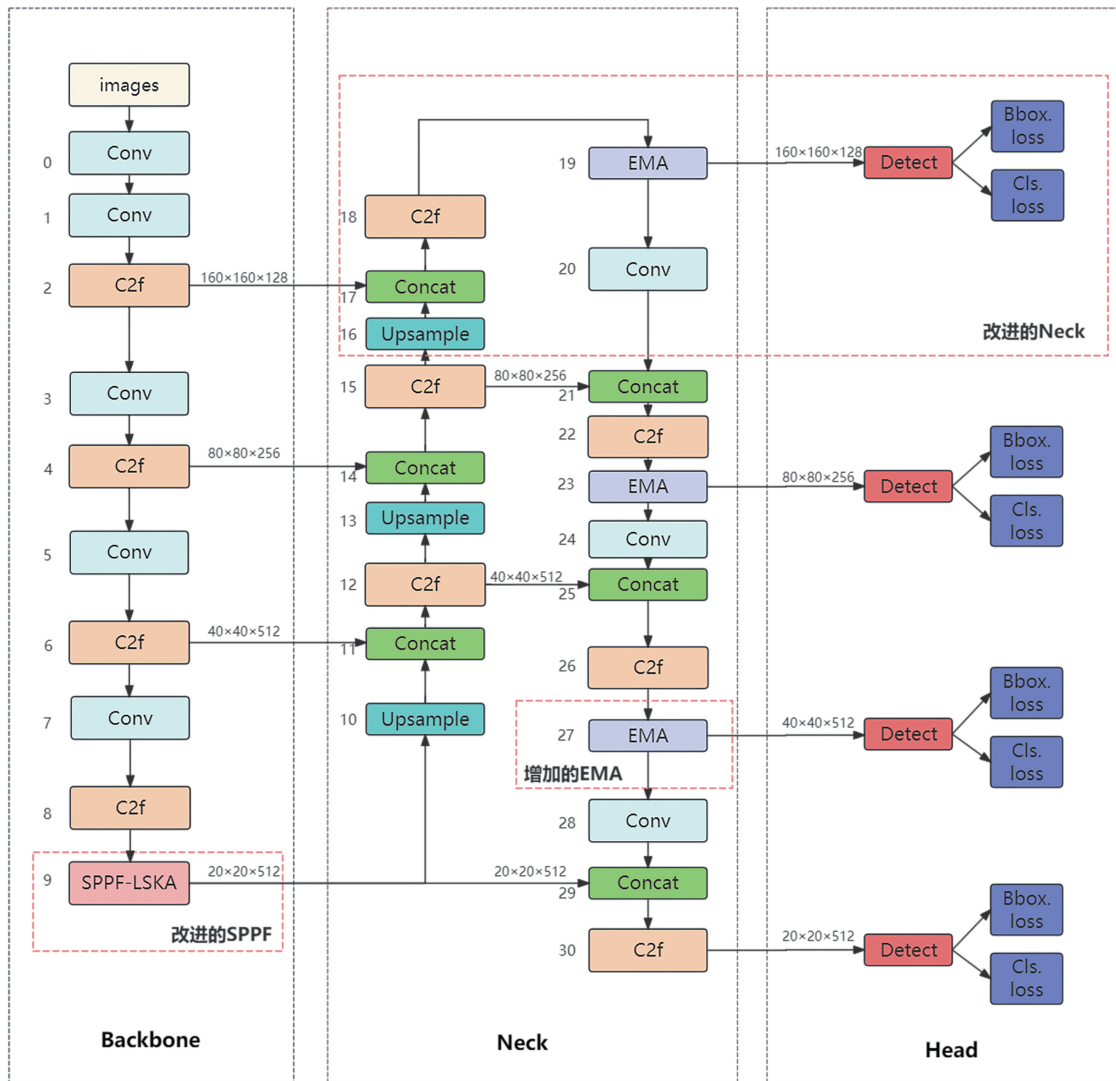


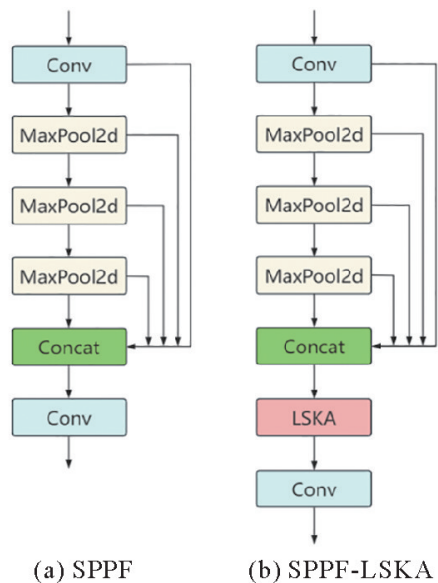
图 3 YOLOv8-LPE 模型结构

### 2.2.2 改进的 Neck 结构

原有的 YOLOv8 模型只有 P3、P4、P5 3 种不同尺度的检测层,但是模型经过了下采样后,小目标的特征已经不够明显,因此在 Neck 中增加一个 P2 层。P2 层通过与 Backbone 中  $160 \times 160$  尺度的特征向量进行残差连接,增强模型对小目标的检测能力。

### 2.2.3 新增 EMA 模块

EMA 模块如图 5 所示,其中“X Avg Pool”和“Y Avg Pool”分别代表一维水平和垂直的全局池化操作。EMA 将输入按跨通道维度方向划分为不同分支进行处理。在  $1 \times 1$  的分支中进行“X Avg Pool”和“Y Avg Pool”,在  $3 \times 3$  的分支中通过  $3 \times 3$  内核以捕获多尺度特征表示。2 个分支的输出特征通过 Sigmoid 函数和归一化操作跨维度交互模块合并,以捕捉像素级的成对关系。EMA 模块增强了模型对目标特征信息的学习能力,提



(a) SPPF (b) SPPF-LSKA

图 4 SPPF 和 SPPF-LSKA 模块示意

高了对小目标的检测精度。

### 3 实验设置和评价指标

#### 3.1 实验设置

在这项工作中,实验环境是 Ubuntu20.04 和 Python3.8。深度学习框架是 Pytorch 1.12.0,所有模型都在 Nvidia GTX 3090 上进行了训练和测试。在训练阶段,本文使用随机梯度下降优化器,权重衰减为  $10^{-3}$ ,动量为 0.98。所有模型的训练和测试输入大小都设置为  $1280 \times 1280$ 。由于 GPU 内存的限制,检测器训练的批处理大小为 16,训练周期为 300。

#### 3.2 评价指标

本文选取如下指标评估模型的检测性能:精确率(Precision, P)、召回率(Recall, R)、 $F_1$ 得分( $F_1$ -Score)、平均精度(Average Precision, AP)、平均精度均值(Mean Average Precision, mAP)、参数量(Params)、计算量(GFLOPs)。

$$P = \frac{P_T}{P_T + P_F} \quad (1)$$

$$R = \frac{P_T}{P_T + N_F} \quad (2)$$

$$P_A = \int_0^1 P(R) dR \quad (3)$$

式中: $P$ 为精确率,指被正确分配的正样本数占总分配的正样本数的比例; $R$ 为召回率,指被正确分配的正样本数与总正样本数的比例; $P_T$ 为真正例,即正类预测为正类; $P_F$ 为假正例,即负类预测为正类; $N_F$ 为假负例,即正类预测为负类; $N_T$ 为真负例,即负类预测为负类; $P_A$ 为平均精度,是某一类别的  $P$ - $R$  曲线的面积,用于计算不同类别的平均精度。

式(4)中  $F_1$  得分是精确率和召回率的调和平均数,同时兼顾了模型的精确率和召回率,其最大值是 1,最小值是 0,值越大意味着模型越好。

$$F_1 = \frac{2 \times P \times R}{P + R} \quad (4)$$

在本文中,使用  $mAP@0.5$  评价置信度阈值为 0.5 时各类别的平均精度, $mAP@0.5 : 0.95$  为置信度阈值从 0.5 到 0.95 按 0.05 步长加权平均后的平均精度( $P_{mA}$ ),计算如式(5):

$$P_{mA} = \frac{\sum_{i=1}^k P_{Ai}}{k} \quad (5)$$

式中: $k$ 为防毒面具佩戴的类别数量; $P_{Ai}$ 为第  $i$  个类别的  $P_A$ 。

## 4 实验结果和分析

#### 4.1 消融实验

为了验证每个改进模块对基线模型的优化效果,在防毒面具佩戴数据集上进行消融实验,实验结果如表 2 所示。将 LSKA 融合进 SPPF 后, $mAP@0.5$  和  $mAP@0.5 : 0.95$  分别提高了 4.5% 和 1.3%;在此基础上,在 Neck 中添加了 P2 层后,相较于基线模型, $mAP@0.5$  和  $mAP@0.5 : 0.95$  分别提高了 4.1%

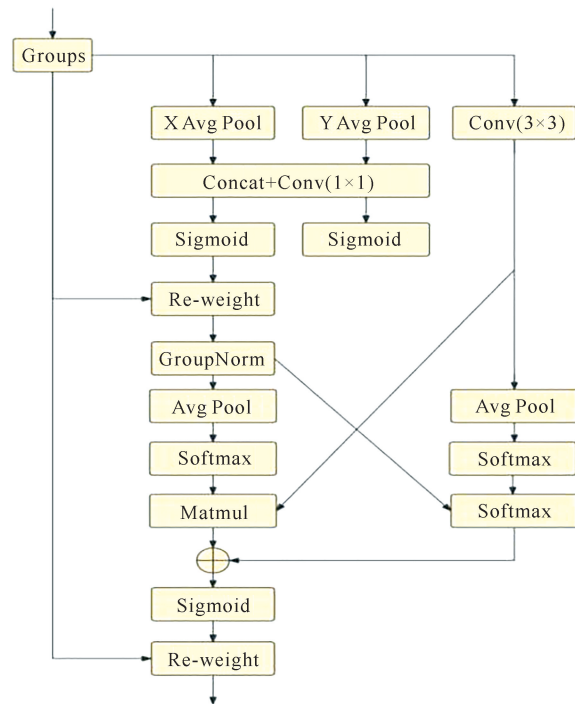


图5 EMA 模块示意

和 2.3%；最后增加 EMA 模块后，YOLOv8-LPE 比基线模型 mAP@0.5 和 mAP@0.5 : 0.95 分别提高了 5.2% 和 3.5%。综上所述，YOLOv8-LPE 在检测防毒面具佩戴情况时相较于 YOLOv8n 基线模型表现出更高的平均精度，证明了本文的改进模型具有有效性和可行性。

表 2 消融实验结果

| 模型               | P    | R    | $F_1$ | mAP@0.5/% | mAP@0.5 : 0.95/% |
|------------------|------|------|-------|-----------|------------------|
| YOLOv8n          | 77.2 | 76.0 | 76.5  | 80.4      | 57.0             |
| YOLOv8n+LSKA     | 78.0 | 77.6 | 77.8  | 84.9      | 58.3             |
| YOLOv8n+P2       | 72.5 | 78.4 | 75.3  | 82.3      | 57.9             |
| YOLOv8n+EMA      | 80.1 | 75.7 | 77.8  | 83.6      | 57.5             |
| YOLOv8n+LSKA+P2  | 74.6 | 85.0 | 76.2  | 84.5      | 59.3             |
| YOLOv8n+LSKA+EMA | 78.0 | 75.3 | 76.6  | 84.3      | 58.5             |
| YOLOv8n+P2+EMA   | 79.6 | 75.5 | 77.5  | 83.1      | 58.0             |
| YOLOv8-LPE       | 78.9 | 79.2 | 79.0  | 85.6      | 60.5             |

## 4.2 对比实验

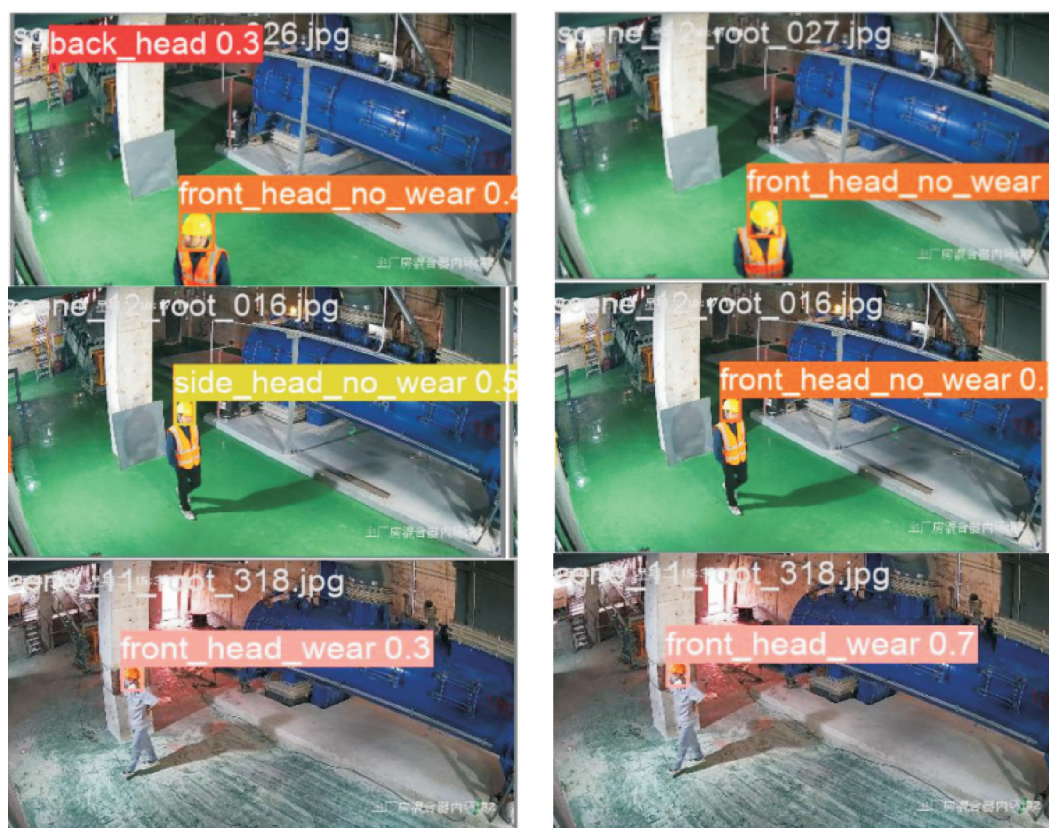
为了进一步验证 YOLOv8-LPE 模型的检测效果，将其与目前目标检测领域中应用广泛且具有代表性的 Faster R-CNN、RetinaNet、YOLOv5 系列、YOLOv7tiny 和 YOLOv8 系列进行对比，对比实验结果见表 3。由表 3 可以看出，本文提出的 YOLOv8-LPE 相较于上述算法，不仅参数量优化至  $3.20 \times 10^6$ 、计算量仅为 12.5，而且 mAP@0.5 : 0.95 也是达到了 60.5%。帧率 (Frames Per Second, FPS) 虽然相较 YOLOv8n 原模型降低了，但仍有 170，满足实际使用需求，实验结果表明，YOLOv8-LPE 更适用于防毒面具佩戴检测。

表 3 对比实验结果

| 模型           | mAP@0.5/% | mAP@0.5 : 0.95/% | GFLOPs | FPS | Params/ $10^6$ |
|--------------|-----------|------------------|--------|-----|----------------|
| Faster R-CNN | 83.7      | 55.3             | 322.34 | 20  | 41.15          |
| RetinaNet    | 83.7      | 54.0             | 331.00 | 21  | 36.23          |
| YOLOv5m      | 81.5      | 56.9             | 47.90  | 110 | 20.89          |
| YOLOv5l      | 85.1      | 58.3             | 107.70 | 70  | 46.16          |
| YOLOv7tiny   | 83.5      | 55.3             | 13.20  | 526 | 6.03           |
| YOLOv8n      | 80.4      | 57.0             | 8.9    | 357 | 3.16           |
| YOLOv8s      | 85.0      | 59.0             | 28.8   | 154 | 11.2           |
| YOLOv8-LPE   | 85.6      | 60.5             | 12.5   | 170 | 3.20           |

## 4.3 实验结果可视化

图 6 展示了 YOLOv8 基准模型与改进后的 YOLOv8 模型的检测效果对比。从图 6(a) 的基准模型检测结果可以看出，原模型存在明显的误检、错检以及检测精度不足的问题，表明其定位与分类能力存在显著不足。相比之下，改进后的 YOLOv8 模型不仅能够准确识别目标类别，还显著提升了检测精度，有效解决了基准模型中的误检和错检问题。



(a) YOLOv8检测效果

(b) YOLOv8-LPE检测效果

图 6 检测效果

## 5 结论

本文将 LSKA 模块融合到 SPPF 中,以提升不同特征层间的语义融合;然后将主干的 P2 层的输出集成到颈部,增加了一个小目标检测层,4 个检测层辅助进行目标检测任务;最后在头部添加了 EMA 模块,在不进行通道降维的情况下学习有效的待检测目标语义信息,YOLOv8-LPE 在没有明显增加参数量和计算量的情况下,mAP@0.5 : 0.95 相较于基线模型增加了 3.5%,达到了 60.5%。

### 参考文献 (References):

- [1] BARLING J,FRONE M R. The psychology of workplace safety[M]. Washington: American Psychological Association,2004.
- [2] SOLHI M,SAKI M,ALIMOHAMMADI I,et al. Effect of health education based on BASNEF pattern on use of personal protective respiratory equipment in Ahvaz carbon block factory workers,2009[J]. Iran Occupational Health,2012,9(2):50-58.
- [3] REDMON J,DIVVALA S,GIRSHICK R,et al. You only look once:Unified,real-time object detection[C]//2016 IEEE conference on Computer Vision and Pattern Recognition. Las Vegas,2016:779-788.
- [4] REDMON J,FARHADI A. YOLOv3: An incremental improvement[J]. arXiv preprints,2018. DOI:10.48550/arXiv.1804.02767.
- [5] WANG C Y,BOCHKOVSKIY A,LIAO H Y M. YOLOv7:Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[C]//2023 IEEE/CVF conference on Computer Vision and Pattern Recognition. Vancouver,2023:7464-7475.
- [6] LIU W,ANGUELOV D,ERHAN D,et al. SSD: Single Shot Multibox Detector[C]//Computer Vision-ECCV 2016. Amsterdam, 2016:21-37.
- [7] 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.
- [8] CAI Z W,VASCONCELOS N. Cascade R-CNN:High quality object detection and instance segmentation[J]. IEEE Transactions on

- Pattern Analysis and Machine Intelligence, 2019, 43(5):1483-1498.
- [9] WANG B R, WANG J, XU X F, et al. Gas mask wearing detection based on Faster R-CNN[J]. Journal of Ambient Intelligence and Smart Environments, 2024, 16(1):57-71.
- [10] 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.
- [11] 魏明军, 周太宇, 纪占林, 等. 基于 Mask-YOLO 的复杂场景口罩佩戴检测[J]. 应用科学学报, 2022, 40(1):93-104.  
WEI Mingjun, ZHOU Taiyu, JI Zhanlin, et al. Mask wearing detection in complex scenes based on Mask-YOLO[J]. Journal of Applied Sciences, 2022, 40(1):93-104.
- [12] LI Z S, XIE W Q, ZHANG L Z, et al. Toward efficient safety helmet detection based on YOLOv5 with hierarchical positive sample selection and box density filtering[J]. IEEE Transactions on Instrumentation and Measurement, 2022, 71:1-14.
- [13] KIM J H, KIM N, WON C S. High-speed drone detection based on YOLO-v8[C]//ICASSP 2023—2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Rhodes Island, 2023:1-2.
- [14] LAU K W, PO L M, REHMAN Y A U. Large separable kernel attention: Rethinking the large kernel attention design in CNN[J]. Expert Systems with Applications, 2024, 236:121352.
- [15] OUYANG D, HE S, ZHANG G Z, et al. Efficient multi-scale attention module with cross-spatial learning[C]//ICASSP 2023—2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Rhodes, 2023:1-5.

(责任编辑 张晓靖; 英文校审 程文华)

---

#### (上接第 148 页)

- [10] DENG W, LIU H, XU J, et al. An improved quantum-inspired differential evolution algorithm for deep belief network[J]. IEEE Transactions on Instrumentation and Measurement, 2020, 69(10):7319-7327.
- [11] ZOUFAL C, SUTTER D, WOERNER S. Error bounds for variational quantum time evolution[J]. Physical Review Applied, 2023, 20(4):044059.
- [12] HAUG T, KIM M S. Natural parametrized quantum circuit[J]. Physical Review A, 2022, 106(5):052611.
- [13] 方正宜, 彭玉旭. 求解旅行商问题的改进型量子蚁群算法[J]. 计算机工程与应用, 2016, 52(22):59-63.  
WAN Zhengyi, PENG Yuxu. Improved quantum ant colony algorithm for traveling salesman problem[J]. Computer Engineering and Applications, 2016, 52(22):59-63.
- [14] NOMURA Y, YOSHIOKA N, NPRI F. Purifying deep Boltzmann machines for thermal quantum states[J]. Physical Review Letters, 2021, 127(6):060601.
- [15] NOMURA Y. Boltzmann machines and quantum many-body problems[J]. Journal of Physics: Condensed Matter, 2023, 36(7):073001.
- [16] ZEGUENDRY A, JRIR Z, UAFAROU M. Quantum machine learning: A review and case studies[J]. Entropy, 2023, 25(2):287.

(责任编辑 张晓靖; 英文校审 程文华)