

廖佳逸, 冯奥, 蒋银萤, 等. 基于改进 YOLOv8n 的口罩佩戴检测算法[J]. 智能计算机与应用, 2025, 15(10): 145-149. DOI: 10.20169/j. issn. 2095-2163. 251022

# 基于改进 YOLOv8n 的口罩佩戴检测算法

廖佳逸, 冯奥, 蒋银萤, 骆进

(贵州大学 电气工程学院, 贵阳 550025)

**摘要:** 针对人工检测口罩佩戴时, 漏检率高和检测速度慢的问题, 提出一种改进的 YOLOv8n 口罩佩戴检测模型。首先, 在 YOLOv8n 网络模型中通过引入 CA 注意力机制, 以提高计算效率。其次, 在 YOLOv8n 网络模型中添加 CBAM 注意力机制, 增强特征图的通道之间的关系。最后, 引入 WIoU 损失函数指导网络训练, 在预测框之间赋予权重系数, 提高模型的推理能力和检测精度。实验结果表明, 本文改进的 YOLOv8n 口罩检测模型相比于原模型 mAP 值高达 91.6%, 提升 2.1%。检测速度高达 85 FPS, 提升 7.5%, 能够满足实际场景应用的精度与速度需求。

**关键词:** 口罩检测; YOLOv8n; 注意力机制; C2f; WIoU

中图分类号: TP391 文献标志码: A 文章编号: 2095-2163(2025)10-0145-05

## Mask wearing detection algorithm based on improved YOLOv8n

LIAO Jiayi, FENG Ao, JIANG Yinying, LUO Jin

(College of Electrical Engineering, Guizhou University, Guiyang 550025, China)

**Abstract:** Aiming at the problems of high leakage rate and slow detection speed when manually detecting mask wearing, an improved YOLOv8n mask wearing detection model is proposed. Firstly, the CA attention mechanism is introduced into the YOLOv8n network model by introducing the CA attention mechanism to improve the computational efficiency. Secondly, the CBAM attention mechanism is added to the YOLOv8n network model to enhance the relationship between the channels of the feature map. Finally, the WIoU loss function is introduced to guide the network training, and the weight coefficients are assigned between the prediction frames to improve the model's inference ability and detection accuracy. The experimental results show that the improved YOLOv8n mask detection model in this paper has a high mAP value of 91.6% compared to the original model, an improvement of 2.1%. The detection speed is as high as 85 FPS, an improvement of 7.5%, which can meet the accuracy and speed requirements of real scene applications.

**Key words:** mouthpiece detection; YOLOv8n; attention mechanism; C2f; WIoU

## 0 引言

佩戴口罩是医疗、建筑和化工等行业的从业人员免受病毒、粉尘和化学物质侵害的重要手段<sup>[1]</sup>。目前, 口罩佩戴检查多为人工检查, 检查效率低, 难以及时发现错误佩戴口罩和未佩戴口罩的行为。随着深度学习的普及和发展, 基于神经网络的目标检测算法在行人识别、口罩检测、智能交通等方面得到广泛的应用<sup>[2-3]</sup>。经典的目标检测算法主要分为 2 类。一类是两阶段(Two Stage)算法, 如 R-CNN<sup>[4]</sup>、Fast R-CNN<sup>[5]</sup>、Mask R-CNN<sup>[6]</sup>; 另一类是单阶段

(One Stage) 检测算法, 如 SSD<sup>[7]</sup>、YOLO<sup>[8]</sup>、Retina-Net<sup>[9]</sup>系列等。李泽琛等学者<sup>[10]</sup>在原始 Faster R-CNN 框架中引入 Res2Net 分组残差结构, 获取更细粒度的特征表征, 验证了模型的可行性和有效性。田勇等学者<sup>[11]</sup>提出基于 Mask R-CNN 算法的口罩佩戴识别检测方法, 实现了对佩戴口罩的目标检测与实例分割。李雨阳等学者<sup>[12]</sup>在 SSD 模型中引入特征融合网络及协调注意力机制, 提高模型对口罩佩戴检测平均精度。刘玉国等学者<sup>[13]</sup>通过替换 Faster R-CNN 中的主干网络, 提高模型检测精度。牛作东等学者<sup>[14]</sup>在 RetinaFace 增加人脸口罩佩戴

基金项目: 贵州大学“SRT 计划”项目(贵大 SRT 字[2022]054 号); 大学生创新创业项目(贵大(省)创字 2023(021), 2023(022))。

作者简介: 廖佳逸(2001—), 女, 本科生, 主要研究方向: 目标检测, 计算机视觉。Email: liaojiayi2024@163.com; 冯奥(2001—), 男, 本科生, 主要研究方向: 基于深度学习的目标检测; 蒋银萤(2002—), 女, 本科生, 主要研究方向: 图像处理; 骆进(2001—), 男, 本科生, 主要研究方向: 计算机视觉。

收稿日期: 2024-01-06

检测任务,改进损失函数,增强网络对图像特征的表征能力。叶子勋等学者<sup>[15]</sup>使用改进的 MobileNetv3 替换 YOLOv4 的主干特征提取网络,提高网络检测精度。孙永豪<sup>[16]</sup>针对室内人脸佩戴口罩检测场景,在 YOLOx 的基础上加入通道注意力机制,改善网络的有效特征利用率,提高人脸目标检测的准确率。

上述研究都实现了口罩佩戴检测,但应用场景不够广泛,在复杂场景下的口罩检测漏检率高,检测精度和速度有待提升。为此,本文提出一种基于改进 YOLOv8n 的口罩佩戴检测算法,并通过与其他目标检测算法对比,确保其性能指标能够满足应用需求,实现复杂环境下的实时检测。

## 1 YOLOv8n 模型

本研究采用缩放系数较小的 YOLOv8n 模型,在图像分辨率为  $640 \times 640$  的 COCO 数据集上的训练结果表明,相较于其它模型,YOLOv8n 模型检测帧率为 79 帧每秒,检测速度快,模型参数量为 5.64M,参数量少,其模型结构如图 1 所示。

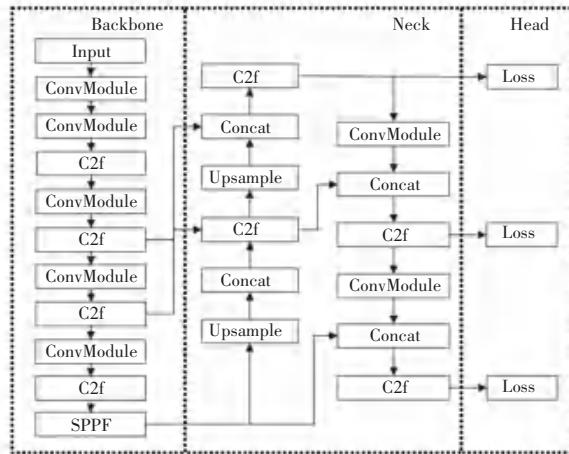


图 1 YOLOv8n 模型结构

Fig. 1 YOLOv8n model structure

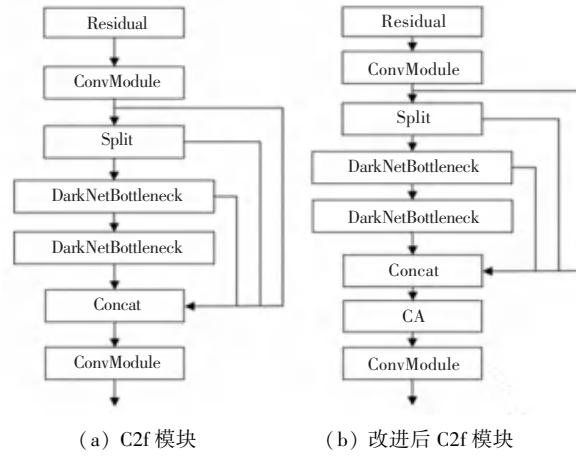
YOLOv8n 模型由 Backbone、Neck 和 Head 三个部分组成。输入图像首先经过 Backbone 部分进行特征提取,然后将特征图传递到 Neck 部分,进一步增强特征并实现多尺度特征融合,最后将融合结果输入到 Head 部分,由 Head 生成预测的目标框位置、类别和置信度。

## 2 YOLOv8n 模型改进

### 2.1 C2f 模块改进

为提高模型的计算效率,减少不必要的计算负担,在 C2f 模块中引入 CA<sup>[17]</sup> 注意力机制,CA 注意

力机制是一种通道注意力机制,能增强特征图中的重要通道并抑制不相关的通道。通过在 C2f 中引入 CA 注意力机制,使模型可以动态地学习每个通道特征的重要性,使其更关注与目标检测任务相关的通道,忽略与任务无关的通道特征,通过减少不重要通道特征的计算,降低计算复杂性,从而提高模型的关键特征提取能力,提高实时性能。改进前后的 C2f 模块如图 2 所示。其中,图 2(a) 为原始 C2f 模块,图 2(b) 为改进后的 C2f 模块。



(a) C2f 模块 (b) 改进后 C2f 模块

Fig. 2 C2f module before and after improvements

### 2.2 CBAM 注意力机制

CBAM<sup>[18]</sup>是一种用于增强卷积神经网络性能的通道与空间注意力机制,可提高网络的表示能力,允许网络更好地关注重要区域的特征,从而提高网络的特征提取能力。CBAM 注意力机制主要由 2 个关键模块组成,分别为通道注意力模块和空间注意力模块,其结构如图 3 所示。

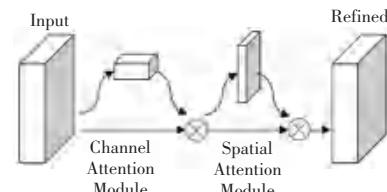


图 3 CBAM 结构

Fig. 3 CBAM structure

通道注意力模块主要用于增强特征图的通道之间的关系。通过学习每个通道的权重,以自适应地调整不同通道的贡献度。空间注意力主要用于增强特征图中的不同空间位置之间的关系。通过学习每个空间位置的权重,以自适应地关注不同位置的信息,公式如下:

$$F' = M_c(F) \otimes F \quad (1)$$

$$F'' = M_s(F') \otimes F' \quad (2)$$

其中,  $F$  表示输入的特征图; “ $\otimes$ ” 表示对应元素乘积, 中间采用广播机制进行维度变换和匹配;  $M_c$  表示通道注意力操作;  $M_s$  表示空间注意力操作;  $F'$  表示输出通道注意力特征图;  $F''$  表示输出通道注意力特征图。输入特征图经过通道注意力平均池化, 得到通道注意力特征图  $F'$ , 再经过空间注意力机制、平均池化和最大池化处理, 将特征图  $F$  与  $F'$  拼接起来、输出  $F''$ , 突出其中的最有用信息, 使模型更有效地关注不同位置的信息, 从而提高其空间关系感知性能。

### 2.3 WIoU 损失函数

YOLOv8 的原模型的回归损失函数采用 CIoU<sup>[19]</sup> 损失函数, CIoU 是 IoU 的一种改进版本。传统的 IoU 只考虑预测框和真实框的重叠部分, 如图 4 所示, 没有考虑两者之间的区域, 导致在评估结果时可能存在偏差。

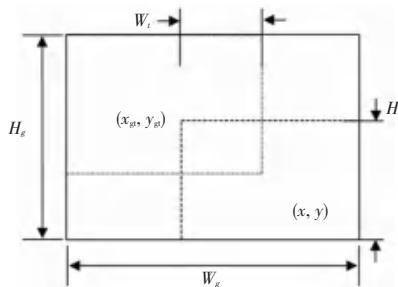


图 4 IoU 图解

Fig. 4 IoU illustration

CIoU 考虑重叠面积、中心点距离、纵横比, 提升回归精确度, 相比于 CIoU, WIoU<sup>[20]</sup> 不仅考虑到预测框与真实框的重叠面积、中心点距离等因素, 还通过引入加权系数衡量 2 个框之间的关系, 可以有效解决计算偏差问题。计算公式如下:

$$\text{WIoU} = \frac{\sum_{i=1}^n w_i \text{IoU}(b_i, g_i)}{\sum_{i=1}^n w_i} \quad (3)$$

其中,  $n$  表示预测框或真值框的总数;  $i$  表示当前预测框或真值框索引值;  $w_i$  表示预测框与第  $i$  真值框的权重值;  $\text{IoU}(b_i, g_i)$  表示预测框与真实标注框之间的 IoU 值;  $b_i$  表示目标预测框的坐标;  $g_i$  表示目标真值框的坐标。2 个框之间重合的面积越大, 权重值越高, 通过引入框之间的区域和权重系数, WIoU 可以更准确地评估目标检测结果, 避免传统 IoU 存在计算偏差的问题。

## 3 实验结果与分析

### 3.1 数据集与实验平台

本研究使用的数据集由 2 000 张原始图像组成, 其中佩戴口罩的图像数量为 1 240 张, 未佩戴口罩的图像量为 760 张, 且包含医院、工厂、街道、学校等多种场景。首先, 对原始图像进行数据增强操作, 包括对图像进行亮度调整、对比度调整、水平翻转、添加高斯噪声等, 得到数据增强后的图像共 8 000 张, 将其作为本实验的数据集。其次, 通过 Labelimg 软件对佩戴口罩和未佩戴口罩的目标分别进行框选, 并标注为“mask”与“face”两类标签。最后, 将图像以 6 : 3 : 1 比例划分为训练集、验证集和测试集。实验平台和相关配置见表 1。

表 1 实验平台配置

Table 1 Configuration of the experimental platform

| 名称   | 配置           |
|------|--------------|
| 操作系统 | Windows 11   |
| GPU  | RTX3060      |
| 模型框架 | PyTorch1.7.1 |
| 编程语言 | Python3.8    |
| CUDA | 11.6         |

### 3.2 评价指标

目标检测模型的性能评价涵盖多个关键指标, 一般采用精确率 ( $P$ )、召回率 ( $R$ )、平均精度均值 (mAP)、帧率 (FPS) 等指标对模型进行评价。本实验选取的模型检测性能评价指标如下。

(1) 精确率 ( $P$ )。是指模型正确预测为正类别的样本数  $T_p$  与所有被模型预测为正类别的样本数  $T_p + F_p$  之比。精确率主要衡量模型在正类别预测上的准确性。定义公式如下:

$$P = \frac{T_p}{T_p + F_p} \quad (4)$$

(2) 召回率 ( $R$ )。是指模型正确预测为正类别的样本数  $T_p$  与所有实际正类别的样本数  $T_p + F_N$  之比。召回率衡量模型检测到的正类别的数量。定义公式如下:

$$R = \frac{T_p}{T_p + F_N} \quad (5)$$

(3) mAP。是所有类别的 AP 的平均值, 用于评估整体性能。AP 是目标检测精度指标, 其大小等于  $P - R$  曲线下与坐标轴围成的面积。AP 的计算公式如下:

$$AP = \int_0^1 p(r) dr \quad (6)$$

其中,  $r$  表示召回率。

(4) FPS。用于描述模型在图像或视频中检测目标的速度。FPS 表示模型每秒可以处理和检测的图像或视频帧数。FPS 越高, 模型处理速度越快。

### 3.3 改进模型实验数据分析

$P-R$  曲线是评估二分类模型性能的一种常用指标。是以精确率和召回率为 2 个主要指标, 其坐标轴的横轴是召回率, 纵轴是精确率, 曲线上的每个点对应于一个不同的阈值。随着阈值的改变, 模型的召回率和精确率会发生变化, 从而在  $P-R$  曲线上形成一条曲线。曲线所围成的面积为 AP 值, mAP 值由 AP 值求平均所得, 本研究中由“mask”的 AP 与“face”的 AP 值求取平均值得到, 结果曲线如图 5 所示。由图 5 可以看出, 改进后模型的 mAP 值大于原模型的 mAP 值。

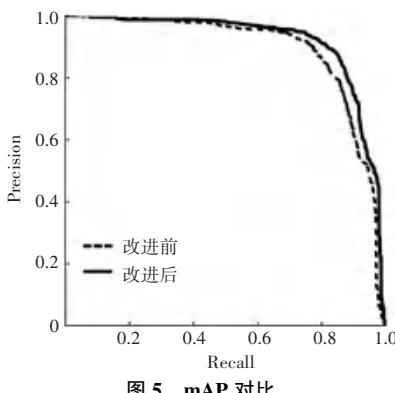


Fig. 5 Comparison of mAP

改进前后的 YOLOv8n 网络模型和原网络模型对口罩的检测效果如图 6 所示。图 6(a)、(b) 中, 左侧为原模型检测结果, 右侧为改进后模型检测结果。图 6(a) 中, 原模型检测口罩的置信度为 0.86, 改进后的模型检测口罩置信度为 0.91; 图 6(b) 中, 原模型重复检测一个口罩目标, 改进后的模型无复检情况出现, 且置信度相较于原模型有所提高。



图 6 检测效果对比

Fig. 6 Comparison of detection effect

### 3.4 不同位置添加 CBAM 注意力机制效果对比

为探究 CBAM 模块接入位置对网络性能的影响, 设计以下实验。在特征提取网络不同位置接入 CBAM 模块, 在相同训练环境下训练模型, 实验结果见表 2。表 2 中, 第 1~3 列为 CBAM 模块不同摆放位置, 第 4 列为 mAP 值, 第 5 列为检测帧率(FPS)。由表 2 中数据可知, 将 CBAM 模块放到前 2 个 C2f 模块之后效果最好, mAP 与 FPS 均高于其他 4 种情况。

表 2 不同位置添加 CBAM 注意力机制实验结果对比

Table 2 Comparison of experimental results of adding CBAM attention mechanisms at different locations

| 第 1 个 C2f 后 | 第 2 个 C2f 后 | 第 3 个 C2f 后 | mAP   | FPS |
|-------------|-------------|-------------|-------|-----|
| ✓           |             |             | 0.899 | 79  |
|             | ✓           |             | 0.907 | 74  |
|             |             | ✓           | 0.898 | 78  |
| ✓           | ✓           |             | 0.916 | 85  |
| ✓           | ✓           | ✓           | 0.895 | 77  |

### 3.5 模型对比

为验证本实验模型的有效性, 在相同参数下, 与其他常见目标检测模型进行对比验证, 结果见表 3。表 3 中, 第 1 列为模型名称, 第 2 列为模型参数数量大小(Params), 第 3 列为输入图像分辨率(size), 第 4 列为 mAP 值, 第 5 列为检测帧率(FPS)。从表 3 中第 4 列、5 列可知, 本模型的 mAP 为 91.6%, FPS 为 85。综上所述, 本模型对口罩的检测速度与精度均优于其余 4 种模型。

表 3 不同模型口罩检测效果实验结果对比

Table 3 Comparison of experimental results of different models

| Model        | Params/M | Size    | mAP   | FPS |
|--------------|----------|---------|-------|-----|
| 本模型          | 5.94     | 640×640 | 0.916 | 85  |
| YOLOv8n      | 5.96     | 640×640 | 0.895 | 79  |
| YOLOv7-tiny  | 11.70    | 640×640 | 0.897 | 44  |
| YOLOv5s      | 13.70    | 640×640 | 0.854 | 23  |
| Faster R-CNN | 108.00   | 640×640 | 0.525 | 15  |

## 4 结束语

本文针对现有口罩佩戴检测算法检测精度低、速度慢的问题, 通过改进网络模块、引入注意力机制、优化损失函数改进 YOLOv8n 模型。提出了一种高精度的口罩佩戴检测算法。首先, 在 C2f 模块中引入 CA 注意力机制, 提高模型的关键特征提取性能。然后, 在主干网络中引进 CBAM 注意力机制, 改善网络的空间关系感知能力。进一步, 设计了

WIoU 损失函数, 约束网络在 IoU 计算中的权重分配, 提高网络的检测精度。最后, 不同模型口罩检测效果实验表明, 改进后的 YOLOv8n 模型精度提高, mAP 达到 91.6%, 检测速度提升, 达到 85 FPS, 能够满足实际场景应用需求。

## 参考文献

- [1] 朱华. 防尘口罩结构/材料及其加工技术探讨[J]. 中国安全生  
产科学技术, 2013, 9(4): 67-71.
- [2] 许德刚, 王露, 李凡. 深度学习的典型目标检测算法研究综述  
[J]. 计算机工程与应用, 2021, 57(8): 10-25.
- [3] 赵永强, 饶元, 董世鹏等. 深度学习目标检测方法综述[J]. 中  
国图象图形学报, 2020, 25(4): 629-654.
- [4] GIRSHICK R, DONAHUE J, DARRELL T, et al. Region-based  
convolutional networks for accurate object detection and  
segmentation [J]. IEEE Transactions on Pattern Analysis and  
Machine Intelligence, 2016, 38(1): 142-158.
- [5] WANG Xiaolong, ABHINAV S, ABHINAV G. A-Fast-RCNN:  
Hard positive generation via adversary for object detection [C]//  
Proceedings of 2017 IEEE Conference on Computer Vision and  
Pattern Recognition (CVPR). Piscataway, NJ: IEEE, 2017:  
3039-3048.
- [6] HE Kaiming, GKIOXARI G, DOLLAR P, et al. Mask R-CNN  
[C]//Proceedings of the IEEE International Conference on  
Computer Vision. Piscataway, NJ: IEEE, 2017: 2980-2988.
- [7] 卢云聪. 基于 SSD 算法的口罩佩戴检测模型[J]. 科技与创新,  
2022, 29(18): 101-103.
- [8] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look  
once: Unified, real-time object detection [C]//Proceedings of the  
IEEE Conference on Computer Vision and Pattern Recognition.  
Piscataway, NJ: IEEE, 2016: 779-778.
- [9] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal Loss for dense  
object detection [J]. IEEE Transactions on Pattern Analysis and  
Machine Intelligence, 2020, 42(2): 318-327.
- [10] 李泽琛, 李恒超, 胡文帅, 等. 多尺度注意力学习的 Faster R-  
CNN 口罩人脸检测模型 [J]. 西南交通大学学报, 2021, 56  
(5): 1002-1010.
- [11] 田勇, 林开司. 基于深度学习的佩戴口罩检测方法研究 [J].  
集宁师范学院学报, 2023, 45(5): 98-102.
- [12] 李雨阳, 沈记全, 翟海霞, 等. 基于改进 SSD 的口罩佩戴检测  
算法 [J]. 计算机工程, 2022, 48(8): 173-179.
- [13] 刘玉国, 张晶. 基于改进的 Faster R-CNN 的行人口罩检测  
[J]. 现代计算机, 2021, 27(26): 73-76.
- [14] 牛作东, 覃涛, 李捍东, 等. 改进 RetinaFace 的自然场景口罩  
佩戴检测算法 [J]. 计算机工程与应用, 2020, 56(12): 1-7.
- [15] 叶子勋, 张红英, 何昱均. 融合注意力机制的轻量级戴口罩人  
脸识别算法 [J]. 计算机工程与应用, 2023, 59(3): 166-174.
- [16] 孙永豪. 基于改进 YOLOx 的口罩佩戴检测系统设计与实现  
[D]. 合肥: 安徽建筑大学, 2023.
- [17] HOU Qibin, ZHOU Dequan, FENG Jiashi. Coordinate attention  
for efficient mobile network design [C]//Proceedings of 2021  
IEEE/CVF Conference on Computer Vision and Pattern  
Recognition (CVPR). Piscataway, NJ: IEEE, 2021: 13708 -  
13717.
- [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). Cham: Springer, 2018: 3-19.
- [19] ZHENG Zhaohui, WANG Ping, REN Dongwei, 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] TONG Zanlia, CHEN Yuhang, XU Zewei, et al. Wise-IoU:  
Bounding box regression loss with dynamic focusing mechanism  
[J]. arXiv preprint arXiv:2301.10051, 2023.