

DOI:10.19651/j.cnki.emt.2417273

改进 YOLOv8 的密集人群口罩检测算法^{*}

伍锡如 梁诗意

(桂林电子科技大学电子工程与自动化学院 桂林 541004)

摘要: 针对公共场合密集人群场景下由于人群遮挡导致的信息缺失及检测目标较小、分辨率低导致人脸佩戴口罩检测算法检测困难的问题,提出了一种改进 YOLOv8 的密集人群口罩检测算法。采用 GD 机制代替 YOLOv8 中 FPN 结构解决跨层信息传输中信息丢失的问题,GD 使用一个统一的模块收集和融合所有层级的信息,实现网络跨层信息的无损传输,增强了网络特征提取能力。在 GD 机制中插入 ODconv 模块对 GD 传输的信息沿 4 个维度进行学习,提高模型低分辨目标和小目标的检测精度。此外,引入了一种轻量化检测头 SCSBD,对占比过高的 YOLOv8 检测头进行轻量化处理,平衡模型大小。实验结果表明,改进后的网络在精确率、召回率和平均精度上分别提升了 13.6%、1.5% 和 6.9%,对人脸口罩的检测精度达到了 81.1%,模型权重文件仅为 25 MB,模型大小介于 YOLOv8s 和 Gold-YOLO-S 之间,达到了大小和精度的平衡。

关键词: 密集人群;口罩检测;YOLOv8;GD 机制;ODConv;SCSBD;特征提取;轻量化

中图分类号: TP391;TN919.8 **文献标识码:** A **国家标准学科分类代码:** 520.20

Improved YOLOv8 for mask detection in dense crowds

Wu Xiru Liang Shiyi

(College of Electronic Engineering and Automation, Guilin University of Electronic Technology, Guilin 541004, China)

Abstract: To address the challenges in mask detection for faces in dense crowd scenarios, particularly due to information loss from crowd occlusion, small detection targets, and low resolution, improved YOLOv8 algorithm for dense crowd mask detection is proposed. This approach replaces the FPN structure in YOLOv8 with a GD mechanism to solve the issue of missing cross-layer information transmission. The GD mechanism uses a unified module to collect and integrate information from all layers, enabling lossless cross-layer information transmission and enhancing the network's feature extraction capabilities. The ODconv module is inserted into the GD mechanism to learn the information transmitted by GD along four dimensions, improving the model's detection accuracy for low-resolution images and small targets. Additionally, a SCSBD is introduced to lighten the YOLOv8 detection head, which occupies a significant proportion, thereby balancing the model size. Experimental results show that the improved network has increased precision, recall, and mean average precision by 13.6%, 1.5% and 6.9%, respectively, with an 81.1% accuracy in mask detection on faces. The model's weight file is only 25 MB, and its size is between YOLOv8s and Gold-YOLO-S, achieving a balance between size and accuracy.

Keywords: dense crowd; mask detection; YOLOv8; GD mechanism; ODConv; SCSBD; feature extraction; lightweight

0 引言

2019年,新型冠状病毒(COVID-19)^[1-2]开始在全球蔓延,新冠主要通过唾液、飞沫传播且传染性极强。口罩作为有效且简单的预防措施是减缓新冠传播的最佳解决方案。

在学校、商场、公交等人流较多的地点会对口罩佩戴情况进行检查。目前很多公众场合还是以人工的方式检测是否佩戴口罩,不仅效率低、存在漏检还浪费人力资源,因此,对口罩佩戴检测方法进行研究具有重要意义。

目标检测的众多算法可大致分为单阶段检测算法与双

收稿日期:2024-11-04

* 基金项目:国家自然科学基金(62263005)、广西自然科学基金(2020GXNSFDA238029)、广西高校人工智能与信息处理重点实验室开放基金重点项目(2022GXZDSY004)、桂林电子科技大学研究生教育创新计划项目(2024YCX119,2024YCX131)资助

阶段检测算法^[3],单阶段检测算法在单一前向传播中完成目标检测任务,无需生成候选区域,典型的算法有 YOLO^[4]、SSD^[5]、RetinaNet^[6]等。双阶段检测算法通常有两个阶段,第一阶段通过区域提议网络(如:RPN^[7])生成候选目标区域再进入第二阶段对这些划分区域进行分类和精细化定位,典型的代表算法包括:Faster R-CNN^[8]、Mask R-CNN^[9]、Cascade R-CNN^[10]等。

为了解决口罩检测过程中的问题,高民等^[11]提出了一种基于 FasterNet 的 FLM-YOLOv8 检测算法,平衡检测精度与速度,但主干网络的特征提取能力较差;魏明军等^[12]以 YOLOv3 为基准改进特征金字塔结构,将底层特征图的传递到中层和高层,但信息在传输的过程中依旧存在丢失;李小波等^[13]提出了基于融合注意力机制的 YOLOv5 口罩检测算法,但未做轻量化处理;张路达等^[14]提出多尺度融合的 YOLOv3 人群口罩佩戴检测方法,但缺少口罩佩戴错误的检测;张烈平等^[15]提出一种基于迁移学习的轻量化 YOLOv2 口罩佩戴检测方法,提高了模型训练速度与检测速度,但检测精度较低。因此,对于这些存在的问题,本文提出了一种改进 YOLOv8 密集人群口罩检测算法,具体改进部分如下:

1)为了解决目标遮挡和检测目标较小的问题,引入聚集分配(gather-and-distribute,GD)机制,GD 较于传统的特征金字塔(feature pyramid networks,FPN)具有更好的特征提取效果,可以融合跨层信息并做到无损传输。

2)为了进一步提高网络特征提取能力,插入全维度动态卷积(omni-dimensional dynamic convolution, ODconv),提高对小物体检测精度。

3)为了平衡 GD 机制和 ODConv 带来的模型增大、复杂度增加的问题,引入共享卷积参数 BN 检测头(shared convolutional separamter BN detection head, SCSBD),在参数量更少、计算量更少的情况下,提高精度并减少模型大小。

1 YOLOv8

YOLOv8 作为 YOLO 系列的模型,它依旧是一种单阶段的目标检测器,YOLOv8 通过单个前向传递来直接预测图像中的目标位置和类别,无需复杂的多阶段过程。

YOLOv8 主要由主干网络(Backbone)、颈部网络(Neck)和检测头(Head)3 个部分构成。在骨干网络和颈部网络部分,YOLOv8 将 YOLOv5^[16-17]的 C3 结构换成了梯度流更为丰富的 C2f 结构。检测头部分从 Anchor-Based^[18]改变为使用 Anchor-Free^[18-19]来处理目标的位置和大小。损失函数部分,YOLOv8 的 Loss 计算包括分类分支和回归分支两个部分,分类分支采用 BCE Loss,回归分支则采用了数据平均保真度(Distribution Focal Loss,DFL)^[20-21]损失函数和 CIoU Loss。YOLOv8 提供了 N/S/M/L/X 5 种不同尺度大小的模型。本文基于对密集人群口罩检测的实时状况,选择 YOLOv8s 模型权衡精度与模型大小。

2 改进的 YOLOv8 算法

在使用 YOLOv8 完成口罩检测任务时,发现模型针对小目标和重叠物体存在漏检和误检现象。针对检测中存在的不足对 YOLOv8 模型进行改进,改进后的模型如图 1 所示。

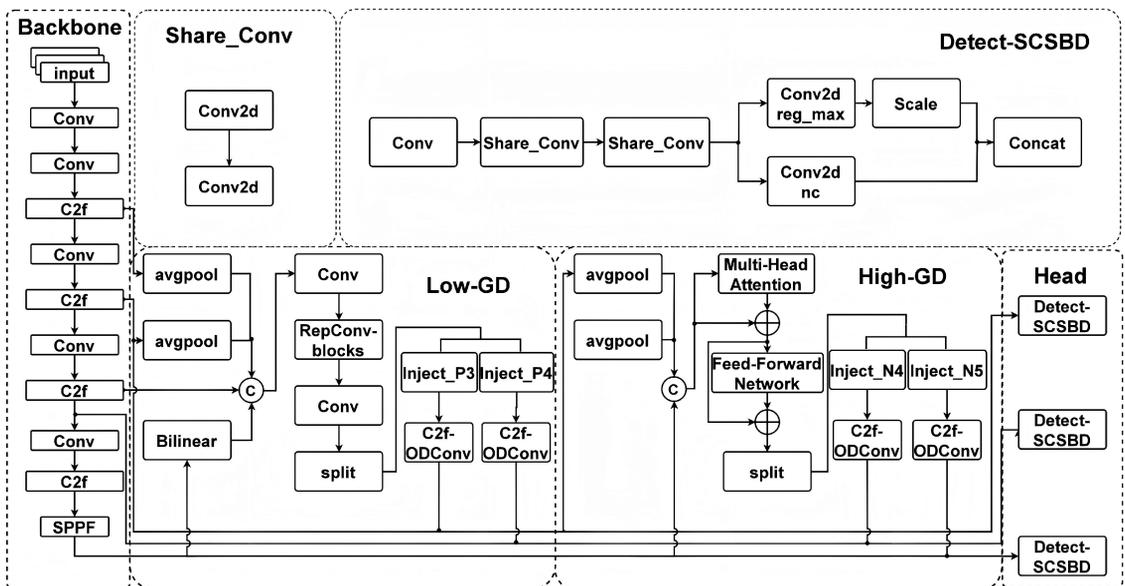


图 1 改进 YOLOv8 网络结构

Fig. 1 Diagram of improved YOLOv8 network structure

GD 机制分为 Low-GD 和 High-GD(如图 1),采用 GD

替换 YOLOv8 的 FPN 结构,使模型更好的进行信息融合。

同时在注入(Inject)模块后插入 C2f-ODConv^[22],引导网络更好关注细节特征信息,提高网络对小目标和信息缺失目标检测的准确性和鲁棒性。采用 SCSBD 替换 YOLOv8 原有的检测头,运用共享卷积和批量归一化优化检测头结构,减轻网络负担并提高检测精度。

2.1 聚集-分配机制

目标检测过程中,通常先经过骨干网络提取得到一系列不同层级的特征。FPN^[23]利用了骨干网络可以提取一系列不同层级特征的这一特点,构建了相应的融合结构:不同层级的特征包含着不同大小物体的位置信息。其层层递进的信息融合模式使相邻的信息能够充分融合,但由于没有直连的交互通路,只能依靠中间层充当“中介”进行融合,导致了跨层信息融合困难,因此参考 Gold-YOLO^[24]的结构引入 GD 机制解决跨层信息融合困难的问题。

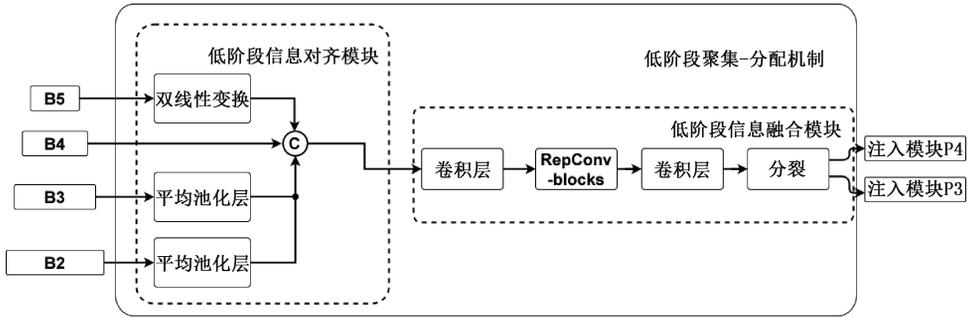


图2 低阶段 GD 机制

Fig.2 Low-GD mechanism

高阶阶段 GD 机制(如图3)中的特征对齐模块由平均池化层组成,将输入特征的维度降低到统一的大小,输入到信息融合模块中,使用 Transformer 块组合从高阶信息对齐模块导出的 F_{align} 获得 F_{fuse} , F_{fuse} 同样在通道维度中被拆

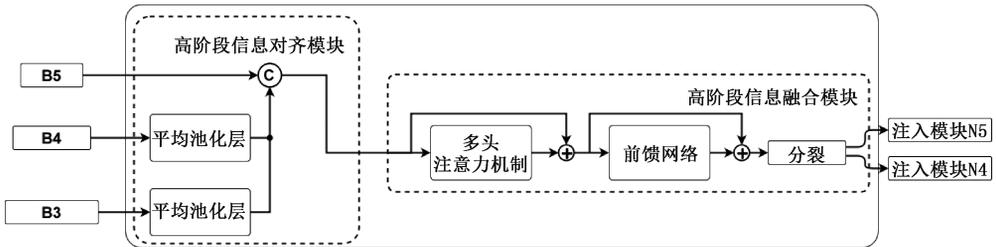


图3 高阶阶段 GD 机制

Fig.3 High-GD mechanism

2.2 全维度动态卷积

现有的动态卷积一般采用单个注意力,通过对多个卷积核进行线性加权,这种加权与输入相关,因此,输入对于输出卷积核有相同的注意力值,但关于卷积核的空间维度、输入通道维度以及输出通道维度3个其他维度都被忽略了。

ODConv^[26]利用一种新颖的多注意机制和并行策略,

GD 机制通过3个模块实现:信息对齐模块(feature alignment module, FAM)、信息融合模块(information fusion module, IFM)和信息注入模块。这3个模块分别负责不同的功能,信息对齐模块负责收集并对齐不同层级不同大小的特征,信息融合模块通过使用卷积或 Transformer^[25]算子对对齐之后的特征进行融合,得到全局信息,信息注入模块将全局信息注入到不同层集中。

低阶段 GD 机制如图2所示,其中,特征对齐模块使用平均池化和双线性变换操作实现统一大小的输入特征,得到信息融合模块中的 F_{align} , RepBlock 块将 F_{align} 作为输入产生 F_{fuse} , 生成的 F_{fuse} 随后在通道维度中被拆分为 F_{P3} 和 F_{P4} 与不同级别的特征尺度融合,其公式如下:

$$F_{align} = \text{Low_FAM}([B2, B3, B4, B5]) \quad (1)$$

$$F_{fuse} = \text{RepBlock}(F_{align}) \quad (2)$$

$$F_{P3}, F_{P4} = \text{Split}(F_{fuse}) \quad (3)$$

分为 F_{N4} 和 F_{N5} 与其当前级别特征融合,公式如下:

$$F_{align} = \text{High_FAM}([P3, P4, P5]) \quad (4)$$

$$F_{fuse} = \text{Transformer}(F_{align}) \quad (5)$$

$$F_{N4}, F_{N5} = \text{Split}(\text{Conv1} \times 1(F_{fuse})) \quad (6)$$

在任意卷积层沿内核空间的4个维度(卷积核的空间大小、输入通道数、输出通道数和卷积核数量)学习卷积核的注意力。ODConv 可以视为 CondConv^[27]的延续, CondConv 如图4所示,其采用单个注意力标量,输出滤波器对于输入具有相同的注意力,但 CondConv 忽视了卷积和的空间维度、输入通道维度以及输出通道维度, ODConv 将 CondConv 中一个维度上的动态特性进行了扩展,同时

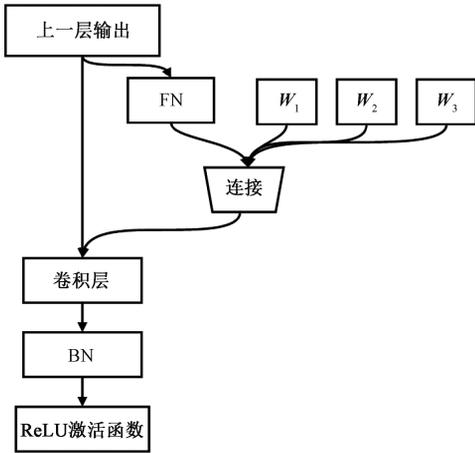


图 4 CondConv 结构图

Fig. 4 CondConv structure diagram

考虑了空域、输入通道、输出通道等维度上的动态特性，故称之为全维度动态卷积。ODConv 的实现流程如图 5 所示。

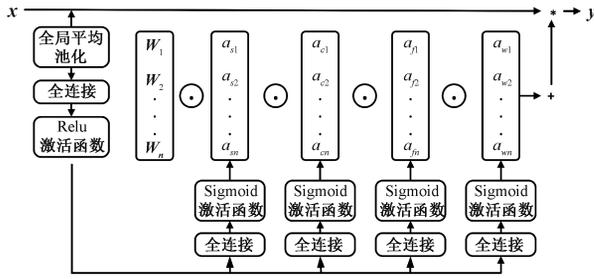


图 5 ODConv 结构图

Fig. 5 ODConv structure diagram

在 ODConv 中，对于卷积核 W_i ， a_{si} 是 $k \times k$ 空间位置的卷积参数分配的注意力标量， a_{ci} 是为每个卷积滤波器 W_i 的 c_{in} 通道分配的不同注意力标量， a_{fi} 为 c_{out} 卷积滤波器分配的不同注意力标量；图 5 说明了这 4 种类型的注意力与不同卷积相乘的过程，这些注意力以位置方式、通道方式、滤波器方式和卷积核方式的顺序与卷积内核 W_i 相乘使得卷积操作相对于卷积核是不同的。具体的说，通过全局平均池化的操作将输入 x 压缩成一个长度为 c_{in} 的特征向量，随后进入一个全连接层和 4 个分支，在全连接层和分支之间设置一个 *Relu* 激活函数，全连接层将压缩后的特征向量映射到具有缩减比 r 的低维空间中。在 4 个头部分支中，每个分支都有一个输出大小分别为 $k \times k, c_{in} \times 1, c_{out} \times 1$ 和 $n \times 1$ 的全连接层以及一个 *Sigmoid* 激活函数，4 个分支分别生成 a_{si}, a_{ci}, a_{fi} 和 a_{wi} 。输出 y 的计算公式：

$$y = (a_{w1} \odot a_{f1} \odot a_{c1} \odot a_{s1} \odot W_1 + \dots + a_{wn} \odot a_{fn} \odot a_{cn} \odot a_{sn} \odot W_n) * x \quad (7)$$

其中， $a_{wi} \in R$ 表示卷积核 W_i 的注意力标量； $a_{si} \in R^{k \times k}$ ， $a_{ci} \in R^{c_{in}}$ 和 $a_{fi} \in R^{c_{out}}$ 表示引入的 3 个新注意力，分

别沿着卷积核 W_i 的核空间的空间维度、输入通道维度和输出通道维度计算； \odot 表示沿着核空间的不同维度的乘法运算。

2.3 轻量化检测头

YOLOv8 中的检测头几乎占据了模型整个计算量的 1/5，其解耦头的结构如图 6 所示，提取特征头后的两个分支分别通过两个 3×3 的卷积以及一个 1×1 的卷积来提取信息，分别结算 Bbox. loss 和 Cls. loss，用于单独预测边界框损失和分类损失，同时对应 3 个不同大小的尺寸对不同大小的特征图进行检测，而 3×3 的卷积在通道加深的同时会导致参数数量和浮点运算量的急剧增加。

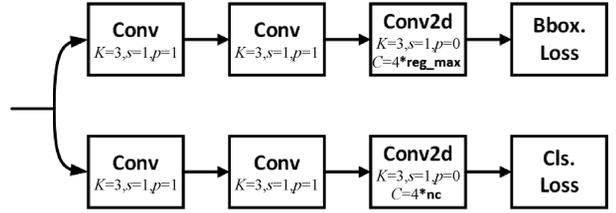


图 6 YOLOv8 检测头

Fig. 6 YOLOv8 detection head

为进一步简化检测头的计算量，对检测头进行改进，改进后的轻量化检测头 SCSBD 结构如图 7 所示。运用共享卷积和 Batch Normalization 对冗余的 YOLOv8 检测头进行优化：

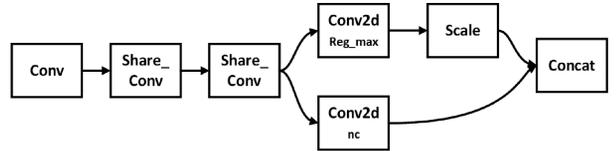


图 7 SCSBD 结构图

Fig. 7 SCSBD structure diagram

1) 共享卷积参数

由图 6 检测头的结构可以看出两个分支的卷积层大小参数是相同的，因此使两个卷积分支输入共享同一个卷积层和两个共享卷积层，为了对应每个检测头所检测的目标尺度不一致的问题，使用 *Scale* 层对特征进行缩放，这种共享参数的设计允许两个分支在计算损失时共享相同的卷积核权重，从而避免重复计算相似的特征图，进一步加速整个模型的推理过程。

2) 批量归一化(batch normalization, BN)

由于不同层级之间特征的统计量仍存在差异，归一化层依然是必须的，BN^[28] 可以优化网络，减少梯度爆炸并提升模型的整体精度。BN 通过批次内计算的均值和方差对特征进行归一化，因此 BN 特别依赖 batch size，减少 batch size 的大小会对估计的批量统一数据产生巨大影响，Wu 等^[29] 在 *Group Normalization* 一文中指出了当 batch size 的数值减小时，使用组归一化(group normalization, GN)产生的误差比 BN 低，BN 与 GN 的归一化方法如图 8 所示，

本文采用的批量大小为32, batch size数值较大,因此BN的效果优于GN,由于直接在共享参数的检测头中引入BN会导致其滑动平均值而产生误差,而引入GN又会增加推理时的开销,参考NASFPN^[30]的做法,引入BN并分别独立计算。

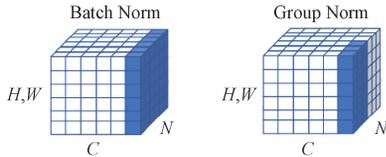


图8 BN与GN归一化方法

Fig.8 BN and GN normalization method

3 实验与结果分析

3.1 数据收集与评价指标

本文使用的数据集共收集了789张图片,来源于Kaggle上公开的Face Mask Detection数据集。本文采用了如图9所示马赛克(Mosaic)数据增强方法扩充数据集至2 232张图片,该方法通过随机缩放、随机裁剪和随机排列不同的图片进行拼接,从而生成新的训练数据。



图9 Mosaic数据增强

Fig.9 Mosaic data augmentation

将扩充后的数据集图片按7:2:1的比例随机划分为训练集、验证集和测试集。数据集被分为3个类别:佩戴口罩、未佩戴口罩和口罩佩戴错误。这种分类方式有助于模型在实际场景中更好地检测不同类型的口罩佩戴情况。

实验在Ubuntu 22.04操作系统上进行,硬件为NVIDIA RTX 4090 GPU。编译环境为Python 3.10、PyTorch 2.1.0和CUDA 12.1。在训练过程中,统一使用SGD优化器,初始学习率设为0.01,权重衰减系数为0.0005,动量参数为0.9375。训练总轮次设置为600个epoch, batch size为32。为了加快模型的训练速度和减少计算资源的消耗,当损失函数趋于稳定时,采用提前终止训练的策略,以确保模型在保持高效训练的同时获得良好的性能。

目标检测的评价指标一般分为:准确率(precision, P)、召回率(recall, R)、F1分数(F1 Score)、平均精确度均值(average precision, mAP)和模型大小。

1)准确率:检测出的目标属于目标类别的比例,正确检测出的目标数量与所有检测出的目标数量比值。公式为:

$$P = \frac{TP}{TP + FP} \quad (8)$$

2)召回率:真正被检测出的目标数量与数据集中所有真是目标的数量之比,可以衡量模型的全局性,反映模型检测目标的能力。公式为:

$$R = \frac{TP}{TP + FN} \quad (9)$$

3)平均精确度均值:计算了模型在不同类别上的精确度并进行平均操作。mAP通常会使用不同的IOU阈值来计算每个类别的精确度。

$$AP = \int_0^1 P(R) dR \quad (10)$$

$$mAP = \frac{\sum_{n=1}^N AP(n)}{N} \quad (11)$$

式中:TP表示预测正确样本的个数,FP表示非样本的个数,FN表示未检测出样本的个数。

4)权重文件大小(Size)和参数量(Parameter)用于检验模型大小以及与硬件的适配性。

3.2 实验结果与可视化分析

为了验证改进方法对口罩检测模型性能的有效性,将改进后的模型与YOLOv8模型进行相同数据集的对比检验。表1中分别罗列使用YOLOv8与改进后的YOLOv8对数据集的3个类别进行检测的平均精度,从表1结果可以看出实验证明了改进YOLOv8模型在进行口罩检测时有显著的精度提升,证明了改进策略的有效性,未佩戴口罩、佩戴口罩、口罩佩戴错误的检测平均精度分别提升了4.4%、1.9%、14.3%,总体性能提升6.9%。

表1 模型各项类别检测的平均精度对比

Table 1 Comparison of AP for each category in model detection

模型	model detection			mAP50
	未佩戴口罩	佩戴口罩	口罩佩戴错误	
YOLOv8s	77.9	91.4	53.3	74.2
本文	82.3	93.3	67.6	81.1

使用YOLOv8模型和改进后的v8模型分别对一部分具有一定复杂度的密集人群口罩图片进行实际检测。检测结果图10(a)和(b)展示了两种模型在远距离和低分辨率目标检测中存在的差异,YOLOv8模型在检测过程中对检测图中远处低分辨率目标存在较多误检情况,而改进后的模型准确性更高,误检率显著降低。这表明改进后的模型在处理低分辨和小目标时,能够更加准确的对口罩佩戴情况进行分类,有效减少误检。



(a) YOLOv8s (b) 本文
(b) Ours

图 10 检测结果

Fig. 10 Detection results

热力图如图 11 所示,进一步验证了这一结论,通过热力图的可视化效果可以更为直观地观察到原始模型如图 11(a)在某些图片中的口罩检测错误分布情况,这些误检通常集中在远处低分辨率小物体区域,尤其是在目标与背景存在一定相似性的情况下,改进后的模型在相同的测试环境下,表现出更高的稳定性和精确度,如图 11(b)。

3.3 消融实验

为了体现本文提出的各个改进模块对 YOLOv8 性能的提升效果,对改进的 YOLOv8 进行消融实验,以 YOLOv8 作为基本模型,分别将 SCSBD、ODConv、GD 逐步添加到 YOLOv8 模型中,并分别给出多个改进模块组合的实验效果,消融结果如表 2 所示。

将颈部替换为 GD 后,由于 GD 使用一个统一的模块收集和融合所有层级的信息并分发到不同的层级,GD 网络通过这种方式避免了传统 FPN 结构固有的信息丢失问题,加强了中间层的部分信息融合能力,使得模型在准确率、召回率和精确率方面都有显著的提高。但这种无损传输的复杂的机制不可避免的带来了参数和计算量的增加的问题,由表 2 结果可知,替换 GD 的模型相比原模型精确度、召回率和平均精确度分别有 5.7%、5.6%、3.8%的提



(a) YOLOv8s



(b) 本文
(b) Ours

图 11 热力图

Fig. 11 Heatmap

表 2 消融实验

Table 2 Ablation test

基准模型	SCSBD	OD	GD	P/%	R/%	mAP50/%	模型大小/M	参数量/ 10^6
YOLOv8s	×	×	×	81.7	67.6	74.2	22.5	11.1
YOLOv8s	√	×	×	91.2	71.2	78.7	19.1	9.4
YOLOv8s	×	√	×	89.0	71.6	77.5	24.2	11.2
YOLOv8s	×	×	√	87.4	73.2	78.0	27.7	13.6
YOLOv8s	√	√	×	89.9	64.9	75.6	19.3	9.5
YOLOv8s	×	√	√	83.9	70.7	79.2	28.0	13.7
YOLOv8s	√	×	√	91.3	70.4	77.0	24.7	12.1
YOLOv8s	√	√	√	95.3	69.1	81.1	25.0	12.1

升,但模型同样增加了 5.2 MB 的权重文件大小和 2.5×10^6 的参数量;为了进一步提升模型动态特征提取能力,提高低分辨率图片和小物体的检测精度,在 GD 网络中融合全维度动态卷积 ODConv,ODConv 采用多注意力机制,考虑空域、输入通道、输出通道等多个维度上的动态特性,对 GD 机制无损传输的信息进行进一步的全面特征提取,根据表 2 的结果可知,添加 ODConv 进一步提升了模型 1.2% 的平均精度,但其权重文件和参数量进一步提高达到 28 MB 和 13.7×10^6 。

为了平衡使用 GD 和 ODConv 带来的权重文件和参数量的增大,对检测头进行轻量化替换为 SCSBD,更换检测头后的模型仅在召回率降低 1.6% 的情况下做到了对准确率、平均精确度、权重文件和参数量的全面优化,准确率、平均精确度分别提升了 11.4%、1.9%,权重文件和参数量分别减轻了 3 MB 和 1.6×10^6 。

通过改良的 YOLOv8 模型可以明显看出,更换颈部网络和引入轻量化检测头,显著优化了网络的整体结构和信息传递方式。这种改进不仅在一定程度上提升了网络的表达能力,同时也在模型参数和大小仅有少量增加的情况下,大幅度提高了对小目标和低分辨率物体的检测能力,改良后的模型大大增强了其在复杂场景中的适用性和鲁棒性。

3.4 对比实验

为了证明本文的改进算法模型的优越性和有效性,将改进后的 YOLOv8 算法和其他算法进行了比较,本文选择了 RT-DETR^[31]、YOLOv5s、YOLOv9s^[32]、YOLOv10s^[33]。由表 3 中结果可以看出,RT-DETR 模型不仅本身权重文件较大,对口罩的检测效果也不如除 YOLOv8s 外的其他 YOLO 系列,这是由于 Transformer 主要关注全局特征,对小物体关注局部特征的检测能力有限;同时,相较于其他一系列 YOLO 系列算法 YOLOv5s、YOLOv8s、YOLOv9s 和 YOLOv10s,本文所提出的算法仅在少量的增大模型权重文件大小和计算量的基础上,mAP 分别获得了 3.7%、6.9%、3.3%、3.9% 的提升。

表 3 不同算法对比结果

Table 3 Comparison of results from different algorithms

模型	P/%	R/%	mAP50/%	模型大小/M
RT-DETR	86.5	68.3	75.1	40.4
YOLOv5s	86.0	72.8	77.4	18.5
YOLOv8s	81.7	67.6	74.2	22.5
YOLOv9s	86.8	73.8	77.8	15.3
YOLOv10s	86.0	71.3	77.2	16.6
本文	95.3	69.1	81.1	25.0

4 结 论

本文针对密集人群口罩检测困难的问题,设计了一种

在复杂密集人群场景下进行口罩检测的改进 YOLOv8 算法。在颈部网络中使用 GD 网络替换原有的 FPN,做到信息的跨层无损传输,增强对小目标特征的提取能力,融入 ODConv 使 GD 提取小目标特征能力进一步提升,在头部网络,采用 SCSBD 轻量化检测头优化 YOLOv8 的检测头并提高检测精度。实验表明,与现有的模型相比,改进后的 YOLOv8s 算法精度更高、对小目标和分辨率低的物体检测效果更好,具有实际的应用价值。

未来将在该模型的基础上优化使模型在面对更加复杂的密集人流下对口罩的检测能力,同时提升模型的检测速度。同时为了适应算力更低的边缘设备,考虑使用新的技术减小模型参数量,以便该模型能更好的适应不同的复杂场景。

参考文献

- [1] 朱浩,施培武,沈群红,等. 新冠疫情前后中国内地重大疫情防控工作职责明确程度变化分析[J]. 中国公共卫生,2024,40(3): 274-279.
ZHU H, SHI P W, SHEN Q H, et al. Changes in clear assignment of departmental responsibilities for major epidemic prevention and control in the Chinese mainland before and after COVID-19 epidemic: An analysis on official documents[J]. Chinese Journal of Public Health, 2024,40(3):274-279.
- [2] 甘雨,吴雨,王建勇. 新冠肺炎疫情趋势预测模型[J]. 智能系统学报,2021,16(3):528-536.
GAN Y, WU Y, WANG J Y. Epidemics trend prediction model of COVID-19[J]. CAAI Transactions on Intelligent Systems, 2021,16(3):528-536.
- [3] SUN Y B, SUN Z, CHEN W T. The evolution of object detection methods[J]. Engineering Applications of Artificial Intelligence, 2024,133: 108458.
- [4] REDMON J, FARHADI A. YOLOv3: An incremental improvement [J]. ArXiv preprints arXiv: 1804.02767, 2018.
- [5] BELABBES M A, OUKDACH Y, SOUAIDI M, et al. Advancements in polyp detection: A developed single shot multibox detector approach [J]. IEEE Access, 2024,12: 19199-19215.
- [6] WANG Y Y, WANG C, ZHANG H, et al. Automatic ship detection based on RetinaNet using multi-resolution Gaofen-3 imagery [J]. Remote Sensing, 2019,11(5): 531.
- [7] REN SH 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] 翟永杰,王璐瑶,郭聪彬. 基于 Faster R-CNN 的复杂背景下绝缘子目标检测[J]. 电子测量技术, 2023, 46(16): 187-194.
ZHAI Y J, WANG L Y, GUO C B. Insulator object detection in complex background based on Faster R-CNN[J]. Electronic Measurement Technology, 2023, 46(16):187-194.
- [9] HE K M, GKIOXARI G, DOLLAR P, et al. Mask R-CNN [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42(2): 386-397.
- [10] CAI Z, VASCONCELOS N. Cascade R-CNN: Delving into high quality object detection[J]. IEEE Conference on Computer Vision and Pattern Recognition, 2018: 6154-6162.
- [11] 高民,陈高华,古佳欣,等. FLM-YOLOv8: 一种轻量级的口罩佩戴检测算法[J]. 计算机工程与应用, 2024, 60(17): 203-215.
GAO M, CHEN G H, GU J X, et al. FML-YOLOv8: Lightweight mask wearing detection algorithm [J]. Computer Engineering and Applications, 2024, 60(17): 203-215.
- [12] 魏明军,周天宇,纪占林,等. 基于 YOLOv3 的公共场所口罩佩戴检测方法[J]. 广西师范大学学报(自然科学版), 2023, 41(1): 76-86.
WEI M J, ZHOU T Y, JI ZH L, et al. Detection method of mask wearing in public places based on YOLOv3[J]. Journal of Guangxi Normal University (Natural Science Edition), 2023, 41(1): 76-86.
- [13] 李小波,李阳贵,郭宁,等. 融合注意力机制的 YOLOv5 口罩检测算法[J]. 图学学报, 2023, 44(1): 16-25.
LI X B, LI Y G, GUO N, et al. Mask detection algorithm based on YOLOv5 integrating attention mechanism[J]. Journal of Graphics, 2023, 44(1): 16-25.
- [14] 张路达,邓超. 多尺度融合的 YOLOv3 人群口罩佩戴检测方法[J]. 计算机工程与应用, 2021, 57(16): 283-290.
ZHANG L D, DENG CH. Multi-scale fusion of YOLOv3 crowd mask wearing detection method[J]. Computer Engineering and Applications, 2021, 57(16): 283-290.
- [15] 张烈平,李智浩,唐玉良. 基于迁移学习的轻量化 YOLOv2 口罩佩戴检测方法[J]. 电子测量技术, 2022, 45(10): 112-116.
ZHANG L P, LI ZH H, TANG Y L. Light-YOLOv2 mask wearing detection method based on transfer learning [J]. Electronic Measurement Technology, 2022, 45(10): 112-116.
- [16] CHEN M, LIU Y Q, WEI X, et al. PO-YOLOv5: A defect detection model for solenoid connector based on YOLOv5[J]. PloS One, 2024, 19(1): e0297059.
- [17] 赵鸿亮,郭佑民,王建鑫,等. 基于改进 YOLOv5 的轨道交通障碍物检测算法[J]. 电子测量技术, 2024, 47(1): 130-135.
ZHAO H L, GUO Y M, WANG J X, et al. Rail transit obstacle detection algorithm based on improved YOLOv5 [J]. Electronic Measurement Technology, 2024, 47(1): 130-135.
- [18] ZHANG S, CHI C, YAO Y, et al. Bridging the gap between anchor-based and anchor-free detection via adaptive training sample selection [C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 9759-9768.
- [19] TIAN ZH, SHEN CH H, CHEN H, et al. FCOS: A simple and strong anchor-free object detector[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022, 44(4): 1922-1933.
- [20] LI X, LYU CH Q, WANG W H, et al. Generalized focal loss: Towards efficient representation learning for dense object detection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2023, 45(3): 3139-3153.
- [21] 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.
- [22] XU W Y, LI X, JI Y CH, et al. BD-YOLOv8s: Enhancing bridge defect detection with multidimensional attention and precision reconstruction[J]. Scientific Reports, 2024, 14(1): 18673.
- [23] LIN T Y, DOLLAR P, GIRSHICK R, et al. Feature pyramid networks for object detection [C]. IEEE Conference on Computer Vision and Pattern Recognition, 2017.
- [24] WANG CH CH, HE W, NIE Y, et al. Gold-YOLO: Efficient object detector via gather-and-distribute mechanism [J]. Advances in Neural Information Processing Systems, 2024, DOI: 10.48550/arXiv.2309.11331.
- [25] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[J]. ArXiv preprint arXiv: 1706.03762, 2017.
- [26] LI CH, ZHOU AO J, YAO AN B. Omni-dimensional dynamic convolution[J]. ArXiv preprint arXiv: 2209.07947, 2022.
- [27] YANG B, BENDER G, NGIAM J, et al. Condconv:

- Conditionally parameterized convolutions for efficient inference [J]. *Advances in Neural Information Processing Systems*, 2019, DOI: 10.48550/arXiv.1904.04971.
- [28] 张婷,王海淇,张认成,等. 基于自归一化神经网络的电弧故障检测方法[J]. *仪器仪表学报*, 2021, 42: 141-149.
ZHANG T, WANG H Q, ZHANG R CH, et al. An arc fault detection method based on the self-normalized convolutional neural network[J]. *Chinese Journal of Scientific Instrument*, 2021, 42: 141-149.
- [29] WU Y, HE K M. Group normalization [J]. *International Journal of Computer Vision*, 2020, 128(3):742-755.
- [30] GHIASI G, LIN T Y, LE Q V. NAS-FPN: Learning scalable feature pyramid architecture for object detection [C] *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019: 7036-7045.
- [31] ZHAO Y, LYU W Y, XU S L, et al. DETRs beat YOLOs on real-time object detection[C]. *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2024:16965-16974.
- [32] WANG C Y, YE H I H, LIAO H Y M. YOLOv9: Learning what you want to learn using programmable gradient information [C]. *European Conference on Computer Vision*. Springer, Cham, 2025:1-21.
- [33] WANG AO, CHEN H, LIU L H, et al. YOLOv10: Real-time end-to-end object detection [J]. *ArXiv preprint arXiv: 2405.14458*, 2024.

作者简介

伍锡如,教授,博士,主要研究方向为深度学习、复杂网络、机器人控制。

E-mail: xiru@guet.edu.cn

梁诗意(通信作者),硕士研究生,主要研究方向为深度学习、图像处理、计算机视觉。

E-mail:1697889521@qq.com