

doi:10.3969/j.issn.1673-9833.2024.06.013

复杂场景下的改进 YOLOv7-tiny 安全帽佩戴检测算法

雷源毅, 朱文球, 廖 欢

(湖南工业大学 计算机学院, 湖南 株洲 412007)

摘 要: 针对建筑工地复杂场景下安全帽图像因背景信息复杂、光照强度和拍摄角度不同等特点, 导致 YOLOv7-tiny 安全帽检测算法识别精度低, 易出现漏检以及误检等问题, 提出了一种 MCWE-YOLO 安全帽检测算法。通过改进 Mosaic 方法进行数据增强, 丰富数据多样性; 将 CoordAttention 注意力模块嵌入 YOLOv7-tiny 模型中, 以增强模型对小目标语义信息的提取能力和全局感知能力; 提出一种新的 FR-ODConv 动态卷积代替 Head 中的静态卷积, 并自适应动态调整卷积核的权值, 使模型能够更好地提取不同形状和大小安全帽的关键特征; 改进算法的损失函数, 以聚焦于普通质量的锚框, 提高模型检测整体性能; 提出用 Res-EDH 解耦头结构替换原始的 Detect 结构, 对分类和回归进行解耦, 在提升精度的同时降低了延时。最后使用开源 GDUT-HWD 安全帽数据集进行训练测试, 实验结果表明, MCWE-YOLO 算法的 *mAP* 达到 87.3%, 相比原始 YOLOv7-tiny 算法提升了 1.6%, 对比其他算法的最优值提升了 2.1%, 能有效提高安全帽检测精度, 实现建筑工地安全帽的自动化检测。

关键词: YOLOv7-tiny; CoordAttention 注意力机制; FR-ODConv 动态卷积; Res-EDH 解耦头

中图分类号: TP391.4

文献标志码: A

文章编号: 1673-9833(2024)06-0093-08

引文格式: 雷源毅, 朱文球, 廖 欢. 复杂场景下的改进 YOLOv7-tiny 安全帽佩戴检测算法 [J]. 湖南工业大学学报, 2024, 38(6): 93-100.

Improved YOLOv7-tiny Helmet Wearing Detection Algorithm in Complex Scenes

LEI Yuanyi, ZHU Wenqiu, LIAO Huan

(College of Computer Science, Hunan University of Technology, Zhuzhou Hunan 412007, China)

Abstract: Aiming at the characteristics of helmet images in complex scenes of construction sites, such as complex background information, different light intensity and shooting angles, which lead to low recognition accuracy of YOLOv7-tiny helmet detection algorithm, and prone to missed detection and false detection, a MCWE-YOLO helmet detection algorithm has been proposed. Data enhancement is performed by improving the Mosaic method to enrich data diversity; the CoordAttention module is embedded in the YOLOv7-tiny model to enhance the model's ability to extract small target semantic information and global perception capabilities; a new FR-ODConv dynamic convolution is proposed to replace the static convolution in the head, and the weights of the convolution kernel are adaptively and dynamically adjusted, so that the model can better extract the key features of helmets of different shapes

收稿日期: 2023-10-18

基金项目: 科技创新 2030—“新一代人工智能”基金资助重大项目(2018AAA0100400); 湖南省自然科学基金资助项目(2021JJ50058, 2022JJ50051)

作者简介: 雷源毅, 男, 湖南工业大学硕士生, 主要研究方向为图像处理与目标检测, E-mail: yuanyi_lei@foxmail.com

通信作者: 朱文球, 男, 湖南工业大学教授, 主要研究方向为人工智能, 计算机视觉, 大数据技术,

E-mail: wenqiu_zhu@126.com

and sizes; the loss function of the algorithm is improved to focus on anchor boxes of ordinary quality and improve the overall detection performance of the model; a new Res-EDH decoupling head structure is proposed to replace the original detect structure, decouple classification and regression, and reduce latency while improving accuracy. The open source GDUT-HWD helmet dataset was used for training and testing. The experimental results show that the mAP of the MCWE-YOLO algorithm reaches 87.3%, which is 1.6% higher than the original YOLOv7-tiny algorithm and 2.1% higher than the optimal value of other algorithms. It can effectively improve the accuracy of helmet detection and realize the automatic detection of helmets on construction sites.

Keywords: YOLOv7-tiny; CoordAttention module; FR-ODConv; Res-EDH decoupling head

1 研究背景

作为生产生活中常见的防护用品,在生产作业中佩戴安全帽可以保护工人的生命安全。然而在实际工程项目中,总会出现工人忘记或不戴安全帽的情况,这一方面会给施工现场工人的安全造成隐患,另一方面也会延误工程项目的工作进度。因此,按要求佩戴好安全帽在建筑工地施工场景中对于施工工人来说必不可少。

安全帽图像在建筑工地中具有光照条件复杂、目标分布密集、目标尺度变化大、背景复杂等特点,给目标检测带来了很大的困难。孙国栋等^[1]通过融合注意力机制改进 Faster-RCNN 算法,以改善安全帽检测难度和各种场景下的高漏检率。吕宗喆等^[2]通过对 YOLOv5 算法中损失函数的优化和引进切片辅助调节,让模型在复杂密集的环境中提高了检测小目标的准确率。丁田等^[3]在 YOLOX 检测模型的特征提取网络中加入注意力机制等方法,提高了目标特征表达能力。赵敏等^[4]通过对 YOLOv7-tiny 采用更为先进的定位损失函数 SIOU Loss 计算损失,添加所需回归的向量角度,提高算法训练过程中预测框的收敛速度及效率。上述改进算法仍存在检测精度不够高、背景特征提取不够丰富、泛化性和实时性较差等情况,未能较好解决小目标密集、安全帽和背部同色等复杂场景下的漏检、误检问题。

针对这些问题,本文结合施工现场的特点,对 YOLOv7-tiny 模型进行了改进,改进后的算法显著提高了安全帽佩戴检测的精度,实现了对施工现场安全帽的有效检测。

2 YOLOv7-tiny 简介

Ultralytics 于 2022 年 7 月发布了 YOLOv7^[5] 算法,与之前的网络相比,YOLOv7 在减少网络参数的同时提高了检测精度和实时性,选择了一个轻量版本

的 YOLOv7-tiny^[6] 作为基础网络进行改进。YOLOv7-tiny 作为单阶段目标检测网络,主要包括三部分:骨干网络(Backbone)、颈部网络(Neck)和预测头(Head),与原始的 YOLOv7 相比,YOLOv7-tiny 在骨干网络中使用更简单的 ELAN 代替 E-ELAN;取消 MPConv 中的卷积运算,只使用池化进行下采样;保留了优化后的 SPP(spatial pyramid pooling)结构,为 Neck 层输入更加丰富的特征图;Neck 部分依然采用 PANet 结构进行特征聚合;在 Head 部分使用标准卷积代替 REPCov 来调整通道数。

在 YOLOv7-tiny 算法基础上,建立了一个 MCWE-YOLO 复杂建筑工地情况下安全帽佩戴检测模型。通过改进 Mosaic 数据增强,嵌入 CoordAttention 坐标注意力机制,提出一种新的多维动态卷积 FR-ODConv 替换 Head 中的静态卷积,改进原有的损失函数,同时,提出一种新的 Res-EDH 解耦头结构替换原始的 Detect 结构,该算法可以适应复杂场景,并获得更好的检测效果,其网络结构如图 1 所示,虚线标明处为具体改进位置。

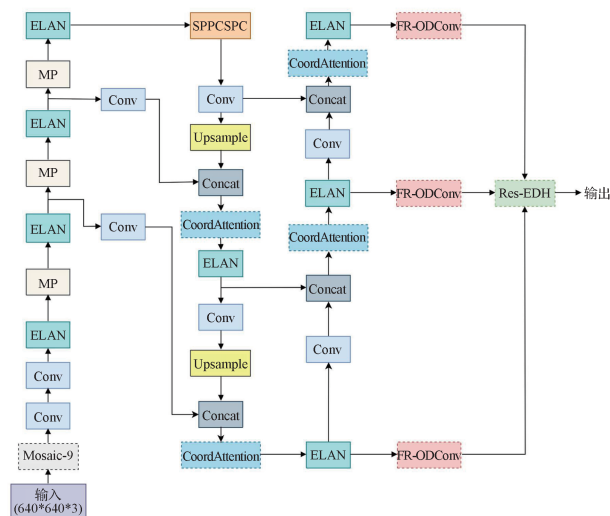


图 1 MCWE-YOLO 网络结构

Fig. 1 MCWE-YOLO network structure

3 模型改进

3.1 改进 Mosaic 数据增强

YOLOv7-tiny 算法中, 默认使用了 Mosaic^[7] 数据增强技术。该方法的过程如下: 首先随机选择 4 张图片, 然后对图像进行随机裁剪, 按特定顺序拼接成一张新的图像, 最后将拼接后的图像调整大小至指定的输入尺寸, 并将其用作网络模型的新训练数据, 具体流程如图 2 所示。

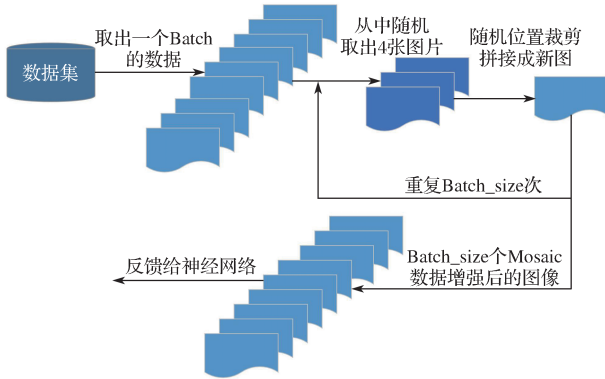


图 2 Mosaic 数据增强示意图

Fig. 2 Diagram of Mosaic data enhance

在实际工程场景中, 安全帽对象的比例通常较小, 常规模型训练无法充分学习到小对象的特征, 因此需要更多的数据扩充来提高模型的泛化能力。原始的 Mosaic 数据增强方法采用随机裁剪, 这种方法会使小目标和模糊密集目标被裁剪概率较高, 导致模型的输入样本仅包含背景信息。然而由于原始图像的比例尺不一致, 拼接后的图像往往会产生大量的白色边界, 这样会让模型受到许多无用特征信息的训练, 进而影响模型的收敛速度。

对 Mosaic 数据增强进行改进: 首先, 将拼接的图片数量从 4 个扩展到 9 个; 其次, 通过包围图像的最小矩形区域而不是原始的随机裁剪来裁剪, 如图 3 所示, 目的是尽可能减少白色边界区域, 并降低对象被裁剪的概率; 最后, 通过随机缩放、随机排列和组合来合成图像。

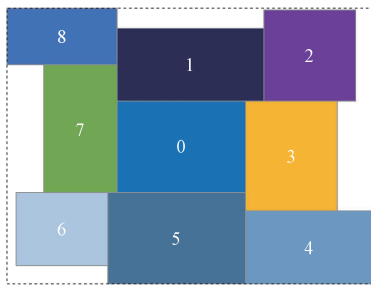


图 3 最小矩形区域

Fig. 3 Minimum rectangular area

Mosaic-9 数据增强方法减少了大量无用的边界区域, 提高了训练效率。图像拼接的数量从 4 个增加到 9 个, 丰富了小规模样本数据集, 混合 9 幅具有不同语义信息的图像可以增强网络的鲁棒性。此外, 9 个图像拼接在一起, 相当于增加 Batch_size。当执行批归一化时, 将同时计算 9 个图像的平均值和方差。因此, 批量样本数越大, 计算结果越接近整个数据集, 检测效果越好。

3.2 注意力机制

注意力机制是通过在全局信息中搜索关键的局部信息来处理数据, 帮助模型增强有用特征, 去除冗余特征, 使网络更多地关注有利于检测的有用信息。为了提高算法在特定区域的识别准确率, 在特征融合的过程中嵌入 CoordAttention 坐标注意力模块^[8-9], 模型不仅能够捕捉跨通道的信息, 还能够获取方向和位置信息, 从而更精准地定位和识别感兴趣的目标。可以提高特征提取性能, 更新组合特征图, 更好地防止连续多次下采样时目标信息的严重丢失, 得到带有方向和位置信息的特征图, 其结构如图 4 所示。

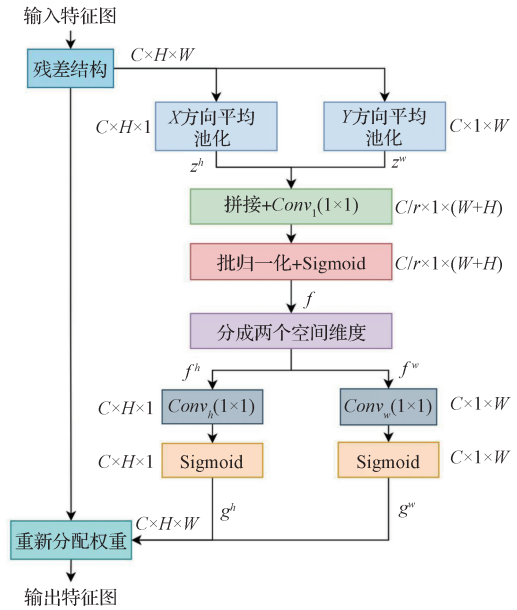


图 4 CoordAttention 结构

Fig. 4 CoordAttention structure

CoordAttention 注意力模块能够增强移动网络学习特征的表达能, 它可以对网络中的任意中间特征张量 X 进行转换变换后, 输出与输入尺寸相同的张量 Y :

$$X = [x_1, x_2, \dots, x_c] \in \mathbf{R}^{H \times W \times C}; \quad (1)$$

式中: H 和 W 分别为输入特征图的高度和宽度; C 为通道数。

$$Y=[y_1, y_2, \dots, y_c] \in \mathbf{R}^{H \times W \times C}. \quad (2)$$

为了获取图像宽度和高度上的注意力并对精确位置信息进行编码, CoordAttention 先将输入特征图分为高度 h 和宽度 w 两个方向进行 2 个 1×1 的卷积加强特征提取, 其中 $h \in H$ 、 $w \in W$ 。再分别进行全局平均池化, 分别得到在高度和宽度两个方向的特征图 z_c^h 和 z_c^w , 如式 (3) (4) 所示:

$$z_c^h(h) = \frac{1}{W} \sum_{0 \leq i < W} x_c(h, i), \quad (3)$$

$$z_c^w(w) = \frac{1}{H} \sum_{0 \leq j < H} x_c(j, w). \quad (4)$$

式中: $i=0, 1, \dots, H-1$; $j=0, 1, \dots, H-1$ 。

接着对池化后的两个通道输出的特征图进行 Concat 拼接, 并使用 1×1 大小的卷积核完成对特征图的卷积。在批归一化 (batch normalization) 和 Sigmoid 激活函数之后, 具有水平和垂直空间信息的输出特征图表示为式 (5):

$$f = \delta \left(\text{Conv}_1 \left([z^h, z^w] \right) \right). \quad (5)$$

式中: f 为具有水平和垂直空间信息的输出特征图; Conv_1 为卷积核为 1×1 的卷积; δ 为 Sigmoid 激活函数。

将上一步输出的特征图按照不同空间维度进行划分, 得到两个独立高度和宽度的特征图 f^h 和 f^w , 经过两次 1×1 卷积计算和调整 f^h 和 f^w 的通道数。通过 Sigmoid 激活函数将它们变换到 0~1 作为表示重要性水平的参数, 计算方法如式 (6) (7) 所示:

$$g^h = \sigma \left(\text{Conv}_h \left(f^h \right) \right), \quad (6)$$

$$g^w = \sigma \left(\text{Conv}_w \left(f^w \right) \right). \quad (7)$$

式中: g^h 和 g^w 分别为高度和宽度的重要性水平参数; σ 为 Sigmoid 激活函数; Conv_h 和 Conv_w 分别为卷积核为 $h \times h$ 、 $w \times w$ 的卷积块。

最后将输入 x 用不同维度的空间信息作为注意力权重进行加权编码, 输出 y 的表达式如式 (8) 所示:

$$y_c(i, j) = x_c(i, j) \times g_c^h(i) \times g_c^w(j). \quad (8)$$

3.3 FR-ODConv 动态卷积

动态卷积基于卷积层的输入确定每个卷积核的权重, 并且使用注意力机制对这些卷积核进行加权和求和, 以获得适合于该输入的动态卷积核。动态卷积的输出如式 (9) 所示。

$$\text{Output}(y) = x \times (\alpha_1 W_1 + \alpha_2 W_2 + \dots + \alpha_n W_n). \quad (9)$$

式中: y 和 x 为输出特征和输入特征; α_i ($i=1, 2, \dots, n$) 为注意力标量; n 为卷积核的数量; 每个卷积核 W_i ($i=1, 2, \dots, n$) 具有与标准卷积核相同的大小。

模型在 ODConv^[10-11] 的基础上结合 FReLU 激活函数提出一种 FR-ODConv 动态卷积结构, 利用多维注意力机制和并行策略来学习沿着内核空间的所有 4 个维度的任何卷积层处的卷积内核的互补注意力。这 4 个不同的注意力为卷积核输入通道的数量、卷积核本身的感受野、卷积核输出通道的数量以及卷积核的数量。这些注意力相互补充, 并按照位置、通道、过滤器和内核的顺序与卷积内核相乘, 使得卷积运算与输入 x 的所有空间位置、所有输入通道、所有过滤器和所有内核都不同, 从而显著增强了对上下文信息的捕获。FR-ODConv 的输出可以用式 (10) 表示:

$$\text{Output}(y) = x \times (\alpha_{w1} \odot \alpha_{f1} \odot \alpha_{c1} \odot \alpha_{s1} \odot W_1 + \dots + \alpha_{wn} \odot \alpha_{fn} \odot \alpha_{cn} \odot \alpha_{sn} \odot W_n). \quad (10)$$

式中: α_{wi} , α_{fi} , α_{si} ($i=1, 2, \dots, n$) 为 4 种注意力; $\text{Output}(y)$ 为卷积核沿着空间中不同维度进行乘法运算的输出结果。

MCWE-YOLO 安全帽佩戴检测模型以“即插即用”的方式运用 FR-ODConv 替代 Head 中 3 个传统的普通静态卷积, 根据不同的输入图像数据自适应调整卷积核, 有针对性地提取安全帽的有效特征。该方法显著增强了模型的特征提取能力, 加强了网络的学习能力, 提高了模型的识别精度, FR-ODConv 结构如图 5 所示。

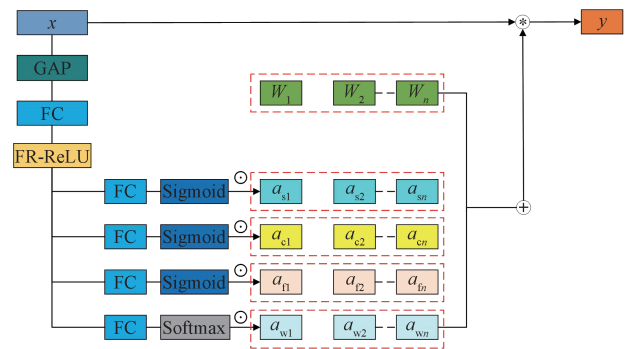


图 5 FR-ODConv 结构

Fig. 5 FR-ODConv structure

3.4 Res-EDH 解耦头

原始 YOLOv5 的检测头是通过合并和共享分类和回归分支来实现的, 而 YOLOX 的检测头则是将分类和回归分支进行解耦, 增加了两个额外的 3×3 的卷积层, 虽然提升了检测精度, 但一定程度上增加了网络延时。

MCWE-YOLO 模型在 Efficient 解耦头^[12] 的基础上提出一种 Res-EDH 解耦头结构, 替换原有的 Detect 结构, 并对其进行了精简设计, 同时考虑到相关算子表征能力与硬件计算成本之间的平衡, 使用混合通道策略重新采用更高效的解耦头结构, 同时降低

了延迟, 减少了解耦头中 3×3 卷积带来的额外延迟开销; 在 3 个 1×1 的卷积之前加入残差结构, 可有效降低梯度损失, 其结构如图 6 所示。

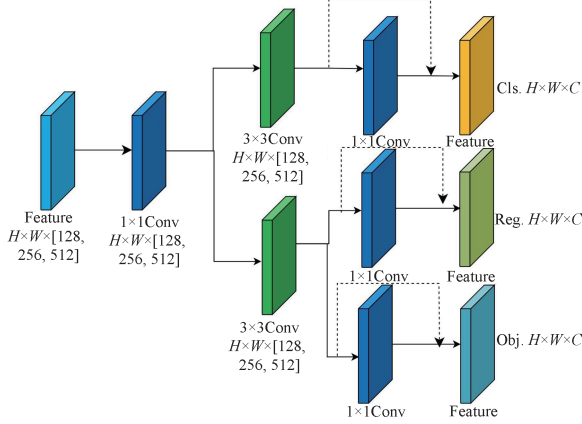


图 6 Res-EDH 结构

Fig. 6 Res-EDH structure

针对深度网络梯度消失过快的问题, 残差网络通过引入跨层连接结构, 使网络在反向传播时能避免梯度消失问题。残差结构如式 (11) 所示:

$$\frac{\partial Loss}{\partial x_1} = \frac{\partial Loss}{\partial x_L} \times \frac{\partial x_L}{\partial x_1} = \frac{\partial Loss}{\partial x_L} \times \left(1 + \frac{\partial}{\partial x_L} \sum_{i=1}^{L-1} F(x_i, W_i) \right). \quad (11)$$

式中: $\frac{\partial Loss}{\partial x_L}$ 为损失函数到达 L 的梯度; x_i 为所求函数偏导的自变量; 括号中的“1”代表短路机制 (identity x) 可以完整传播梯度, 而另一个残差梯度则需要通过带有权重的层。残差梯度不会全部为 -1, 即使很小, 由于“1”的存在梯度不会消失, 这有助于简化残差学习过程。

3.5 损失函数

目前的大部分研究都假设训练数据中的样本是高质量的, 旨在增强边界框损失的拟合能力, 然而, 目标检测的训练集包含低质量的样本。如果盲目地提高对低质量示例的边界框回归, 则会损害模型检测性能的提高。Focal-EIOU v1^[13] 被提出来解决这个问题, 但由于其聚焦机制是静态的, 非单调聚焦机制的潜力没有得到充分发挥。

YOLOv7-tiny 中的分类损失仍然使用二分类交叉熵损失 f_{BCEL} (BCEL, binary crossentropy loss), 回归部分使用双焦点损失 f_{DFL} (DFL, dual focal loss) 和边界框回归损失 f_{BBRL} (BBRL, bounding box regression loss)。完整的损失函数 f_{loss} 可以表示为式 (12):

$$f_{loss} = \lambda_1 f_{BCEL} + \lambda_2 f_{DFL} + \lambda_3 f_{BBRL}. \quad (12)$$

式中 λ 为不同损失函数所占权重。

预测类别损失本质上是交叉熵损失, 表达式为

$$f_{BCEL} = \text{weight}[class] \left(-x[class] + \log \left(\sum_j \exp(x[j]) \right) \right). \quad (13)$$

式中: $class$ 为类别的数量; $\text{weight}[class]$ 为每个类的权重; x 为 Sigmoid 激活之后的概率值。

DFL 是对 focal loss 损失函数的优化, 通过积分将离散的分类结果推广为连续的结果, 表达式如式 (14) 所示:

$$f_{DFL}(S_i, S_{i+1}) = -((y_{i+1} - y) \log(S_i) + (y - y_i) \log(S_{i+1})). \quad (14)$$

式中: y_i 、 y_{i+1} 为在连续标签 y 附近的来自左侧和右侧的值, 满足 $y_i < y < y_{i+1}$, $y = \sum_{i=0}^n P(y_i) y_i$; 其中, P 可以通过 softmax 层 $P(y_i)$ 即 S_i 来实现。

与 YOLOv7-tiny 中使用的 CIOW^[14] 损失不同, 模型引入 WiseIOU^[15-16] 损失函数作为边界框回归损失。在训练数据标记质量较低时, 损失函数结合动态非单调聚焦机制, 通过使用“离群值”来评估锚框质量, 从而避免几何因素 (如距离和纵横比) 对模型造成过度干扰。WiseIOU 原理见图 7。

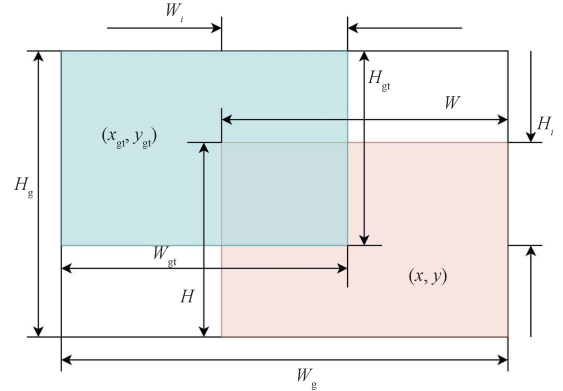


图 7 WiseIOU 原理

Fig. 7 WiseIOU principle

表达式如下:

$$f_{BBRL} = \left(1 - \frac{W_i H_i}{S_u} \right) \exp \left[\frac{(x_p - x_{gt})^2 + (y_p - y_{gt})^2}{(W_g^2 + H_g^2)^*} \right] \gamma, \quad (15)$$

$$\gamma = \beta / \delta \alpha^{\beta - \delta}. \quad (16)$$

式 (15) (16) 中: β 为预测框的异常程度, 异常程度越小意味着锚框的质量越高, 利用 β 构建非单调聚焦点数, 可以将较小的梯度增益分配给异常值较大的预测框, 从而有效减少低质量训练样本带来的有害梯度; α 和 δ 为超参数, 其他参数的含义如图 7 所示; x_p 和 y_p 为预测框的坐标值; x_{gt} 和 y_{gt} 为真实框的坐标

值,对应的 H 值和 W 值分别表示两框的宽度和高度; $S_u=wh+w_{gt}h_{gt}-W_iH_i$; * 代表将 w_g 和 w_h 从计算图中分离出来,目的是防止 WiseIOU 产生阻碍收敛的梯度。

4 实验结果与分析

4.1 实验环境配置及数据准备

模型基于 Pytorch 框架,使用 GPU 进行训练,实验环境配置具体如表 1 所示。

表 1 实验环境

Table 1 Experimental environment

名称	相关配置
操作系统	Windows 10 (x64)
CPU	i5-9400F CPU@GHz
GPU	NVIDIA GTX 2070 (8 GB)
内存	32 GB
Python 版本	3.7.12
CUDA	11.1

实验所用安全帽数据集为 GDUT-HWD,该数据集共有 3 174 张图像,实验分别以 7:2:1 的比例划分为训练集、验证集和测试集,共包含 18 893 个对象,包括建筑工地常见的大部分场景,代表性和广泛性较强。所有对象实例根据安全帽的颜色分为蓝色、白色、黄色、红色和无类别,不同颜色代表不同人员的工作属性,无类别表示未佩戴安全帽,共 5 个类别。设置迭代次数为 100,训练批次为 32,学习率为 0.01。

4.2 评价指标

评估网络性能的指标主要有准确率 (precision) P , 召回率 (recall, R), 平均准确度 (average precesion, AP), 平均准确率均值 (mean average precesion, mAP), mAP 越大表示模型整体精度越高。模型使用 $mAP@0.5$ 作为评估模型的准确性指标,代表将 IOU (intersection over union) 阈值取 50% 时的 mAP 值,各指标的计算方法如式 (17)~(20) 所示:

$$P = \frac{TP}{TP + FP}, \quad (17)$$

$$R = \frac{TP}{TP + FN}, \quad (18)$$

$$AP = \int_0^1 P(r) dr, \quad (19)$$

$$mAP = \frac{1}{n} \sum_{i=0}^n AP_i. \quad (20)$$

式 (17)~(20) 中: TP 为预测正确的正样本数量; FN 为预测错误的负样本数; FP 为预测错误的正样本数; $P(r)$ 为 r 点的检测精度。

4.3 消融实验及对比分析

针对 GDUT-HWD 安全帽数据集设计消融实验,实验均采用 4.1 节所述的实验参数,消融实验结果如表 2 所示。其中,模型 1 表示原始 YOLOv7-tiny 的实验结果,模型 2 表示改进 Mosaic 数据增强,模型 3 表示嵌入 CoordAttention 坐标注意力机制,模型 4 表示加入多维动态卷积 FR-ODConv 代替 head 中的静态卷积,模型 5 表示改进损失函数,模型 6 表示用 Res-EDH 解耦头替换原始的 Detect 解耦头。

表 2 本文算法在 GDUT-HWD 数据集的消融实验结果比较

Table 2 Algorithm compared with the discretion experiments of the GDUT-HWD dataset

模型	改进点	blue	white	yellow	red	none	$mAP@0.5$
模型 1	YOLOv7-tiny	84.2	86.3	91.5	83.1	83.4	85.7
模型 2	YOLOv7-tiny+Mosaic-9	86.5	87.7	92.5	82.1	84.6	86.5
模型 3	YOLOv7-tiny+CoordAttention	85.7	87.7	91.5	84.5	84.7	86.8
模型 4	YOLOv7-tiny+FR-ODConv	84.8	87.7	91.8	83.0	83.6	86.2
模型 5	YOLOv7-tiny+WiseIOU	84.7	88.0	92.2	84.4	84.3	86.7
模型 6	YOLOv7-tiny+Res-EDH	84.3	87.9	92.6	83.5	84.5	86.5
改进模型	本文算法	85.6	87.9	93.8	84.1	84.9	87.3

从表 2 中数据可以看出,与原始 YOLOv7-tiny 算法相比,每修改一部分都有不同程度的提高,将它们结合起来,最终改进算法的 mAP 提高了 1.6%。改进后的 MCWE-YOLO 算法训练结果如图 8 所示。

通过在模型中嵌入 CoordAttention 坐标注意力机制,增强网络模型对深层次信息的关注,与

CBAM^[17]、SimAM^[18] 等常规注意力机制相比, mAP 提升更高,如表 3 所示。

为验证所改进模块的有效性,采用相同的验证集,做了相关对比实验,将其与目前主流的安全帽目标检测算法进行对比,结果如表 4 所示,表明改进的实验方法获得了最优的 mAP 值结果。

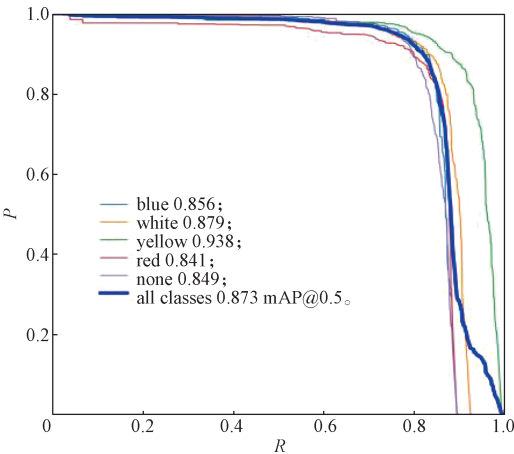


图 8 MCWE-YOLO 方法训练结果
Fig. 8 MCWE-YOLO method training results

表 3 注意力机制对比结果
Table 3 Comparison of attention mechanisms

算 法	<i>mAP</i> /%
YOLOv7-tiny+CBAM	86.5
YOLOv7-tiny+GAM ^[19]	86.2
YOLOv7-tiny+SimAM	86.4
YOLOv7-tiny+ECA ^[20]	86.3
YOLOv7-tiny+CoordAttention	86.8

表 4 不同算法在 GDUT-HWD 数据集的检测结果比较
Table 4 Different algorithms compared in the detection results of GDUT-HWD dataset %

模 型	<i>P</i>	<i>R</i>	<i>mAP</i> @0.5
YOLOv3-tiny	88.7	68.0	74.6
YOLOv5s	90.6	79.7	85.0
YOLOv7	85.8	76.4	83.3
YOLOv7-tiny	88.8	81.2	85.7
YOLOv8n	89.9	78.3	84.4
YOLOv8s	89.7	77.9	85.2
本文算法	90.7	82.6	87.3

为更加直观地进行评价, 模型使用 GDUT-HWD 安全帽数据集中部分图片进行测试。对比图 9a 和 9b 可以看出, YOLOv7-tiny 在小目标场景下存在漏检, 没有检测到图中的黄色安全帽; 对比图 9c 和 9d 可以看出, YOLOv7-tiny 在遮挡场景下存在漏检, 没有检测到图中被遮挡的蓝色安全帽; 对比图 9e 和 9f 可以看出, YOLOv7-tiny 在反光场景下存在误检, 将图中白色的安全帽误检为黄色安全帽; 对比图 9g 和 9h 可以看出, YOLOv7-tiny 在黑暗场景下存在漏检, 没有将最上方未戴安全帽的人员检测到。可以看出, 改进后的 MCWE-YOLO 模型检测准确率更高, 整体检测效果的置信度更高, 检测性能明显提升。



a) YOLOv7-tiny 小目标检测 b) MCWE-YOLO 小目标检测



c) YOLOv7-tiny 遮挡下检测 d) MCWE-YOLO 遮挡下检测



e) YOLOv7-tiny 反光下检测 f) MCWE-YOLO 反光下检测



g) YOLOv7-tiny 黑暗下检测 h) MCWE-YOLO 黑暗下检测

图 9 改进前后效果对比图

Fig. 9 Comparison of effect before and after improvement

5 结语

为解决建筑工地复杂条件下安全帽佩戴检测存在的模型识别精度低, 易出现漏检以及误检等问题, 提出了一种 MCWE-YOLO 安全帽检测算法。实验结果表明, 所提出的 MCWE-YOLO 算法相比于原始的 YOLOv7-tiny 算法在 GDUT-HWD 数据集上 *mAP* 提升了 1.6%, *P* 提高了 1.9%, 表明了改进算法在安全帽佩戴检测中的有效性, 实现了建筑工地安全帽佩戴的自动化检测。

参考文献:

[1] 孙国栋, 李 超, 张 航. 融合自注意力机制的安全帽佩戴检测方法 [J]. 计算机工程与应用, 2022, 58(20): 300-304.

- SUN Guodong, LI Chao, ZHANG Hang. Safety Helmet Wearing Detection Method Fused with Self-Attention Mechanism[J]. Computer Engineering and Applications, 2022, 58(20): 300–304.
- [2] 吕宗喆, 徐 慧, 杨 骁, 等. 面向小目标的 YOLOv5 安全帽检测算法 [J]. 计算机应用, 2023, 43(6): 1943–1949.
- LÜ Zongzhe, XU Hui, YANG Xiao, et al. Small Object Detection Algorithm of YOLOv5 for Safety Helmet[J]. Journal of Computer Applications, 2023, 43(6): 1943–1949.
- [3] 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.
- [4] 赵 敏, 杨国亮, 王吉祥, 等. 改进 YOLOv7-tiny 的安全帽实时检测算法 [J]. 无线电工程, 2023, 53(8): 1741–1749.
- ZHAO Min, YANG Guoliang, WANG Jixiang, et al. Improved YOLOv7-Tiny Algorithm for Safety Helmet Real-Time Detection[J]. Radio Engineering, 2023, 53(8): 1741–1749.
- [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[EB/OL]. [2023–08–27]. <http://arxiv.org/abs/2207.02696>.
- [6] 刘浩翰, 樊一鸣, 贺怀清, 等. 改进 YOLOv7-tiny 的目标检测轻量化模型 [J]. 计算机工程与应用, 2023, 59(14): 166–175.
- LIU Haohan, FAN Yiming, HE Huaqing, et al. Improved YOLOv7-Tiny's Object Detection Lightweight Model[J]. Computer Engineering and Applications, 2023, 59(14): 166–175.
- [7] CHEN Z H, ZHANG F, LIU H B, et al. Real-Time Detection Algorithm of Helmet and Reflective Vest Based on Improved YOLOv5[J]. Journal of Real-Time Image Processing, 2023, 20(1): 4.
- [8] HOU Q B, ZHOU D Q, FENG J S. Coordinate Attention for Efficient Mobile Network Design[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Nashville: IEEE, 2021: 13708–13717.
- [9] 程德强, 陈 杰, 寇旗旗, 等. 融合层次特征和注意力机制的轻量化矿井图像超分辨率重建方法 [J]. 仪器仪表学报, 2022, 43(8): 73–84.
- CHENG Deqiang, CHEN Jie, KOU Qiqi, et al. Lightweight Super-Resolution Reconstruction Method Based on Hierarchical Features Fusion and Attention Mechanism for Mine Image[J]. Chinese Journal of Scientific Instrument, 2022, 43(8): 73–84.
- [10] QIAN J J, LIN J, BAI D, et al. Omni-Dimensional Dynamic Convolution Meets Bottleneck Transformer: A Novel Improved High Accuracy Forest Fire Smoke Detection Model[J]. Forests, 2023, 14(4): 838.
- [11] CHENG S X, ZHU Y S, WU S H. Deep Learning Based Efficient Ship Detection from Drone-Captured Images for Maritime Surveillance[J]. Ocean Engineering, 2023, 285: 115440.
- [12] LI C Y, LI L L, JIANG H L, et al. YOLOv6: A Single-Stage Object Detection Framework for Industrial Applications[EB/OL]. [2023–08–27]. <http://arxiv.org/abs/2209.02976>.
- [13] ZHANG Y F, REN W Q, ZHANG Z, et al. Focal and Efficient IOU Loss for Accurate Bounding Box Regression[J]. Neurocomputing, 2022, 506(C): 146–157.
- [14] 赵梓杉, 桑海峰. 基于改进的 YOLOv5 的交通锥标检测系统 [J]. 电子测量与仪器学报, 2023, 37(2): 56–64.
- ZHAO Zishan, SANG Haifeng. Traffic Cone Detection System Based on Improved YOLOv5[J]. Journal of Electronic Measurement and Instrumentation, 2023, 37(2): 56–64.
- [15] TONG Z J, CHEN Y H, XU Z W, et al. Wise-IoU: Bounding Box Regression Loss with Dynamic Focusing Mechanism[EB/OL]. [2023–08–27]. <http://arxiv.org/abs/2301.10051>.
- [16] LI Y T, FAN Q S, HUANG H S, et al. A Modified YOLOv8 Detection Network for UAV Aerial Image Recognition[J]. Drones, 2023, 7(5): 304.
- [17] LI R, WU Y P. Improved YOLO v5 Wheat Ear Detection Algorithm Based on Attention Mechanism[J]. Electronics, 2022, 11(11): 1673.
- [18] HAO K, CHEN G K, ZHAO L, et al. An Insulator Defect Detection Model in Aerial Images Based on Multiscale Feature Pyramid Network[J]. IEEE Transactions on Instrumentation and Measurement, 2022, 71: 3522412.
- [19] WANG C Q, ZHANG B, CAO Y, et al. Mask Detection Method Based on YOLO-GBC Network[J]. Electronics, 2023, 12(2): 408.
- [20] ZHANG M, YIN L J. Solar Cell Surface Defect Detection Based on Improved YOLO v5[J]. IEEE Access, 2022, 10: 80804–80815.

(责任编辑: 申 剑)