

基于改进YOLO的安全帽佩戴检测算法研究

刘 聪

五邑大学, 智能制造学部, 广东 江门

收稿日期: 2023年11月26日; 录用日期: 2023年12月22日; 发布日期: 2023年12月30日

摘 要

高危场景下的安全帽佩戴检测任务是保障施工人员生命财产安全的重要一环。本文针对复杂场景下传统安全帽检测中小目标检测性能不佳、模型参数量大等问题, 提出一种基于改进YOLOv7-Tiny的安全帽检测方法。该方法设计一种针对小目标检测的特征提取网络, 从而加强网络对小目标的识别能力; 引入轻量化上采样算子CARAFE, 改进上采样过程从而提高特征融合质量; 使用K-means++聚类算法对数据集中目标的锚框进行重新聚类, 得到适应安全帽检测的锚框尺寸; 并引入CA注意力机制, 让网络更加关注有用的坐标信息, 加强网络对小目标的识别能力。实验结果表明, 在公开数据集SHWD上, 改进后的算法相比YOLOv4-Tiny、YOLOX-Tiny等传统算法精度分别提升了14.1%和2.8%, 相比原算法在参数量减少了31.7%的同时精度提升了1.5%, 满足实际场景中对安全帽的检测要求。

关键词

安全帽佩戴检测, 上采样算子, 聚类算法, 注意力机制

Research on Helmet Wear Detection Algorithm Based on Improved YOLO

Cong Liu

Department of Intelligent Manufacturing, Wuyi University, Jiangmen Guangdong

Received: Nov. 26th, 2023; accepted: Dec. 22nd, 2023; published: Dec. 30th, 2023

Abstract

The task of helmet wear detection in high-risk scenarios is an important part of safeguarding the lives and properties of construction workers. In this paper, we propose a helmet detection method based on an improved YOLOv7-Tiny, addressing challenges related to the poor performance of small target detection and the large number of model parameters associated with traditional helmet detection in complex scenes. Our method incorporates a feature extraction network designed

specifically for small target detection, enhancing the network's capability to recognize small targets. Additionally, we introduce a lightweight up-sampling operator, CARAFE, to optimize the up-sampling process, thereby improving the quality of feature fusion. To adapt anchor frames to helmet detection, we employ the K-means++ clustering algorithm to recluster the anchor frames of targets in the dataset. Furthermore, we integrate the CA attention mechanism, allowing the network to prioritize useful coordinate information and reinforce its ability to recognize small targets. Experimental results on the SHWD public dataset demonstrate that the improved algorithm enhances mAP by 14.1% and 2.8% compared to traditional algorithms like YOLOv4-Tiny and YOLOX-Tiny, respectively. Moreover, it improves mAP by 1.5% compared to the baseline while reducing the number of parameters by 31.7%. These results meet the requirements for safety helmet detection in real scenarios.

Keywords

Safety Helmet Wearing Detection, Up-Sampling Operator, Clustering Algorithm, Attention Mechanism

Copyright © 2023 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

近年来生产安全事故态势来看,我国房屋市政工程安全形势不容乐观,建筑行业安全生产监督管理工作需进一步加强[1]。每年因未正确佩戴安全帽而导致的安全事故数不胜数,给人们的生命财产安全带来了严重的损害。为此,工人在施工期间必须佩戴安全帽,检查工人是否按要求佩戴安全帽工作成为生产安全管理工作中的一项重要环节[2]。但传统人工监管随监控时长增加,监控范围扩大,容易产生视觉疲劳,继而导致误判,既耗费大量人力、物力资源,也无法满足目前安全管理要求[3]。因此,采用高效的目标检测方法进行自动化安全帽检测十分必要。

国内外广大学者对安全帽检测技术展开了广泛的研究。冯国臣等人[4]使用机器视觉方法进行安全自动识别研究;Le Vanbang [5]等人通过 HOG 特征提取来检测人体,接着采用颜色直方图识别安全帽;胡恬[6]在研究小波变换和 BP 神经网络的基础上,进行了人脸识别的研究;刘晓慧等人[7]利用安全帽与人头发颜色的色差进而判断工人是否佩戴安全帽。然而,以上传统目标检测方法通过人工设计特征实现检测任务,效率较为低下,并且存在模型复杂、检测精度不高等问题。

随着卷积神经网络(Convolutional Neural Network, CNN)的发展,基于传统机器学习的目标检测算法正在慢慢被基于深度学习的卷积神经网络算法替代。Redmon J 等提出的 YOLO 系列算法[8] [9] [10] [11]使深度卷积神经网络在目标识别精度和速度上都有了较大提升。在此基础上,徐印赟等[12]在传统 YOLO 算法的基础上通过设计特征融合模块,强化不同尺度间的信息融合从而降低误检率。李振华等[13]在 YOLOv3 中引入自注意力,并对聚类算法进行优化,提升了锚框质量。孙备等[14]在 SSD 算法上通过对浅层网络进行多尺度卷积融合,提高对小目标的检测性能。丁田等[15]在 YOLOX 网络基础上引入跨通道交互的注意力机制,并优化损失函数,提高安全帽检测精度的同时保持较好的检测速度。

但以上现有的安全帽检测方法仍然容易受到复杂多变的环境影响,导致检测准确率较低,此外模型参数量普遍较大。本文提出改进 YOLOv7-Tiny 的安全帽检测方法,针对复杂场景下小目标检测效果不佳、模型参数量大等问题,对特征融合网络进行改进,并引入了注意力机制,引导模型关注小目标信息,并

对锚框重新聚类，使其适应安全帽检测任务尺寸。实验结果表明：改进的模型具有环境适应性好、检测准确率高、模型参数量小等特点，满足实际复杂场景下的安全帽检测任务。

2. YOLOv7-Tiny 算法简介

YOLOv7 [16]是 Alexey Bochkovskiy 团队于 2022 年提出的目标检测模型。作为当今 YOLO 系列中的基本模型，其检测速度和精度也十分优秀。为保证模型大小与检测精度之间的平衡，选择使用 YOLOv7-Tiny 作为基础网络进行改进。其网络架构如图 1 所示。

YOLOv7-Tiny 网络主要由 Backbone、Neck、Head 三部分组成。Backbone 部分主要由 CBL、ELAN 模块以及 MP 下采样模块构建而成。其中 Backbone 部分通过多个 CBL 模块对输入图像进行卷积下采样，再经过若干 ELAN 和 MP 下采样模块进行特征提取。Neck 部分通过 SPPCSPC 模块对不同尺度的特征进行融合，同时该网络采用了 YOLOv5 中使用的 PAN 结构，加强特征提取网络中的特征融合。Head 通过 CBL 卷积检测大、中、小目标。

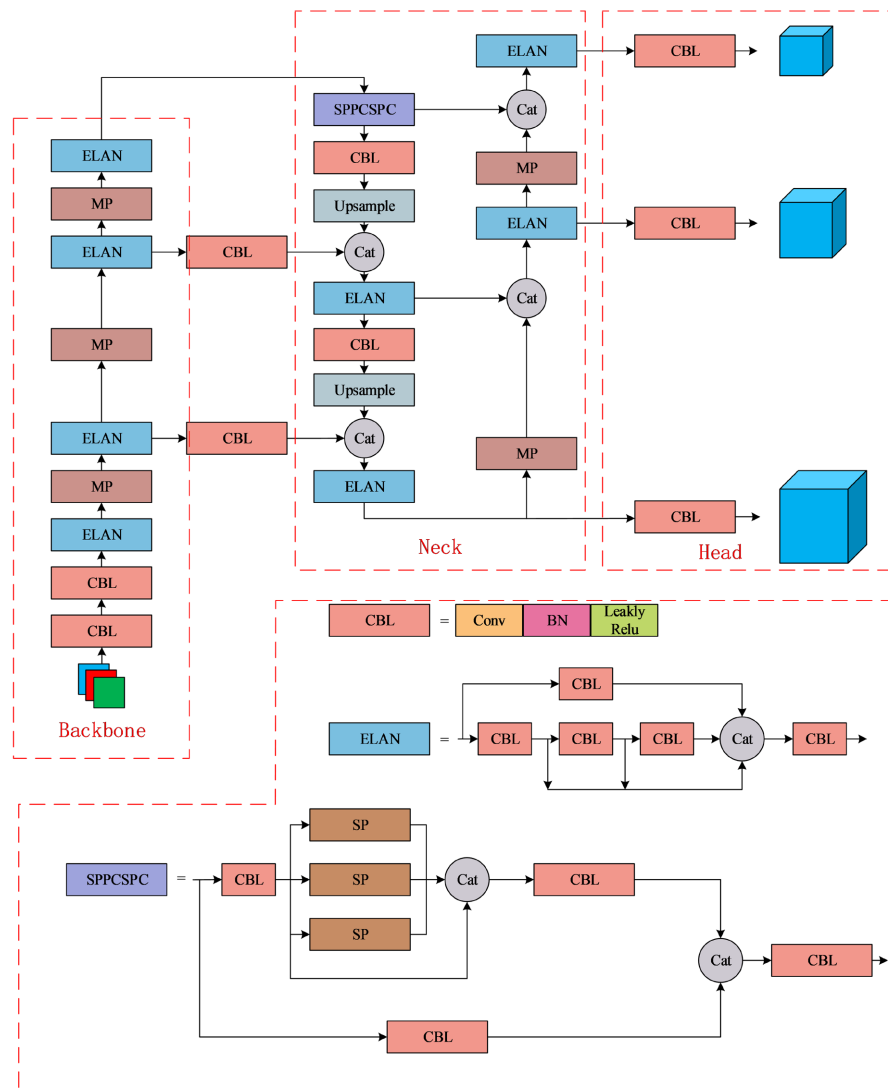


Figure 1. YOLOv7-Tiny network structure diagram

图 1. YOLOv7-Tiny 网络结构图

3. YOLOv7-Tiny 算法的改进

3.1. 引入全局注意力机制

在安全帽检测任务中，安全帽和头部在图像中像素占比较小，同时物体间普遍存在遮挡等问题。此外随着网络加深，细节信息丢失严重，网络中识别小目标等的能力较弱而造成漏检、误检等问题。为改善网络对小目标的识别能力，在骨干网络的最后方和检测头前引入坐标注意力[17]模块，在几乎不带来计算开销的前提下，让网络更加关注空间上的坐标信息，从而更精准的识别被检测的目标，提升网络性能。

坐标注意力通过将横向和纵向的坐标信息通过编码融合进通道注意力中，使其能更加关注到物体的坐标信息，从而提升网络的检测性能。其具体操作主要分为两个步骤：坐标信息编码和坐标注意力生成，结构如图 2 所示。

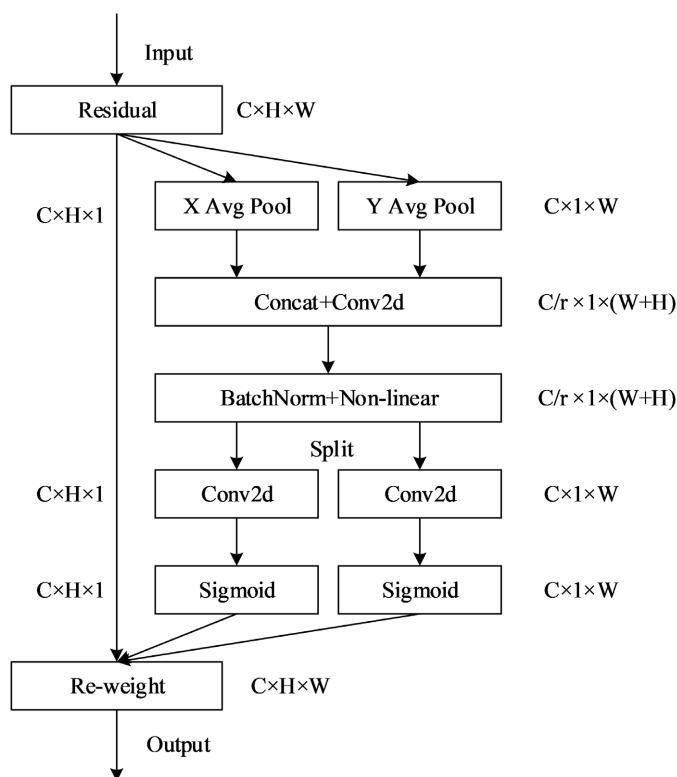


Figure 2. CA coordinate attention

图 2. CA 坐标注意力

首先，对全局位置信息进行编码。对任意给定输入 X ，使用尺寸为 $(H,1)$ 和 $(1,W)$ 的池化核来编码水平和垂直方向特征，其第 c 维特征在垂直方向高度为 h 时和水平方向宽度为 w 时的输出分别如公式(1)、(2)所示：

$$Z_c^h(h) = \frac{1}{W} \sum_{0 < i < w} X_c(h, i) \quad (1)$$

$$Z_c^w(w) = \frac{1}{H} \sum_{0 < i < H} X_c(j, w) \quad (2)$$

上述公式沿水平和垂直两方向集成特征，输出一对方向确定的特征图，使得网络能关注到某一方向上的长距离关系，从而建立坐标关系，使网络能重点关注空间位置信息。再对上述变换后的坐标信息进行注意力生成，将公式(1)、(2)的输出进行 Concatenate 操作，再对其使用卷积函数 F_1 进行变换，最后再经过非线性激活函数 δ 得到包含水平方向和垂直方向的特征图输出记为 f ，其表达式如公式(3)所示：

$$f = \delta\left(F_1\left(\left[Z^h, Z^w\right]\right)\right) \quad (3)$$

随后将 f 分解为两个独立的特征 f^h 和 f^w ，再分别通过 1×1 卷积和 Sigmoid 激活函数进行特征转换，使输出维度与给定输入 X 一致，其输出表达式如公式(4)、(5)所示：

$$g^h = \sigma\left(F_h\left(f^h\right)\right) \quad (4)$$

$$g^w = \sigma\left(F_w\left(f^w\right)\right) \quad (5)$$

将两个独立方向的输出 g^h 和 g^w 合并为权重矩阵，最终 CA 注意力模块的输出可以写成如公式(6)所示：

$$y_c(i, j) = x_c(i, j) \times g_c^h(i) \times g_c^w(j) \quad (6)$$

3.2. 特征提取网络的改进

YOLOv7-Tiny 原网络在输入图像尺寸在(640,640)的情况下，进行连续多次 2 倍下采样得到(80,80)、(40,40)、(20,20)三种尺度的特征图。然而，原网络在训练的过程中，随着不断进行卷积和下采样，抽象的语义信息不断增强，而对小目标而言，细节和纹理信息也在不断减弱。为缓解网络加深导致的细节纹理信息丢失，提升网络对小目标的检测性能，在特征提取网络中，增加针对小目标的特征提取层，其结构如图 3 所示。

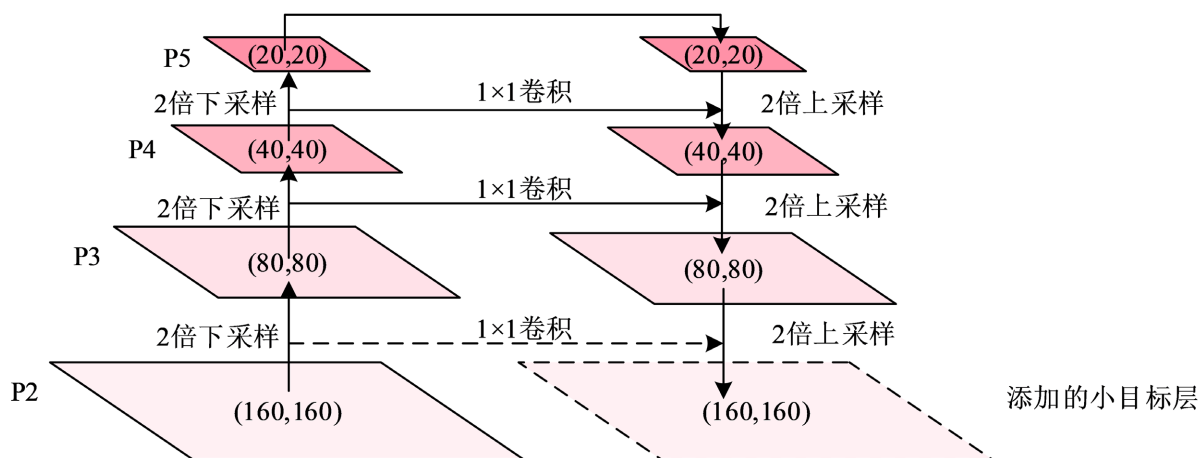


Figure 3. Adding the feature extraction network of the small target feature extraction layer

图 3. 添加小目标特征提取层的特征提取网络图

图中虚线部分为增加的针对小目标的特征提取层，通过对浅层特征层中的特征进行特征融合，加强网络对小目标的识别能力。同时，为减少参数量，删除用于预测大目标的预测层和预测头，由于在安全帽检测任务中，检测对象基本为中小目标，大目标在其中占比较少，因此删减用于预测大目标的检测层和检测头对于安全帽检测性能影响较小。其改进的网络结构如下图 4 所示。

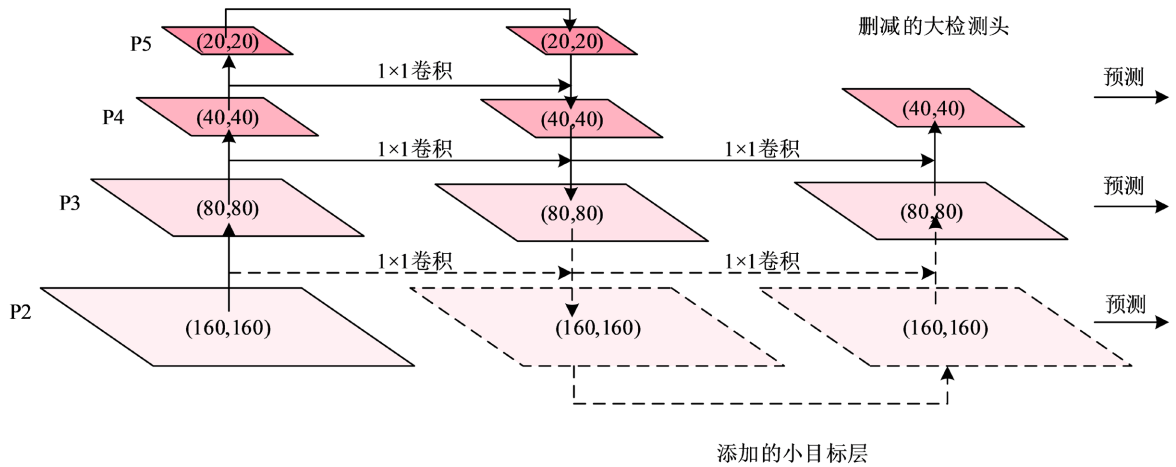


Figure 4. Feature extraction network of deleted large target prediction head
图 4. 删减大目标预测头的特征提取网络图

3.3. 引入 CARAFE 上采样算子

YOLOv7-Tiny 采用最邻近上采样方式进行上采样，该采样方式仅通过空间位置决定上采样核，并没有充分利用特征图中的抽象语义信息，且通常感知域都较小。针对最邻近上采样方式的不足，引入轻量的上采样算子[18] Content-Aware Re Assembly of Features，简称 CARAFE 算子。

CARAFE 主要由上采样核预测模块和特征重组模块组成，其网络结构简图如图 5 所示：

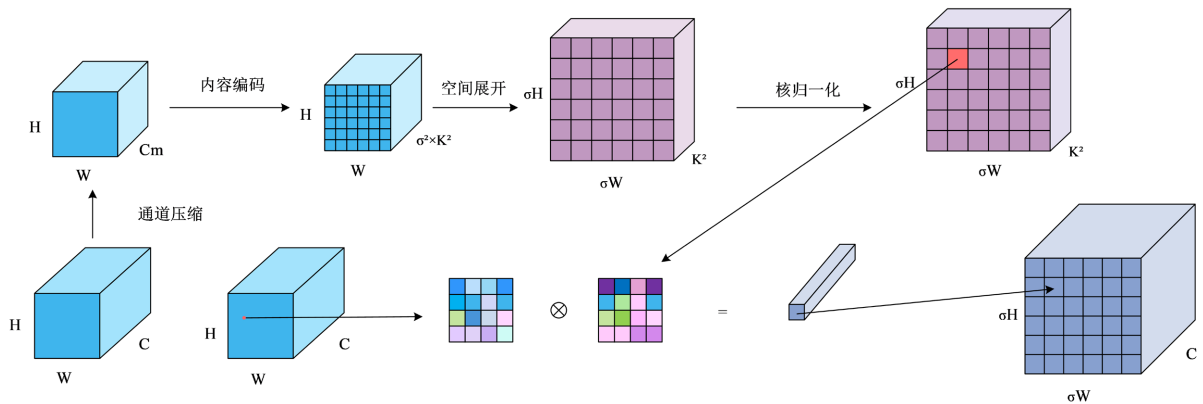


Figure 5. CARAFE up-sampled network structure diagram
图 5. CARAFE 上采样网络结构图

尺寸为 $H \times W \times C$ 的输入特征图，若上采样倍数为 σ ，输出得到尺寸为 $\sigma H \times \sigma W \times C$ 的输出特征图。其主要过程为：首先用一个 1×1 卷积将通道进行降维到 C_m ，接着对输入特征图内容进行编码，假设上采样核的尺寸为 $K \times K$ ，利用一个卷积层来预测上述采样核，输入通道数为 C_m ，输出通道数为 $\sigma^2 K^2$ ，然后再将通道维度展开到空间维度，得到形状为 $\sigma H \times \sigma W \times K^2$ 的上采样核。再将上述上采样核的各项权重利用 Softmax 进行归一化。最后，将上述输出特征图中的每个像素点的位置映射回输入特征图中，找出以该映射点为中心点的 $K \times K$ 区域，与该点上采样核的通道维度展开成 $K \times K$ 的空间区域作点积，最后得到尺寸为 $\sigma H \times \sigma W \times C$ 的输出特征图。

在原网络中利用 CARAFE 算子替换传统的最邻近上采样算子，让网络在特征融合过程中拥有更大感

受野，同时根据输入特征图自适应预测采样核，根据输入特征图进行高质量上采样，保证特征信息的完整性。

改进的 YOLOv7-Tiny 网络如下图 6 所示：

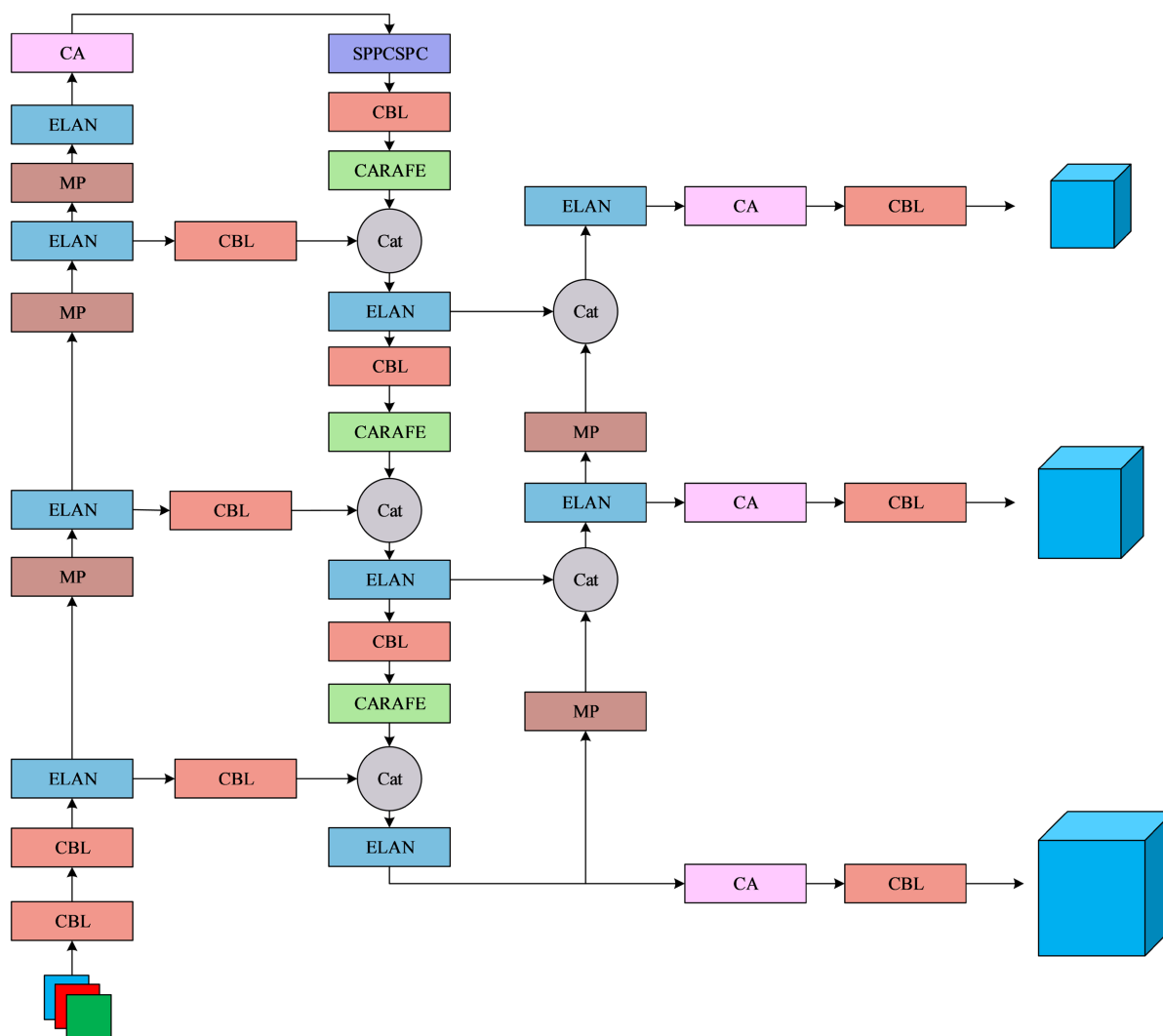


Figure 6. Structural diagram of the improved YOLOv7-Tiny network

图 6. 改进的 YOLOv7-Tiny 网络结构图

3.4. K-Means++ 聚类锚框

YOLOv7-Tiny 默认设置了针对 COCO 数据集的 9 个 anchor 框，按照大小顺序依次为[10,13], [16,30], [33,23], [30,61], [62,45], [59,119], [116,90], [156,198], [373,326]分别用来检测小、中、大目标。然而不同数据集的目标尺寸分布不一样，在安全帽检测任务中，针对 COCO 数据集预设的 9 个锚框并不符合安全帽检测场景。因此，需要针对 SHWD 数据集对锚框进行重新聚类分析。为了使网络更好的学习安全帽的大小尺寸信息，本文采用 K-Means++ 算法对锚框进行重新聚类。最终得到 9 个针对安全帽尺寸的锚框[9,10], [11,13], [16,19], [26,28], [40,44], [60,67], [97,110], [174,198], [396,332]，最后选定用这 9 个锚框替换预设锚框进行安全帽检测训练。

4. 实验及结果分析

4.1. 实验环境与参数

本文实验的操作系统为 Windows10, GPU 型号为 NVIDIA GeForce RTX 3090 (24G 显存), 编程语言选择 Python, 采用 Pytorch 框架, CUDA 版本为 11.4, 实验超参数的设置如表 1 所示。

Table 1. Experimental hyperparameter settings

表 1. 实验超参数设置

参数名	参数值
初始学习率(initial learning rate)	0.01
学习率衰减因子(learning rate factor)	0.1
动量(momentum)	0.937
批量大小(batch_size)	16
迭代轮数(epoch)	300
权重衰减(weight_decay)	0.0005

4.2. 实验数据集

本文采用公开数据集 SHWD (Safety Helmet Wearing Dataset)进行安全帽佩戴检测研究, 该数据集中共有 9044 个人佩戴安全帽的目标和 111,514 个未佩戴安全帽的头部目标。数据集中一共包含 7581 张图片, 其中将 5457 张图片作为训练集, 607 张图片作为验证集, 1517 张图片作为测试集。

4.3. 评价指标

本文使用准确率(Precision, P)、召回率(Recall, R)、平均精确度(Average Precision, AP)、平均精确度均值(mean Average Precision, mAP)等指标评价模型的检测精度, 使用参数量(Parameter, $Params$)指标评价模型的大小。

其中准确率 P 和召回率 R 计算公式如下(7)、(8)所示:

$$P = \frac{TP}{TP + FP} \times 100\% \quad (7)$$

$$R = \frac{TP}{TP + FN} \times 100\% \quad (8)$$

式中 TP 为预测正确的正样本数, FP 为预测错误的正样本数, FN 为预测错误的负样本数。

平均精确度 AP 和平均精确度均值 mAP 计算公式如下所示:

$$AP = \int_0^1 P(R) dR \quad (9)$$

$$mAP = \frac{\sum_{i=0}^N AP(i)}{N} \quad (10)$$

式中 N 为总类别数, AP 用于衡量模型对于某一类别的平均精确度, mAP 为所有类别的 AP 的平均值。模型在实验结束后会根据预测结果得出准确率 P 和召回率 R , 将此两种指标作为横纵坐标可以绘制出 P - R 曲线, 其曲线下面积就是模型对某一类的预测平均精确度 AP , 所有类别 AP 的均值就是 mAP , mAP 越大, 模型的检测性能越好。

模型的参数量计算公式如下(11)所示:

$$Params = (K \times K \times C_i) \times C_o \quad (11)$$

$Params$ 表示一个卷积层不考虑偏重情况下的参数量, 其中 K 为卷积核的尺寸, C_i 和 C_o 为输入输出通道数。

4.4. 实验结果

4.4.1. 消融实验

为分析每一部分模块的改进对于模型整体系统的性能影响, 同时为了保证实验的科学性, 进行消融实验, 将实验一共分为 6 组。编号 1 表示使用原网络进行检测; 编号 2 表示引入坐标注意力模块; 编号 3 表示改进特征提取网络; 编号 4 表示利用 K-Means++ 算法对锚框进行重新聚类; 编号 5 表示引入 CARAFE 上采样算子; 编号 6 表示将上述改进均添加得到的最终模型, 实验结果如表 2 所示。

Table 2. Results of ablation study

表 2. 消融实验结果

编号	添加注意力机制	改进特征提取网络	重新聚类锚框	引入上采样算子	mAP@0.5/%
1	×	×	×	×	93.4
2	√	×	×	×	93.6
3	×	√	×	×	94.1
4	×	×	√	×	94.3
5	×	×	×	√	93.9
6	√	√	√	√	94.9

从表中数据可以看出, 引入注意力机制, 模型检测精度上升 0.2%, 改进特征提取网络后, 模型的检测精度相比原网络上升了 0.7%, 由此证明提出的针对小目标的特征提取网络具有较好的检测性能, 重新聚类锚框后, 算法精度上升 0.9%, 改进上采样算子后, 算法精度提升 0.5%。实验证明各个模块对模型整体性能起良好的改进作用, 通过改进后的算法拥有高达 94.9% 的检测精度, 其优良的检测性能满足安全帽检测任务中的检测需求。

4.4.2. 对比实验

为了进一步验证改进后算法的有效性和实用性, 将改进后的算法与目前主流的目标检测算法 SSD、YOLOv4-Tiny、YOLOX-Tiny、YOLOv5s、及原网络 YOLOv7-Tiny 进行对比实验。其结果如表 3 所示。

Table 3. Results of comparison to others methods

表 3. 对比实验结果

模型	参数量/M	准确率/%	召回率/%	mAP@0.5/%
SSD	23.9	93.5	67.1	89.3
YOLOv4-Tiny	5.9	82.8	77.7	80.8
YOLOX-Tiny	5.0	91.9	87.9	92.1
YOLOv5s	7.0	92.1	89.4	93.9
YOLOv7-Tiny	6.0	92.8	87.9	93.4
Ours	4.1	92.2	89.0	94.9

由上表 3 可知, 本文提出的改进 YOLOv7-Tiny 算法在检测性能和模型参数量上相比当前一众主流算法有着较为明显的优势。在保证高检测精度的同时有着较小的参数量, 满足实际应用场景中对安全帽检测的需求。为了更直观地验证改进后算法的优越性, 将原 YOLOv7-Tiny 算法(左)与改进的算法(右)检测效果分别进行可视化分析, 检测效果如图 7~9 所示。通过对比检测结果可以看出, 本文提出的改进的 YOLOv7-Tiny 算法在远距离的小目标上以及遮挡等不利条件下, 检测性能均优于原算法, 减少了漏检情况, 提升了算法的精度。



Figure 7. Target detection of the occlusion
图 7. 遮挡目标检测



Figure 8. Small target detection at long distances
图 8. 远距离小目标检测



Figure 9. Blurred target detection
图 9. 模糊目标检测

5. 结束语

在建筑施工地等高危环境下, 由于人员密集、物体遮挡目标、像素点较小等一系列因素的影响, 进

行安全帽检测任务较为困难, 普遍存在小目标检测困难或检测模型参数量较大等问题。本文针对以上问题, 提出了基于 YOLOv7-Tiny 的改进模型, 通过改善特征提取网络, 加强对小目标的特征提取能力, 并引入了坐标注意力机制, 让网络更加关注坐标信息, 此外, 引入 CARAFE 上采样算子, 优化了网络上采样的质量。实验结果表明, 改进的 YOLOv7-tiny 算法相比原网络, 在参数量下降 31.7% 的同时, 检测精度高达 94.9%, 相比原网络提升了 1.5%, 满足复杂场景下对安全帽检测的需求。

参考文献

- [1] 唐凯, 陈陆, 张洲境, 张善勇. 我国建筑施工行业生产安全事故统计分析对策[J]. 建筑安全, 2020, 35(9): 40-43.
- [2] 张仕廉, 王黎明, 叶贵, 王玉合. 建筑安全全要素投入研究: 以重庆市为例[J]. 中国安全科学学报, 2018, 28(3): 161-166.
- [3] 王玲敏, 段军, 辛立伟. 引入注意力机制的 YOLOv5 安全帽佩戴检测方法[J]. 计算机工程与应用, 2022, 58(9): 303-312.
- [4] 冯国臣, 陈艳艳, 陈宁, 李鑫, 宋程程. 基于机器视觉的安全帽自动识别技术研究[J]. 机械设计与制造工程, 2015, 44(10): 39-42.
- [5] Vanbang, L.E., 朱煜, Anhtu, N. 深度图像手势分割及 HOG-SVM 手势识别方法研究[J]. 计算机应用与软件, 2016, 33(12): 122-126.
- [6] 胡恬. 利用几何分析法和 BP 神经网络进行人脸识别的研究[J]. 计算机工程与设计, 2002(9): 18-21.
- [7] 刘晓慧, 叶西宁. 肤色检测和 Hu 矩在安全帽识别中的应用[J]. 华东理工大学学报(自然科学版), 2014, 40(3): 365-370.
- [8] Redmon, J., Divvala, S., Girshick, R., et al. (2016) You Only Look Once: Unified, Real-Time Object Detection. 2016 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, 27-30 June 2016, 779-788. <https://doi.org/10.1109/CVPR.2016.91>
- [9] Redmon, J. and Farhadi, A. (2017) YOLO9000: Better, Faster, Stronger. 2017 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, 21-26 July 2017, 7263-7271. <https://doi.org/10.1109/CVPR.2017.690>
- [10] Redmon, J. and Farhadi, A. (2018) Yolov3: An Incremental Improvement. <https://doi.org/10.48550/arXiv.1804.02767>
- [11] Bochkovskiy, A., Wang, C.Y. and Liao, Y.M. (2020) Yolov4: Optimal Speed and Accuracy of Object Detection. <https://doi.org/10.48550/arXiv.2004.10934>
- [12] 徐印赞, 江明, 李云飞, 等. 基于改进 YOLO 及 NMS 的水果目标检测[J]. 电子测量与仪器学报, 2022, 36(4): 114-123.
- [13] 李振华, 张雷. 改进 YOLOv3 的安全帽佩戴检测方法[J]. 国外电子测量技术, 2022, 41(12): 148-155.
- [14] 孙备, 左震, 吴鹏, 等. 面向无人艇环境感知的改进型 SSD 目标检测方法[J]. 仪器仪表学报, 2021, 42(9): 52-61.
- [15] 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. <https://doi.org/10.19651/j.cnki.emt.2209425>
- [16] Wang, C.Y., Bochkovskiy, A. and Liao, H.Y.M. (2022) YOLOv7: Trainable Bag-Of-Freebies Sets New State-Of-The-Art for Real-Time Object Detectors. 2023 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Vancouver, 17-24 June 2023. <https://doi.org/10.1109/CVPR52729.2023.00721>
- [17] Hou, Q., Zhou, D. and Feng, J. (2021) Coordinate Attention for Efficient Mobile Network Design. 2021 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Nashville, 20-25 June 2021, 13713-13722. <https://doi.org/10.1109/CVPR46437.2021.01350>
- [18] Wang, J., Chen, K., Xu, R., et al. (2019) Carafe: Content-Aware Reassembly of Features. 2019 *IEEE/CVF International Conference on Computer Vision (ICCV)*, Seoul, 27 October-02 November 2019, 3007-3016. <https://doi.org/10.1109/ICCV.2019.00310>