

# 基于改进 YOLOv7 的安全帽佩戴检测算法

杨大为, 张成超

(沈阳理工大学 信息科学与工程学院, 沈阳 110159)

**摘要:** 为提高工作场所安全帽佩戴的检测精度, 提出一种基于 YOLOv7 网络架构的改进算法。首先, 在特征提取网络中引入卷积块注意力机制(CBAM)取代 YOLOv7 中主干网络部分原有的卷积模块(CBS), 增强网络的特征提取能力, 加强网络对目标和背景的分辨能力; 其次, 为解决由于网络层数的加深导致小目标特征减弱甚至消失的问题, 增加一个小目标层, 通过将浅层网络特征与深层网络特征融合, 进一步保留小目标特征。实验结果表明, 原 YOLOv7 对安全帽佩戴检测的均值平均精度为 86.1%, 改进后到达 93.4%, 实现了检测精度的提高。

**关键词:** 安全帽; 特征提取网络; 注意力机制; 小目标

中图分类号: TP391.41 文献标志码: A DOI:10.3969/j.issn.1003-1251.2024.01.003

## Algorithm of Safety Helmet Wearing Detection Based on Improved YOLOv7

YANG Dawei, ZHANG Chengchao

(Shenyang Ligong University, Shenyang 110159, China)

**Abstract:** In order to improve the detection accuracy of safety helmet wearing in workplace, an improved algorithm based on YOLOv7 network architecture is proposed. Firstly, a convolutional block attention module(CBAM) is introduced into the feature extraction network to replace the original convolutional module(CBS) in the backbone network of YOLOv7, so as to enhance the feature extraction capability of the network and enhance the recognition ability of the target and background. Secondly, in order to solve the problem that the small target features weaken or even disappear due to the deepening of the number of network layers, a small target layer is added to further preserve the small target features through the integration of shallow network features and deep network features. The mean average precision of the original YOLOv7 detection model for safety helmet wearing is 86.1%. The experimental results show that detection accuracy has been improved to 93.4%.

**Key words:** safety helmet; feature extraction network; attention mechanism; small target

安全帽在建筑工地或矿区等场所是一种必不可少的防护工具, 也是我国生产规模和使用量最大的个体防护产品之一, 是安全工业和应急产品领域的重要组成部分。针对工人在作业期间不佩戴安全帽带来的安全隐患, 我国大部分场所都是

通过人工形式对安全帽佩戴情况进行监督, 难免出现疏忽等情况<sup>[1]</sup>。为此, 研究一种智能化且实时有效的安全帽佩戴检测方法有着重要的意义。

基于深度学习的目标检测主要分为两类: 一类是两阶段目标检测算法; 另一类是一阶段目标

检测算法。前者先由卷积神经网络生成一系列作为样本的候选框,之后进行样本分类和回归<sup>[2]</sup>,代表算法有 R-CNN、Faster R-CNN<sup>[3]</sup>;后者直接通过主干网络得到目标的类别和位置信息,代表算法有 YOLO<sup>[4]</sup>、SSD 和 RetinaNet<sup>[5]</sup>。两阶段的目标检测算法在检测精度上优于一阶段的检测算法,但是检测速度慢。随着一阶段算法的发展,其检测精度不断上升,本文在一阶段检测算法 YOLOv7 的基础上进行改进,以期在保证实时性的同时获得更高的检测精度。

在安全帽检测算法的研究中,文献[6]通过先检测人脸再使用边缘检测算法检测人脸上方是否有安全帽的半圆形轮廓,但此检测算法计算量较大且抗干扰能力较差,漏检和错检的情况较多。文献[7]在 YOLOv5s 模型的主干网络中引入注意力机制并替换主干网络中原有残差块,达到增强细粒度融合的能力,提高了算法的精度和速度。文献[8]通过改进的 YOLOv3 将整个人体作为检测对象,完成安全帽的检测,但由于人体特征较多,其检测精度较低。在工地或高危工作厂区大都存在人员流动大,目标存在遮挡或目标较小的情况,文献[9]通过使用并行网络模块增加网络宽度,并使用四尺度特征预测代替三尺度特征预测等方式保留更多特征信息,以提高 YOLOv3 模型检测精度。文献[10]通过在 Faster R-CNN 网络中引入样本注意力机制和可变卷积等方法优化样本不均衡等问题,以提高目标检测精度,但其检测速度只达到每秒 5 帧,无法做到实时检测。

综上所述,为保证实时性的同时获得更高检测精度,本文对 YOLOv7 进行两处改进:首先,针对特征提取不充分等问题,在 YOLOv7 中使用卷积块注意力机制(CBAM)替换原有卷积模块(CBS),增强特征的提取能力,使网络能够关注更多有用的特征;其次,为防止网络层数加深导致小目标的特征信息过多损失,采用增加小目标层的方法进行浅层和深层的特征信息融合,保留更多的目标特征信息。

## 1 YOLOv7 模型

YOLOv7 结构如图 1 所示,包含输入端(Input)、主干网络(Backbone)、检测头(Head)三个部分。

### 1.1 Input

Input 部分采用数据增强、自适应图片缩放等

技术对图像进行预处理,数据增强方式能丰富数据集,自适应图像缩放则可有效地统一图像的尺寸。

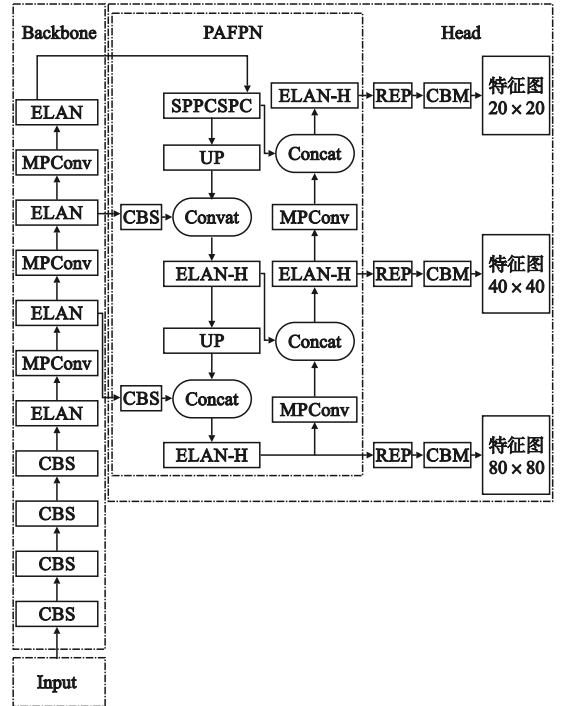


图 1 YOLOv7 结构图

Fig. 1 YOLOv7 structure diagram

### 1.2 Backbone

Backbone 的作用是提取输入图像的特征信息,为特征融合做准备。其主要由若干 CBS 模块、高效层聚合网络(ELAN)模块以及双分支下采样(MPCConv)模块组成。其中 CBS 模块由卷积层、批量归一化层、激活函数组成。

ELAN 结构通过控制不同长度的梯度路径,使更深层次的网络在训练过程中能够有效收敛,其结构如图 2 所示。

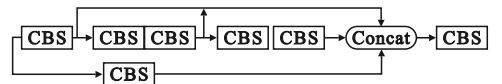


图 2 ELAN 结构图

Fig. 2 ELAN structure diagram

MPCConv 模块分为上下两支,上分支通过最大池化(Maxpool)结合 CBS 的结构进行一次下采样,下分支使用卷积步长为 2 的 CBS 模块进行下采样,最后将两者进行拼接(Concat)操作,防止在下采样的过程中丢失过多的特征信息,同时保持输入输出通道数相同,其结构如图 3 所示。

### 1.3 Head

Head 部分首先使用特征融合网络(PAFPN)

进行两次特征融合,保留更多的浅层信息,其中 SPPCSPC 模块通过不同尺寸的最大池化增大目标感受野,并提高了算法的计算速度。UP 模块调整输入的特征图通道数并进行上采样,为与浅层特征融合做准备。Concat 操作将输入特征图按通道维度进行拼接,以保留更多特征信息。ELAN-H 的结构与 ELAN 类似,唯一的不同是 ELAN-H 将每一层 CBS 都进行了 Concat 操作,保留更多的特征信息。PAFPN 中的 MPCConv 模块与 Backbone 中的类似,通过调整输入特征图通道数为原来的两倍,并进行下采样与浅层特征尺寸对应,以便于进行特征图的 Concat 操作。

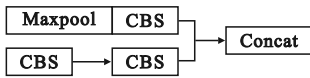


图3 MPCConv 结构图

Fig.3 MPCConv structure diagram

其次,将输出的三个不同尺度的特征经过结构重参数化(Rep)网络进一步提炼并经过卷积进行通道数的调整后,进行目标预测。其中 Rep 网络在训练时为一个三支网络,通过不同尺寸卷积核进行特征提取,增加模型对特征信息的提取能力,在部署时三支结构会转为单分支结构,增加模型的运行速度。CBM 模块整体结构与 CBS 类似,主要作用是通过卷积层调整特征通道数。

## 2 YOLOv7 模型的改进

YOLOv7 虽有一定的先进性,但其在 Backbone 中连续使用 CBS 模块导致特征提取不够充分,丢失信息过多,且在特征融合时损失较多小目标的特征信息。本文通过 CBAM 替换原有 CBS 模块和增加小目标层的方式对模型进行改进,以期达到更好的检测效果。

### 2.1 CBAM

注意力机制是模拟人脑、将有限的算力聚焦于重要区域的算法<sup>[11]</sup>,能够高效地分析复杂场景信息。CBAM 结构如图 4 所示。CBAM 整体又可分为通道注意力模块(CAM)和空间注意力模

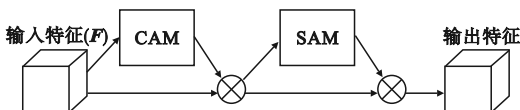


图4 CBAM 结构图

Fig.4 CBAM structure diagram

块(SAM)两部分,CAM 结构如图 5 所示,SAM 结构如图 6 所示。

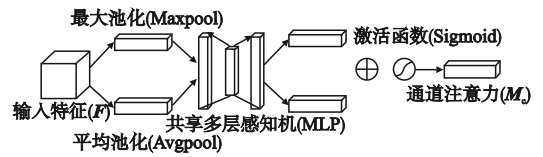


图5 CAM 结构图

Fig.5 CAM structure diagram

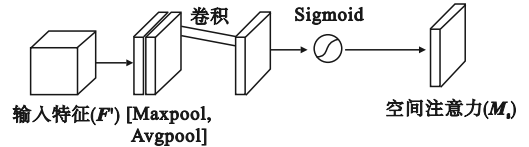


图6 SAM 结构图

Fig.6 SAM structure diagram

在 CAM 部分,首先将输入的特征图分别经过平均池化和最大池化,得到两个长、宽均为 1 的特征图,再将两者送入共享多层感知机中,将得到的两个特征图相加之后经过激活函数得到通道特征图,整体过程用公式表示为

$$M_c(F) = \sigma(\text{MLP}(\text{AvgPool}(F)) + \text{MLP}(\text{MaxPool}(F))) \quad (1)$$

式中: $M_c$  为输出的通道特征图; $F$  为输入特征;MLP 为多层感知机;AvgPool 为平均池化;MaxPool 为最大池化; $\sigma$  为激活函数。

将 CAM 部分所得特征图与  $F$  相乘得到特征图  $F'$  作为 SAM 部分的输入,分别经过最大池化和平均池化,得到两个通道数为 1 的特征图,将两者 Concat 到一起,通过卷积降维操作和 Sigmoid 激活函数生成空间特征图,整体过程用公式表示为

$$M_s(F') = \sigma(f^{7 \times 7}(\text{AvgPool}(F'); \text{MaxPool}(F'))) \quad (2)$$

式中: $M_s$  为输出的空间特征; $F'$  为输入的空间特征; $f^{7 \times 7}$  为  $7 \times 7$  卷积。在 CBAM 前添加一个卷积模块,用于调整输入特征图的通道数,使 CBAM 能够直接替换原 CBS 模块,简化网络的修改。

### 2.2 小目标层

小目标的划分主要分为两类:一类是以相对尺寸划分,要求检测目标的高宽要小于图像高宽的十分之一;另一类是以绝对尺寸进行划分,如 MS COCO 数据集要求小目标的像素值小于  $32 \times 32$ 。小目标相比于大、中目标在特征提取时信息过少,匹配锚点框的训练效果较差,影响检测精度<sup>[12]</sup>。文献[13]通过改进 Faster R-CNN 算法提

高无人机对小目标的识别精度,文献[14]通过改进 YOLOv5s 的网络模型提高算法对安全帽的检测精度,文献[15]通过改进 SSD 算法提高模型对于小目标的检测精度和速度。在目标检测过程中,小目标的特征信息会随着网络模型的加深逐渐消失,在环境复杂多变、人员流动性强的情况下,小目标检测精度仍需提高。

本文提出在 YOLOv7 模型的基础上增加一个

小目标层的方法。该方法的思想源于 PAFPN,在原 YOLOv7 结构中,通过 PAFPN 会输出三种尺寸的特征,该方法是将浅层特征与深层特征融合,达到即使在深层网络中依然可以保留小目标信息的作用。本文增加的小目标层可以将更浅层特征与深层特征相融合,进一步保留小目标特征。本文改进后的 YOLOv7 模型结构如图 7 所示。

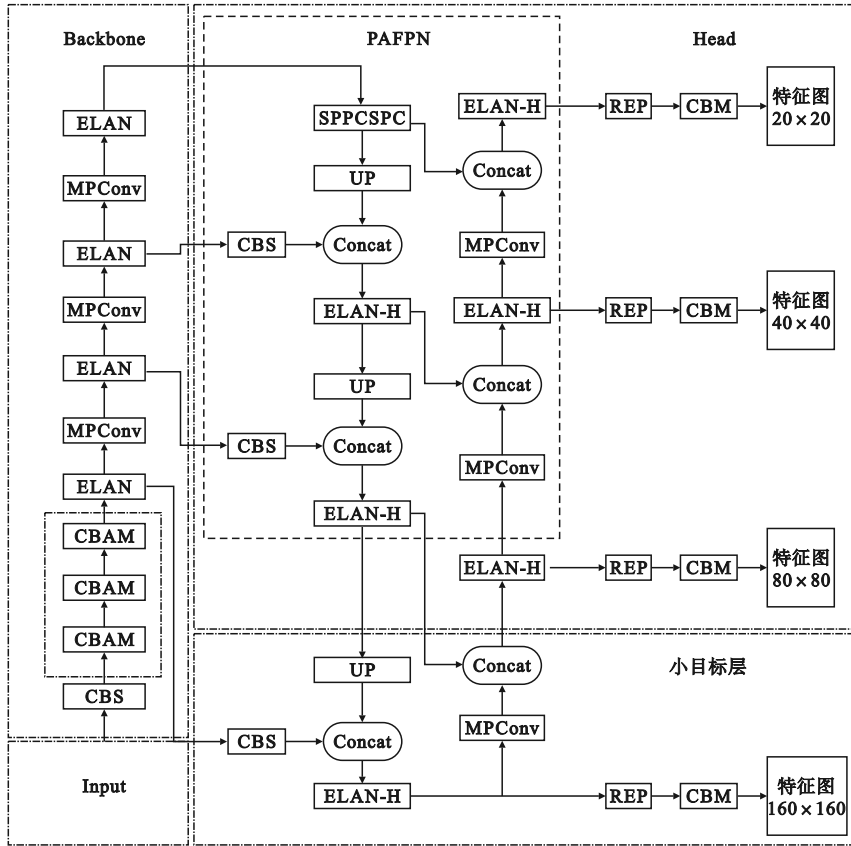


图 7 改进后 YOLOv7 结构图

Fig.7 Structure diagram of improved YOLOv7

### 3 实验与结果分析

实验运行的服务器环境配置如下:CPU 型号为 E5-2680 v4@2.4 GHz,显卡型号为 RTX2080Ti,系统为 Ubuntu 20.04,深度学习框架采用 pytorch 1.9.0。

#### 3.1 数据集与评估指标

本次实验采用的数据是从开源的 SHWD 数据集中抽取 1 500 张图片,该数据集中被检测对象共分为两类:佩戴安全帽的人(hat)和未佩戴安全帽的人(person),两者样本个数比约为 1:13。将本次实验所用图片按照 8:2 的比例划分为训练集和验

证集进行实验。采用数据集样本如图 8 所示。



图 8 样本图像

Fig.8 Sample images

SHWD 数据集虽然不是完全的小目标数据集,但小目标占比较大,一定程度上能够反映算法对小目标的检测能力。

本次实验所采用的评估指标包括准确率(Precision),召回率(Recall),均值平均精度(mAP)。

### 3.2 实验结果分析

为验证不同个数的 CBAM 替换对原 YOLOv7 模型的影响,本文进行四组对比实验,结果如表 1 所示。

从表 1 可见,随着 CBAM 替换数量的增加,算法的检测精度并未一直随之上升,替换三层后的结果相对较好。

为验证本文算法的改进效果,将本文改进算法与其他检测算法进行对比实验。具体结果如表 2 所示。

从表 2 可见,本文提出的改进算法的检测精度优于其他算法。在研究过程中,还进行了不同改进方式的对比实验,具体实验结果如表 3 所示。

表 1 四组实验结果对比

Table 1 Comparisons of four experimental results

%

模型	hat			person		
	Precision	Recall	mAP	Precision	Recall	mAP
替换一层	83.2	86.9	88.6	93.2	81.4	91.5
替换二层	84.4	87.6	87.8	92.4	78.4	89.9
替换三层	84.0	89.2	91.8	95.6	85.8	95.2
替换四层	81.5	89.6	90.4	94.1	84.8	94.2

表 2 不同算法检测结果对比

Table 2 Comparisons of detection results with different algorithms

%

模型	hat			person		
	Precision	Recall	mAP	Precision	Recall	mAP
YOLOv5	84.1	90.3	90.1	95.2	88.9	94.5
YOLOv7	80.9	84.9	86.1	89.7	78.7	87.6
YOLOv8	85.0	90.3	91.0	95.4	81.9	93.3
本文算法	85.4	90.7	93.4	96.6	87.8	96.6

表 3 不同改进方式检测结果对比

Table 3 Comparisons of detection results with different improvement methods

%

模型	hat			person		
	Precision	Recall	mAP	Precision	Recall	mAP
YOLOv7	80.9	84.9	86.1	89.7	78.7	87.6
增加 CBAM	75.3	87.1	79.0	86.9	78.1	84.4
增加一个小目标层和 CBAM	73.6	85.3	74.2	89.7	78.3	87.2
增加两个小目标层和替换 CBAM	78.6	84.9	86.6	93.3	85.2	93.0
本文算法	85.4	90.7	93.4	96.6	87.8	96.6

在改进方式的对比研究中,首先在 Backbone 中每相邻的两个 CBS 模块之间增加一层 CBAM,具体增加位置如图 9 所示。

在此基础上增加小目标层再次进行测试,通过表 3 中数据可以看出,本文改进算法目标检测

效果更好;增加多个小目标层后精度不升反降,并且会导致模型参数量增加,降低模型的检测速度。

经实验测试,原 YOLOv7 模型检测速度约为每秒 60 帧,本文改进算法约为每秒 45 帧,虽然检测速度有所下降,但较大提升了目标的检测精度。

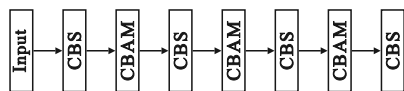


图 9 CBAM 增加位置结构图

Fig. 9 CBAM added position structure diagram

## 4 结论

针对工地或矿区等场所因环境复杂多样,工人在工作期间时常会出现人员聚集、人员遮挡等不利于安全帽佩戴情况检测,造成漏检或错检的问题,本文提出了改进的 YOLOv7 检测模型,该模型在 YOLOv7 的 Backbone 中使用 CBAM 替换 CBS 模块,使算法能够提取更多有效特征,再加入小目标层以防止小目标特征随着模型深度的加深而损失过多。与原 YOLOv7 模型所进行的对比实验表明,对安全帽佩戴检测的 mAP 指标由 86.1% 提高到 93.4%,实现了目标检测精度的提升。

## 参考文献 (References):

- [1] 王忠玉. 智能视频监控下的安全帽佩戴检测系统的设计与实现[D]. 北京:北京邮电大学,2018.
- [2] 戚玲珑,高建瓴. 基于改进 YOLOv7 的小目标检测[J]. 计算机工程,2023,49(1):41-48.  
QI L L, GAO J L. Small object detection based on improved YOLOv7[J]. Computer Engineering, 2023, 49(1):41-48. (in Chinese)
- [3] 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.
- [4] 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. Las Vegas, USA: IEEE, 2016:779-788.
- [5] 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.
- [6] SHRESTHA S, SHRESTHA P P, DINESH B, et al. Hard-hat detection for construction safety visualization [J]. Journal of Construction Engineering, 2015, 2015(1):1-8.
- [7] 宋晓凤,吴云军,刘冰冰,等. 改进 YOLOv5s 算法的安全帽佩戴检测[J]. 计算机工程与应用,2023,59(2):194-201.

- SONG X F, WU Y J, LIU B B, et al. Improved YOLOv5s algorithm for helmet wearing detection [J]. Computer Engineering and Applications, 2023, 59(2):194-201. (in Chinese)
- [8] 施辉,陈先桥,杨英. 改进 YOLO v3 的安全帽佩戴检测方法 [J]. 计算机工程与应用, 2019, 55(11):213-220.  
SHI H, CHEN X Q, YANG Y. Safety helmet wearing detection method of improved YOLO v3 [J]. Computer Engineering and Applications, 2019, 55(11):213-220. (in Chinese)
- [9] SONG H R. Multi-scale safety helmet detection based on RSSE-YOLOv3 [J]. Sensors, 2022, 22(16):6061.
- [10] 储开斌,叶托,张继. 基于改进 Faster R-CNN 的头盔检测算法研究 [J]. 国外电子测量技术, 2022, 41(6):86-92.  
CHU K B, YE T, ZHANG J. Research on helmet detection algorithm based on improved Faster R-CNN [J]. Foreign Electronic Measurement Technology, 2022, 41(6):86-92. (in Chinese)
- [11] 李盼盼,王朝立,孙占全. 基于注意力机制的多特征融合人脸活体检测 [J]. 信息与控制, 2021, 50(5):631-640.  
LI P P, WANG C L, SUN Z Q. Face liveness detection based on multi-feature fusion with an attention mechanism [J]. Information and Control, 2021, 50(5):631-640. (in Chinese)
- [12] 张艳,张明路,吕晓玲,等. 深度学习小目标检测算法研究综述 [J]. 计算机工程与应用, 2022, 58(15):1-17.  
ZHANG Y, ZHANG M L, LYU X L, et al. Review of research on small target detection based on deep learning [J]. Computer Engineering and Applications, 2022, 58(15):1-17. (in Chinese)
- [13] 王君,张德育,康鑫英. 改进 Faster-RCNN 的低空小型无人机检测方法 [J]. 沈阳理工大学学报, 2021, 40(4):23-28.  
WANG J, ZHANG D Y, KANG X Y. Improved detection method of low altitude small UAV by Faster-RCNN [J]. Journal of Shenyang Ligong University, 2021, 40(4):23-28. (in Chinese)
- [14] 李嘉信,胡杨,黄协舟,等. 面向小目标的多空间层次安全帽检测 [J/OL]. 计算机工程与应用, 2022:1-9 [2022-11-29]. <https://kns.cnki.net/kcms/detail/11.2127.TP.20221129.1158.004.html>.  
LI J X, HU Y, HUANG X Z, et al. Small target-oriented multi-space hierarchical helmet detection [J/OL]. Computer Engineering and Applications, 2022:1-9 [2022-11-29]. <https://kns.cnki.net/kcms/detail/11.2127.TP.20221129.1158.004.html>. (in Chinese)
- [15] 康鑫英,张德育,王君. 压缩 SSD 低空小目标检测算法研究 [J]. 沈阳理工大学学报, 2022, 41(6):1-5.  
KANG X Y, ZHANG D Y, WANG J. Research on detection algorithm of small target at low altitude based on compressed SSD [J]. Journal of Shenyang Ligong University, 2022, 41(6):1-5. (in Chinese)

(责任编辑:和晓军)