

丰富上下文捕捉多层特征的安全帽检测算法

齐向明, 宋雨格

(辽宁工程技术大学 软件学院, 辽宁 葫芦岛 125105)

摘要: 佩戴安全帽是保证复杂环境安全生产的必要措施。为此, 提出丰富上下文捕捉多层特征的安全帽检测算法。该算法以YOLOv10n为基线, 将CFBlock与CGLU结合后替换C2f中的Bottleneck, 用于传递空间特征, 调整通道重要程度, 丰富上下文信息, 降低参数数量和计算量; 在主干网SPPF中引入LSKA, 用于扩大卷积核感受野, 捕捉更多空间特征, 提高检测精度; 采用Weighted Fusion改进BiFPN, 融合张量方式后替换Neck部分, 用于双向连接多尺度信息, 捕捉多层特征, 降低参数数量; 将WIoU与MPDIoU融合后替换基线损失函数CIoU, 以最小化预测框与真实框的距离, 缓解正负样本不平衡, 增强泛化能力。实验结果表明, 与基线算法和次优算法相比, 所提算法在颜色检测数据集GDUT-HWD上的Parameters、FLOPs分别降低18.1%、13.8%, 在整体检测数据集SHEL5K上分别降低17.6%、13.8%, 且精确度、平均精度均值和召回率均有明显提升; 在佩戴检测数据集SHWD上的泛化实验结果表明, 所提方法鲁棒性良好。

关键词: 安全帽检测; YOLOv10n; CFBlock; CGLU; LSKA; BiFPN; WIoU; MPDIoU

DOI: 10.11907/rjtk.251142

扫描二维码阅读全文:



中图分类号: TP391

文献标识码: A

文章编号: 1672-7800(2026)005-0177-09

Helmet Detection Algorithm with Rich Context and Multi-Layer Feature Capture

QI Xiangming, SONG Yuge

(School of Software, Liaoning Technical University, Huludao 125105, China)

Abstract: Wearing safety helmet is a necessary measure to ensure safety production in complex environment. For this reason, a helmet detection algorithm is proposed to enrich the multi-level features of context capture. The algorithm takes YOLOv10n as the baseline, replaces Bottleneck in C2f after combining CFBlock and CGLU, which is used to transfer spatial features, adjust channel importance, enrich context information, reduce parameter amount and computational complexity; LSKA is introduced into the backbone SPPF to expand the convolutional nuclear receptive field, capture more spatial features and improve the detection accuracy; Weighted Fusion is used to improve the BiFPN fusion tensor method and replace the Neck part, which is used to connect multi-scale information in both directions, capture multi-layer features and reduce the number of parameters; The baseline loss function CIoU is replaced after the fusion of WIoU and MPDIoU to minimize the distance between the prediction box and the real box, alleviate the imbalance between positive and negative samples, and enhance the generalization ability. The experimental results show that, compared with the baseline algorithm and the suboptimal algorithm, the parameters and FLOPs of the proposed algorithm on the color detection dataset GDUT-HWD are reduced by 22.2% and 16.1% respectively, and the overall detection dataset SHEL5K is reduced by 21.4% and 16.1% respectively, and the accuracy, average precision mean and recall rate are significantly improved; The generalization experiment results on SHWD wear detection dataset show that the proposed method is robust.

Key Words: safety helmet detection; YOLOv10n; CFBlock; CGLU; LSKA; BiFPN; WIoU; MPDIoU

0 引言

施工现场往往存在光线昏暗、物体相互遮挡以及目标

分布密集等情况,极大增加了工作中的潜在危险。在复杂施工现场,佩戴安全帽是保障安全生产的一项必要举措。然而,受复杂场景的制约,现有安全帽智能检测算法存在参数量多、计算量大的问题,难以在实时性要求较高或资

收稿日期: 2025-03-25

基金项目: 国家自然科学基金项目(62173171)

作者简介: 齐向明(1966-),男,硕士,辽宁工程技术大学软件学院副教授,研究方向为图像与视觉信息计算;宋雨格(1999-),女,CCF学生会员,辽宁工程技术大学软件学院硕士研究生,研究方向为图形图像处理与计算机视觉。本文通讯作者: 宋雨格。

源受限的设备上部署应用。

在深度学习领域,目标检测算法是主流研究方向,该类算法以YOLO(You Only Look Once)^[1-5]、SSD(Single Shot Multibox Detector)^[6]、Cascade-RCNN(Cascade Region-based Convolutional Neural Network)^[7]等为代表。其中,YOLO系列算法通过直接回归的方式预测目标的类别和位置,不仅简化了算法结构,而且优化了训练过程,在检测速度和精度方面表现出色,相较于其他算法更具优势。目前,许多学者将该类算法应用于安全帽的检测中。例如,杜晓刚等^[8]提出一种基于YOLOv5的安全帽佩戴检测算法,将Swin Transformer作为主干网络,替代原有特征提取结构,强化了全局与局部特征的提取能力,取得了更准确的安全帽佩戴检测结果,但存在误检率较高的问题。伏德栗等^[9]在YOLOv5的基础上引入CA(Coordinate Attention)机制与双向特征金字塔网络(Bidirectional Feature Pyramid Network, BiFPN)结构,提升了对安全帽的关注度与捕获能力,减少了漏检和误检现象,但模型计算量有一定提升。祁泽政等^[10]提出一种基于YOLOv5s的安全帽检测模型,将混合池化融入空间金字塔结构,同时嵌入CA注意力机制,显著增强了对安全帽细节信息的提取能力,提升了检测准确率,但模型参数量未得到精简,不利于轻量化部署。白培瑞等^[11]提出一种基于YOLOv5的安全帽检测与跟踪模型,结合DeepSORT多目标跟踪算法,同时融入BiFPN特征融合结构与Transformer注意力机制,有效提高了复杂环境下安全帽的识别准确率与跟踪稳定性,但模型结构过于复杂,难以适配实时检测场景。赵敏等^[12]提出一种基于YOLOv7-tiny的安全帽检测模型,引入EPSANet Block注意力模块,强化了模型对安全帽目标信息的聚焦程度,有效提升了检测性能,但模型浮点运算量大幅上升,牺牲了轻量化优势。杨大为等^[13]提出一种基于YOLOv7的安全帽检测模型,采用卷积注意力机制替换原有卷积模块,同时新增小目标检测层适配安全帽尺寸特性,有效提升了检测精度,但模型计算量增加,检测速度较慢。凌港等^[14]提出一种基于YOLOv8的安全帽检测模型,在网络Neck层融入effectiveSE注意力机制,同时设计SPS检测头优化目标预测能力,提升了复杂场景下的安全帽检测稳定性,但模型参数量较大,轻量化程度不足。冯勇等^[15]在YOLOv8模型中引入轻量化自适应权重下采样LAD模块与高效多尺度卷积结构,在显著降低模型参数量与计算量的同时保持了较高的检测精度,但对不同尺度安全帽目标的检测适应性仍有提升空间。王新良等^[16]基于YOLOX改进遮挡场景下的安全帽检测模型,通过设计融合空洞卷积组与卷积注意力机制的感受野增强模块,有效降低了被遮挡目标的漏检率,但模型参数量与浮点运算量同步增加,影响了检测效率。

综上所述,安全帽智能检测模型仍存在复杂度高、参数量大、计算速度慢、检测精度不足等问题。为此,本文以

YOLOv10n为基线,提出丰富上下文捕捉多层特征的安全帽检测算法,主要创新点为:①将卷积变换块(ConvFormer Block, CFBBlock)与卷积门控线性单元(Convolutional GLU, CGLU)结合起来,置换基线算法所有C2f中的Bottleneck模块,用于调整空间维度和通道维度,控制信息流动,丰富上下文信息,降低参数量和计算量;②在主干网络SPPF(Spatial Pyramid Pooling-Fast)模块中引入大型可分离核注意力机制(Large Separable Kernel Attention, LSKA),用于扩大卷积核感受野,捕捉更多空间特征,提高检测精度;③采用Weighted Fusion改进BiFPN后再替换Neck部分,用于捕捉多层特征,提高特征融合能力,降低模型参数量;④将WIoU(Wise Intersection over Union)与MPDIoU(Minimum Point Distance Intersection over Union)融合后替换原损失函数CIoU,以最小化预测边界框与真实边界框之间的距离,缓解正负样本不平衡,增强模型的泛化能力。

1 理论基础

1.1 基线算法

YOLOv10在YOLOv8的基础上消除非最大抑制并优化各种模型组件,有n、s、m、b、l和x等版本,均包括Backbone、Neck和Head 3个部分,分4个通道提取特征图,分别为S0(640×640×3)、S1(80×80×256)、S2(40×40×512)、S3(20×20×1024);分3个通道输出预测特征图,分别为S4(20×20×45)、S5(40×40×45)、S6(80×80×45)。具体结构见图1。YOLOv10的Backbone部分继承了YOLOv8的C2f和SPPF(Spatial Pyramid Pooling-Fast)模块,并新增SCDown(Spatial and Channel Down-sampling)、C2fCIB(C2f with Channel-wise Interaction Block)、PSA(Point-wise Spatial Attention)模块。其中,SCDown用于调节通道数,解耦空间特征和通道特征,降低计算成本,保留更多信息;C2fCIB通过卷积层提取空间特征,然后通过全连接层进行特征融合与分类;PSA模块自适应加权分配像素,聚合局部特征向量,捕捉多尺度细节,提高特征提取能力。Neck部分通过FPN(Feature Pyramid Network)和PAN(Path Aggregation Network)双向感知特征传递目标。Head部分由一对一检测头、一对多检测头和统一匹配度量 $m(\alpha, \beta)$ 组成,其中一对一检测头生成一个最佳预测,减少推理延迟;一对多检测头生成多个预测,提高学习准确性; $m(\alpha, \beta)$ 统一匹配度量,感知匹配,计算公式为:

$$m(\alpha, \beta) = s * p^\alpha * IoU(b, \hat{b})^\beta \quad (1)$$

式中: p 为分类分数; b 和 $\hat{b}$ 分别为预测和实例的边界框; s 为空间先验; IoU 为重叠面积交并比。

1.2 相关技术

1.2.1 CFBBlock

CFBlock基于MetaFormer框架,将Token Mixer替换为SepConv,由Norm、SepConv、Channel MLP和Add组成^[17]。

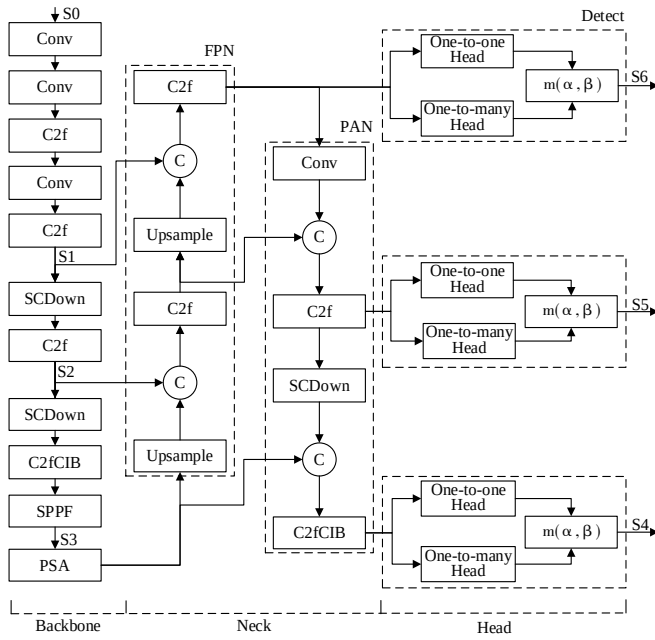


Fig. 1 YOLOv10 structure

图1 YOLOv10 结构

结构见图2。具体流程为:①由 Norm 对特征图进行归一化处理,以减少协变量偏移,提高模型的收敛速度;②由 Sep-Conv 可分离卷积对归一化后的特征图进行处理,调整其空间维度与通道维度,进而捕捉局部特征;③通过残差操作合并局部特征图与原始特征图,以传播梯度,同时保持信息的完整性;④借助 Norm 对残差合并后的特征图进行归一化;⑤由 Channel MLP 在通道维度上进行多层感知运算,提取高维度特征,以提高特征的表达能力;⑥由 Add 操作将高维度特征图、局部特征图与原始特征图的残差合并,进一步传播梯度,同时保持特征表达的完整性。

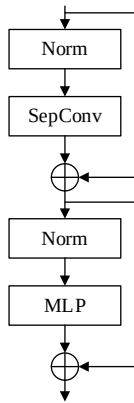


Fig. 2 CFBBlock structure

图2 CFBBlock 结构

1.2.2 CGLU

CGLU 由 Linear、DW-Conv、Sigmoid、Hadamard、Add 组成,其基于最近邻特征深度卷积提供位置编码,动态调整通道重要度,捕捉不同尺度空间信息,增强算法的表达能力^[18]。结构见图3。具体流程为:①由 Linear 梯度最小化损失函数运算对特征图进行线性加权,以改善模型的预测

能力;②由 DW 轻量化分离卷积与逐点卷积捕获局部特征的位置编码,同时提取空间特征;③由 Sigmoid 学习特征图的非线性关系,生成信息流动的门控信号,并将空间特征的相关性表达为空间特征图;④进行 Hadamard 运算,对特征图与空间特征图进行逐元素乘积,实现自适应权重分配,增强空间特征的表达效果;⑤由 Linear 对空间特征图进行线性加权,进一步强化空间特征;⑥由 Add 残差连接合并再次线性加权后的空间特征图与原始特征图,以缓解梯度消失问题,同时保持信息的完整性。

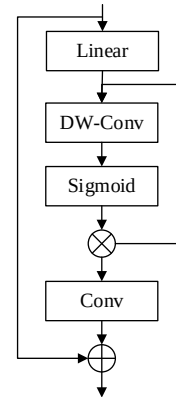


Fig. 3 CGLU structure

图3 CGLU 结构

1.2.3 LSKA

LSKA 由 DW-Conv、DW-D-Conv 和 Conv 和 Hadamard 组成,其可将 2D 卷积核分解为级联 1D 水平卷积核和 1D 垂直卷积核,以扩大卷积核感受野,降低计算复杂性和内存占用^[19]。结构见图4。具体流程为:①由 DW-Conv 分别对水平与垂直方向卷积,提取两个方向的特征,生成初步的注意力图矩阵;②由 DW-D-Conv 再次对水平与垂直方向卷积,在提取空间特征的同时扩大感受野,捕捉上下文信息;③由 Conv 对空间特征图矩阵进行融合,生成最终的注意力图矩阵,实现对重要空间特征的关注;④进行 Hadamard 运算,对最终注意力图矩阵与原始特征图矩阵进行逐元素乘积,自适应调整特征权重,从而突出重要特征。

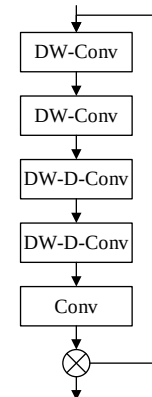


Fig. 4 LSKA structure

图4 LSKA 结构

1.2.4 Weighted Fusion

Weighted Fusion由Conv和Add组成,其通过融合多个特征图来调整空间和通道维度,降低参数数量和计算量。结构见图5。具体流程为:①由Conv对多通道进行线性组合,调整不同通道的权重,以提取高级特征;②由Add操作对多个特征图融合,同步调整空间维度和通道维度信息,提升特征的表达能力。

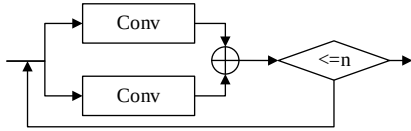


Fig. 5 Weighted Fusion structure

图5 Weighted Fusion结构

1.2.5 损失函数

CIoU采用纵横比相对差异 v 衡量边界框间的相似性,不考虑真实差异,可能会导致计算偏差和正样本优化困难。其计算公式为:

$$CIoU = IoU - \frac{\rho^2(B^{pred}, B^{gt})}{c^2} - \alpha v \quad (2)$$

式中: B^{pred} 、 B^{gt} 分别为预测框和真实框中心点; $\rho^2(B^{pred}, B^{gt})$ 为预测框与真实框中心距离; c 为对角线距离; IoU 为交并比; v 为纵横比; α 为平衡参数。

WIoU通过引入加权机制,对预测框与真实框的交集区域赋予不同权重,以更精准地衡量二者之间的匹配程度。该函数有效解决了传统IoU在目标不相交时值为零,

无法反映预测偏差,以及在目标相同同时对不同方向或程度的预测偏差区分度不足的问题^[20]。其计算公式为:

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

式中: b_i 为目标框坐标; g_i 为真实框坐标; IoU 为交并比; w_i 为权重。

MPDIoU通过最小化预测框与真实框之间的对角距离,构建了一种更为精细的框体匹配评估机制。该函数不仅在常规场景下能够显著提升目标检测的定位准确性,而且在面对复杂场景中的遮挡、形变等干扰因素时亦展现出更强的鲁棒性^[21]。其计算公式为:

$$MPDIoU = IoU - \frac{\rho^2(P_1^{pred}, P_1^{gt})}{w^2 + h^2} - \frac{\rho^2(P_2^{pred}, P_2^{gt})}{w^2 + h^2} \quad (4)$$

式中: P_1^{pred} 、 P_1^{gt} 和 P_2^{pred} 、 P_2^{gt} 分别为预测框和真实框的对角点; $\rho^2(P_1^{pred}, P_1^{gt})$ 和 $\rho^2(P_2^{pred}, P_2^{gt})$ 分别为预测框和真实框的对角距离; w 和 h 分别为边界框宽和高的预测值; IoU 为交并比。

2 算法构建

本文算法分5个通道提取特征图S0(640×640×3)、S1(160×160×128)、S2(80×80×256)、S3(40×40×512)、S4(20×20×1024),分3个通道输出预测特征图S5(20×20×45)、S6(40×40×45)、S7(80×80×45)。具体结构见图6。以下详细介绍该算法的改进点。

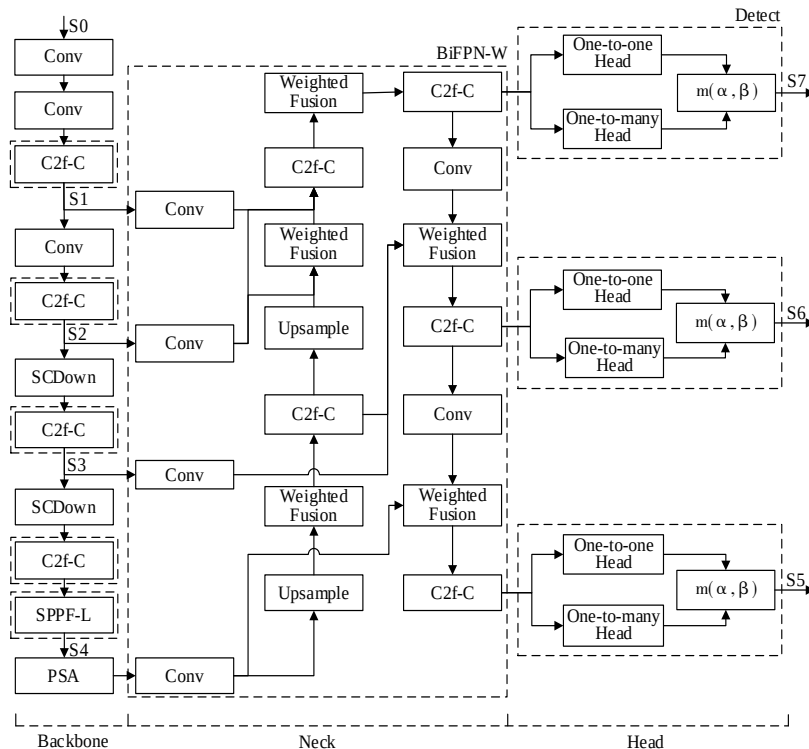


Fig. 6 Structure of the proposed algorithm

图6 本文算法结构

2.1 丰富上下文信息

2.1.1 增强表达能力

CFBlock中的Channel MLP无法捕捉空间结构,易丢失空间信息,特征提取能力减弱,将其替换为CGLU,以动态调整通道重要程度,捕捉多尺度空间信息,增强表达能力,标记为CF-CGLU。结构见图7。

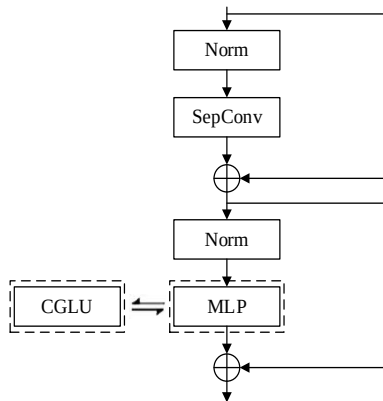


Fig. 7 CF-CGLU structure

图7 CF-CGLU结构

2.1.2 加强上下文传递

C2f中的Bottleneck特征提取相对复杂,参数量和计算量大,将全部C2f中的Bottleneck置换为CF-CGLU后,采用CFBlock中的分离卷积调整空间维度和通道维度,CGLU控制信息流动,丰富全局上下文信息,同时减少参数量和计算量,标记为C2f-C。结构见图8。

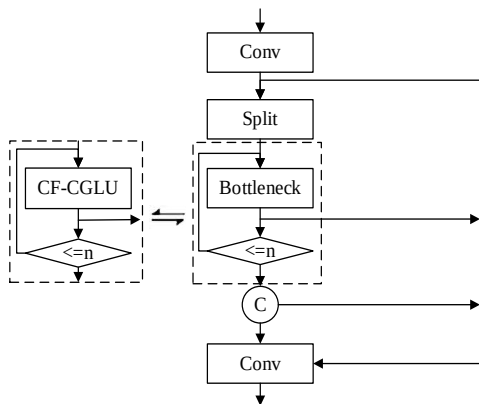


Fig. 8 C2f-C structure

图8 C2f-C结构

2.2 增强特征提取

SPPF裁剪输出固定维度信息,易忽略细粒度特征,引入LSKA注意力机制后,DW卷积参数量减少,工作效率提高。DW-D卷积将DW卷积产生的多个特征图合并成一个新特征图,以广泛捕捉上下文信息,增强特征提取能力,标记为SPPF-L。结构见图9。

2.3 捕捉多层特征

原始算法的Neck部分串行FPN和PAN双向传递特征图,信息分离,参数量大,融合不充分。本文算法采用Weighted Fusion改进BiFPN的融合张量方式,替换Neck部

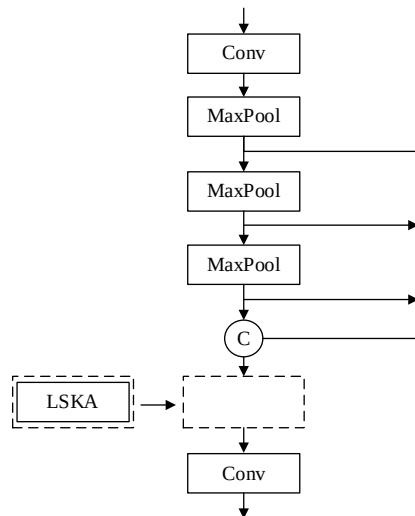


Fig. 9 SPPF-L structure

图9 SPPF-L结构

分融合多尺度特征图,用于捕捉多层特征、降低参数量^[22]。BiFPN删除多尺度特征单向节点,同层节点添加额外边,自顶向下传递高层语义特征,自底向上传递低层位置特征,优化跨尺度连接。同时,采用Weighted Fusion替换BiFPN中的Fast Normalized Fusion,轻量化加权融合多个特征图,标记为BiFPN-W,结构见图6的Neck部分。

2.4 增强泛化能力

将WIoU与MPDIoU后融合替换CIoU,最小化预测边界框与真实边界框的距离,缓解正负样本不平衡,提高复杂场景下的目标位置预测准确率,增强泛化能力,标记为W-MPDIoU。计算公式为:

$$W - MPDIoU = WIoU - \frac{\rho^2(P_1^{pred}, P_1^{gt})}{w^2 + h^2} - \frac{\rho^2(P_2^{pred}, P_2^{gt})}{w^2 + h^2} \quad (5)$$

式中: P_1^{pred} , P_1^{gt} 和 P_2^{pred} , P_2^{gt} 分别为预测框和真实框的对角点; $\rho^2(P_1^{pred}, P_1^{gt})$ 和 $\rho^2(P_2^{pred}, P_2^{gt})$ 分别为预测框和真实框的对角距离; w 和 h 分别为边界框宽和高的预测值。

3 实验方法与结果分析

3.1 实验环境与参数设置

硬件环境:显卡为RTX3090,显存为24 GB,CPU为14 vCPU Intel(R) Xeon(R) Platinum 8362 CPU@ 2.80GHz,内存为45 GB。软件环境:操作系统为Ubuntu20.04,编程模型为CUDA11.8,编程环境为Pytorch2.0.0,编程语言为Python3.8.0。参数设置:采用Mosaic方式增强数据,迭代批量为16,优化器为SGD,初始学习率为0.01,动量参数为0.937,学习算法为余弦退火,训练轮数为200。

3.2 数据集

采用以下3种数据集进行实验:①颜色检测公开数据集GDUT-HWD。包含蓝、黄、白、红和无5类安全帽场景共3 174张图像,按照7:2:1的比例为2 221张训练图片、635张验证图片和318张测试图片;②整体检测公开数据

集 SHEL5K。包含戴安全帽头、没戴安全帽头、戴安全帽人、没戴安全帽人、安全帽、人脸共 6 类 5 000 张图像,按照 8:1:1 的比例划分为 4 000 张训练图片、500 张验证图片和 500 张测试图片;③佩戴检测公开数据集 SHWD。含有佩戴安全帽人和未佩戴安全帽人两类共 7 581 张图像,按照 8:1:1 的比例划分为 6 065 张训练图片、758 张验证图片和 758 张测试图片。

3.3 评价指标

采用精确率(Precision)、召回率(Recall)、平均精度(mAP)、参数量(Params)、计算量(FLOPs)评价算法性能。其中,精确率为模型预测出的目标中被正确识别的比例。计算公式为:

$$\text{Precision} = \frac{TP}{TP + FP} \quad (6)$$

式中:TP 表示预测为正类,且实际标签也为正类的样本数量;FP 表示预测为正类,但实际标签为负类的样本数量。

召回率为所有真实存在的正类目标中被模型成功预测出的比例。计算公式为:

$$\text{Recall} = \frac{TP}{TP + FN} \quad (7)$$

式中:FN 表示预测为负类,但实际标签为正类的样本数量。

mAP 反映各类别的平均精度。计算公式为:

$$AP = \int_0^1 P(R) dR \quad (8)$$

$$\text{mAP} = \frac{1}{c} \sum_{i=1}^c AP_i \quad (9)$$

式中:AP 为数据集中单一类别的平均检测精度;c 为检测目标总类别数;i 为检测次数。

3.4 实验结果分析

3.4.1 比较实验

在 GDUT-HWD 和 SHEL5K 数据集上对本文算法与基线算法 YOLOv10n 进行 P-R 曲线的比较,结果见图 10、图 11。可以看出,在 GDUT-HWD 数据集上,本文算法对蓝色、黄色、白色、红色、没有佩戴安全帽的各类别 P-R 图曲线均在基线算法之上,与坐标轴围成的面积更大,表明本文算法检测精度更高。在 SHEL5K 数据集上,本文算法对戴安全帽头、没戴安全帽头、戴安全帽人、人脸、没戴安全帽人各类别的 P-R 图曲线均在基线算法之上,与坐标轴围成的面积更大,表明本文算法检测精度更高。

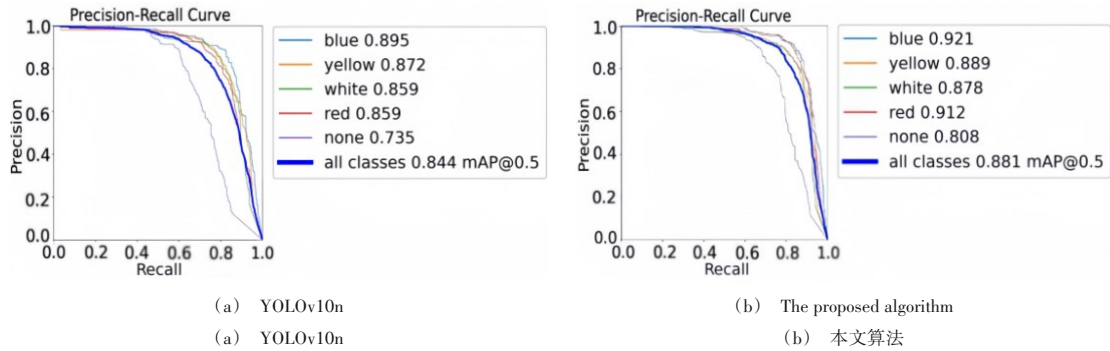


Fig. 10 P-R curve on GDUT-HWD dataset

图 10 在 GDUT-HWD 数据集上的 P-R 曲线

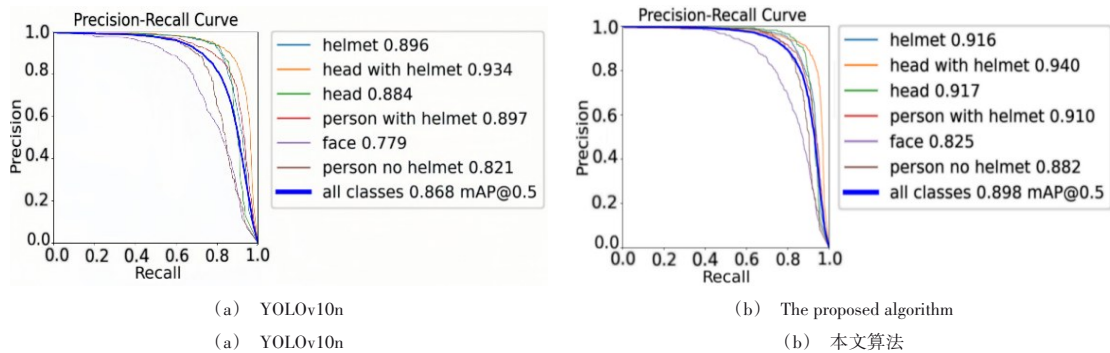


Fig. 11 P-R curve on SHEL5K dataset

图 11 在 SHEL5K 数据集上的 P-R 曲线

在 GDUT-HWD 数据集上进行密集、昏暗和遮挡场景下的可视化比较,结果见图 12,其中第一行为密集场景,第二行为昏暗场景,第三行为遮挡场景。可以看出,针对密集场景,基线算法没有检测出左侧中间、左侧下角以及右下角位

置目标,而本文算法准确检测出以上目标;针对昏暗场景,基线算法没有检测到中间位置昏暗处戴黄色安全帽的人,本文算法准确检测出该目标;针对遮挡场景,基线算法未检测出左侧被遮挡的人,而本文算法准确检测出该目标。



Fig. 12 Visual test result comparison
图 12 可视化检测结果比较

在 GDUT-HWD 和 SHEL5K 数据集上,将本文算法与 Cascade-RCNN、SSD、YOLOv5n、YOLOv8n、YOLOv8s、YOLOv9s、YOLOv10n、YOLOv10s 进行比较,结果见表 1 和表 2。可以看出,在 GDUT-HWD 数据集上,本文算法的 Parameters 和 FLOPs 分别为 1.85 M 和 5.6 G,显著低于多数算法;

mAP@0.5、Precision 和 Recall 分别为 88.1%、87.7%、80.1%,高于所有算法。在 SHEL5K 数据集上,本文算法的 Parameters 和 FLOPs 分别为 1.87 M 和 5.6 G,显著低于多数算法;mAP@0.5、Precision 和 Recall 分别为 89.8%、88.3%、81.6%,高于所有算法。综合来看,本文算法整体性能最佳。

Table 1 Performance of different algorithms on GDUT-HWD dataset
表 1 不同算法在 GDUT-HWD 数据集上的表现

算法	Parameters/M	FLOPs/G	mAP@0.5/%	Precision/%	Recall/%
Cascade-RCNN	68.24(↑ 66.39)	79.1(↑ 73.5)	84.1(↓ 4.0)	87.1(↓ 0.6)	76.5(↓ 3.6)
SSD	23.05(↑ 21.20)	52.5(↑ 46.9)	81.8(↓ 6.3)	83.5(↓ 4.2)	76.9(↓ 3.2)
YOLOv5n	1.76(↓ 0.09)	4.1(↓ 1.5)	84.9(↓ 3.2)	87.4(↓ 0.3)	78.3(↓ 1.8)
YOLOv8n	3.00(↑ 1.15)	8.1(↑ 2.5)	83.4(↓ 4.7)	87.3(↓ 0.4)	77.9(↓ 2.2)
YOLOv8s	11.13(↑ 9.28)	28.4(↑ 22.8)	85.2(↓ 2.9)	87.6(↓ 0.1)	77.7(↓ 2.4)
YOLOv9s	7.13(↑ 5.28)	26.5(↑ 20.9)	86.3(↓ 1.8)	87.5(↓ 0.2)	77.9(↓ 2.2)
YOLOv10n	2.26(↑ 0.41)	6.5(↑ 0.9)	84.4(↓ 3.7)	85.9(↓ 1.8)	75.3(↓ 4.8)
YOLOv10s	7.22(↑ 5.37)	21.4(↑ 15.8)	86.0(↓ 2.1)	86.8(↓ 0.9)	77.8(↓ 2.3)
本文算法	1.85	5.6	88.1	87.7	80.1

3.4.2 消融实验

在 GDUT-HWD 数据集上进行消融实验,结果见表 3。本文算法同时结合 4 个创新点,Parameters、FLOPs 不高于单个创新点、2 个创新点组合和 3 个创新点组合,比基线算法分别降低 18.1%、13.8%;mAP@0.5、Precision 和 Recall 高于单个创新点、2 个创新点组合和 3 个创新点组合,比基线

算法分别提升 3.7%、1.8%、4.8%,验证了各模块的协同有效性。

在 SHEL5K 数据集上进行消融实验,结果见表 4。本文算法同时结合 4 个创新点,Parameters、FLOPs 不高于单个创新点、2 个创新点组合和 3 个创新点组合,比基线算法分别降低 17.6%、13.8%;mAP@0.5、Precision 和 Recall 高于单个

Table 2 Performance of different algorithms on SHEL5K dataset**表 2 不同算法在 SHEL5K 数据集上的表现**

算法	Parameters/M	FLOPs/G	mAP@0.5/%	Precision/%	Recall/%
Cascade-RCNN	68.76(↑ 66.89)	79.5(↑ 73.9)	85.9(↓ 3.9)	87.7(↓ 0.6)	77.2(↓ 4.4)
SSD	23.49(↑ 21.62)	52.7(↑ 46.1)	82.4(↓ 7.4)	87.4(↓ 0.9)	76.9(↓ 4.7)
YOLOv5n	1.77(↓ 0.10)	4.2(↓ 1.4)	83.3(↓ 6.5)	87.9(↓ 0.4)	75.4(↓ 6.2)
YOLOv8n	3.01(↑ 1.14)	8.1(↑ 2.5)	86.2(↓ 3.6)	87.9(↓ 0.4)	79.5(↓ 2.1)
YOLOv8s	11.13(↑ 9.26)	28.4(↑ 22.8)	87.6(↓ 2.2)	87.6(↓ 0.7)	77.7(↓ 3.9)
YOLOv9s	7.13(↑ 5.26)	26.5(↑ 20.9)	87.7(↓ 2.1)	87.7(↓ 0.6)	78.1(↓ 3.5)
YOLOv10n	2.27(↑ 0.40)	6.5(↑ 0.9)	86.8(↓ 3.0)	85.3(↓ 3.0)	79.0(↓ 2.6)
YOLOv10s	7.22(↑ 5.35)	21.4(↑ 15.8)	87.1(↓ 2.7)	86.6(↓ 1.7)	80.0(↓ 1.6)
本文算法	1.87	5.6	89.8	88.3	81.6

Table 3 Ablation experiment on GDUT-HWD dataset**表 3 在 GDUT-HWD 数据集上的消融实验**

C2f-C	SPPF-L	BiFPN-W	W-MPDIoU	Parameters/M	FLOPs/G	mAP@0.5/%	Precision/%	Recall/%
-	-	-	-	2.26	6.5	84.4	85.9	75.3
√	-	-	-	1.91(↓ 0.35)	5.7(↓ 0.8)	85.2(↑ 0.8)	86.3(↑ 0.4)	77.4(↑ 2.1)
-	√	-	-	2.53(↑ 0.27)	6.7(↑ 0.2)	85.8(↑ 1.4)	86.8(↑ 0.9)	76.2(↑ 0.9)
-	-	√	-	1.87(↓ 0.39)	5.9(↓ 0.6)	85.3(↑ 0.9)	86.4(↑ 0.5)	75.7(↑ 0.4)
-	-	-	√	2.26(↓ 0.00)	6.5(↓ 0.0)	85.1(↑ 0.7)	86.9(↑ 1.0)	76.7(↑ 1.4)
√	√	-	-	2.21(↓ 0.05)	6.4(↓ 0.1)	86.7(↑ 2.3)	86.9(↑ 1.0)	78.2(↑ 2.9)
√	√	√	-	1.85(↓ 0.41)	5.6(↓ 0.9)	87.3(↑ 2.9)	87.0(↑ 1.1)	78.7(↑ 3.4)
√	√	√	√	1.85(↓ 0.41)	5.6(↓ 0.9)	88.1(↑ 3.7)	87.7(↑ 1.8)	80.1(↑ 4.8)

创新点、2个创新点组合和3个创新点组合,比基线算法分

别提升3.0%、1.0%、2.6%,验证了各模块的协同有效性。

Table 4 Ablation experiment on SHEL5K dataset**表 4 在 SHEL5K 数据集上的消融实验**

C2f-C	SPPF-L	BiFPN-W	W-MPDIoU	Parameters/M	FLOPs/G	mAP@0.5/%	Precision/%	Recall/%
-	-	-	-	2.27	6.5	86.8	87.3	79.0
√	-	-	-	1.91(↓ 0.36)	5.7(↓ 0.8)	87.5(↑ 0.7)	87.4(↑ 0.1)	79.2(↑ 0.2)
-	√	-	-	2.54(↑ 0.27)	6.7(↑ 0.2)	87.7(↑ 0.9)	87.5(↑ 0.2)	80.0(↑ 1.0)
-	-	√	-	1.88(↓ 0.39)	5.9(↓ 0.6)	87.6(↑ 0.8)	87.6(↑ 0.3)	79.8(↑ 0.8)
-	-	-	√	2.27(↓ 0.00)	6.5(↓ 0.0)	87.4(↑ 0.6)	87.9(↑ 0.6)	79.7(↑ 0.7)
√	√	-	-	2.21(↓ 0.06)	6.4(↓ 0.1)	88.1(↑ 1.3)	87.6(↑ 0.3)	80.1(↑ 1.1)
√	√	√	-	1.87(↓ 0.40)	5.6(↓ 0.9)	89.1(↑ 2.3)	87.8(↑ 0.5)	80.8(↑ 1.8)
√	√	√	√	1.87(↓ 0.40)	5.6(↓ 0.9)	89.8(↑ 3.0)	88.3(↑ 1.0)	81.6(↑ 2.6)

3.4.3 泛化实验

在 SHWD 数据集上对基线算法 YOLOv10n 和本文算法进行泛化性实验,结果见表 5。可以看出,本文算法的 Parameters 和 FLOPs 均低于基线算法, mAP@0.5、Precision 和 Recall 均高于基线算法,表明本文算法泛化性良好。

Table 5 Generalization experiment result**表 5 泛化性实验结果**

算法	Parameters/M	FLOPs/G	mAP@0.5/%	Precision/%	Recall/%
基线算法	2.26	6.5	93.2	91.8	87.4
YOLOv10n	1.84(↓ 0.42)	5.6(↓ 0.9)	95.2(↑ 2.0)	93.5(↑ 1.7)	89.7(↑ 2.3)

4 结语

针对复杂场景下安全帽检测算法精度与参数量失衡的问题,本文提出一种基于 YOLOv10n 的丰富上下文捕捉多层特征的安全帽检测算法。该算法通过结合 CFBLOCK 与 CGLU 来置换 C2f 中的 Bottleneck;采用 CFBLOCK 调整空间维度与通道维度,增强模型表达能力,降低参数量和计算量;

采用 CGLU 控制信息流动,丰富全局上下文信息,提高检测精度;在 SPPF 结构中引入 LSKA 注意力机制,扩大卷积核感受野,关注重要空间特征,提高检测精度;采用 Weighted Fusion 改进 BiFPN 融合方式替换颈部网络,双向连接多尺度信息,轻量化加权融合多个特征图,降低参数量和计算量;采用 WIoU 与 MPDIoU 的融合函数替换 CIoU,缓解正负样本不平衡,增强泛化能力。实验结果表明,本文算法的综合性能优于基线算法和参考文献算法,在不过多增加参数量和计算量的同时有效提升了检测精度等,且具有良好的鲁棒性。未来研究将引入更有效的特征提取模块来提高检测精度,并通过剪枝技术进一步降低模型参数量和计算量,实现安全帽检测算法的实时性部署。

参考文献:

- [1] 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, 2016: 779-788.
- [2] REDMON J, FARHADI A. YOLO9000: better, faster, stronger[C]//Pro-

- ceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017: 7263–7271.
- [3] REDMON J, FARHADI A. Yolov3: an incremental improvement [DB/OL]. <https://arxiv.org/abs/1804.02767>.
- [4] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. Yolov4: optimal speed and accuracy of object detection [DB/OL]. <https://arxiv.org/abs/2004.10934>.
- [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 [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 7464–7475.
- [6] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot multibox detector [C]//14th European Conference on Computer Vision, 2016: 21–37.
- [7] CAI Z W, VASCONCELOS N. Cascade R-CNN: delving into high quality object detection [C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018: 6154–6162.
- [8] DU X G, WANG Y Q, YAN R B, et al. Accurate detection algorithm for helmet wearing based on YOLO-ST [J]. Journal of Shaanxi University of Science and Technology, 2022, 40(6): 177–183, 191.
杜晓刚, 王玉琪, 晏润冰, 等. 基于 YOLO-ST 的安全帽佩戴精确检测算法 [J]. 陕西科技大学学报, 2022, 40(6): 177–183, 191.
- [9] FU D L, GAO L, LIU W, et al. Detection of key safety equipment for electric workers based on improved YOLOv5 algorithm [J]. Journal of Hubei Minzu University (Natural Science Edition), 2022, 40(3): 320–327.
伏德粟, 高林, 刘威, 等. 基于改进 YOLOv5 算法的电力工人作业安全关键装备检测 [J]. 湖北民族大学学报 (自然科学版), 2022, 40(3): 320–327.
- [10] QI Z Z, XU Y X. Research on helmet wearing detection of improved YOLOv5s algorithm [J]. Computer Engineering and Applications, 2023, 59(14): 176–183.
祁泽政, 徐银霞. 改进 YOLOv5s 算法的安全帽佩戴检测研究 [J]. 计算机工程与应用, 2023, 59(14): 176–183.
- [11] BAI P R, WANG R, LIU Q Y, et al. DS-YOLOv5: a real-time detection and recognition model for helmet wearing [J]. Chinese Journal of Engineering, 2023, 45(12): 2108–2117.
白培瑞, 王瑞, 刘庆一, 等. DS-YOLOv5: 一种实时的安全帽佩戴检测与识别模型 [J]. 工程科学学报, 2023, 45(12): 2108–2117.
- [12] ZHAO M, YANG G L, WANG J X, et al. Improved YOLOv7-tiny algorithm for safety helmet real-time detection [J]. Radio Engineering, 2023, 53(8): 1741–1749.
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.
- [13] YANG D W, ZHANG C C. Algorithm of safety helmet wearing detection based on improved YOLOv7 [J]. Journal of Shenyang Ligong University, 2024, 43(1): 16–21.
杨大为, 张成超. 基于改进 YOLOv7 的安全帽佩戴检测算法 [J]. 沈阳理工大学学报, 2024, 43(1): 16–21.
- [14] LING G, ZHAO J, MO D J, et al. Wearing safety helmet detection undermine based on improved YOLOv8s [J/OL]. Computer Systems & Applications, 2024–12–20. <http://www.c-s-a.org.cn/1003-3254/9743.html>.
凌港, 赵杰, 莫定界, 等. 基于改进 YOLOv8s 的矿井下安全帽佩戴检测 [J/OL]. 计算机系统应用, 2024–12–20. <http://www.c-s-a.org.cn/1003-3254/9743.html>.
- [15] FENG Y, YANG S Z, XU H Y. Light weighted safety helmet wearing detection algorithm based on YOLOv8 [J/OL]. Journal of Computer Applications, 2024–09–01. <http://kns.cnki.net/kcms/detail/51.1307.TP.20240418.1338.002.html>.
冯勇, 杨思卓, 徐红艳. 基于 YOLOv8 的轻量化安全帽佩戴检测算法 [J/OL]. 计算机应用, 2024–09–01. <http://kns.cnki.net/kcms/detail/51.1307.TP.20240418.1338.002.html>.
- [16] WANG X L, WANG L Y. Helmet detection algorithm based on multiple receptive field enhancement in blasting scene [J/OL]. Computer Engineering and Applications, 2024–7–19. <http://kns.cnki.net/kcms/detail/11.2127.TP.20240129.1050.012.html>.
王新良, 王璐莹. 多感受野增强的爆破现场安全帽检测算法 [J/OL]. 计算机工程与应用, 2024–07–19. <http://kns.cnki.net/kcms/detail/11.2127.TP.20240129.1050.012.html>.
- [17] YU W, SI C, ZHOU P, et al. MetaFormer baseline for vision [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022, 46(2): 896–912.
- [18] SHI D. TransNeXt: robust foveal visual perception for vision transformers [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2024: 17773–17783.
- [19] LAU K W, LAI-MAN P, REHMAN Y A U. Large separable kernel attention: rethinking the large kernel attention design in CNN [J]. Expert Systems with Applications, 2024, 236: 121352.
- [20] TONG Z, CHEN Y, XU Z, et al. Wise-IoU: bounding box regression loss with dynamic focusing mechanism [DB/OL]. <https://arxiv.org/abs/2301.10051>.
- [21] MA S L, XU Y. MPDIoU: a loss for efficient and accurate bounding box regression [DB/OL]. <https://arxiv.org/abs/2307.07662>.
- [22] TAN M X, PANG R M, LE Q V. EfficientDet: scalable and efficient object detection [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 10781–1079.

(责任编辑:尹晨茹)