

# 光电工程

## Opto-Electronic Engineering

中文核心期刊 中国科技核心期刊  
Scopus CSCD

### THI-YOLO: 改进YOLOv8的非机动车驾驶员头盔检测

孙光灵, 王薪博

#### 引用本文:

孙光灵, 王薪博. THI-YOLO: 改进YOLOv8的非机动车驾驶员头盔检测[J]. *光电工程*, 2024, 51(12): 240244.

Sun G L, Wang X B. THI-YOLO: Improved non-motorized drivers helmet detection of YOLOv8[J]. *Opto-Electron Eng*, 2024, 51(12): 240244.

<https://doi.org/10.12086/oe.2024.240244>

收稿日期: 2024-10-18; 修改日期: 2024-11-15; 录用日期: 2024-11-16

### 相关论文

#### 基于改进YOLOv5的交通标志识别方法

曲立国, 张鑫, 卢自宝, 刘玉玲, 陈国豪

*光电工程* 2024, 51(6): 240055 doi: 10.12086/oe.2024.240055

#### PAW-YOLOv7: 河道微小漂浮物检测算法

栾庆磊, 常昕昱, 吴叶, 邓从龙, 史艳琼, 陈梓华

*光电工程* 2024, 51(4): 240025 doi: 10.12086/oe.2024.240025

#### 图像引导和点云空间约束的公路洒落物检测定位方法

蔡怀宇, 杨朝乾, 崔子扬, 汪毅, 陈晓冬

*光电工程* 2024, 51(3): 230317 doi: 10.12086/oe.2024.230317

更多相关论文见光电期刊集群网站 



<http://cn.ojournal.org/oe>



 OE\_Journal



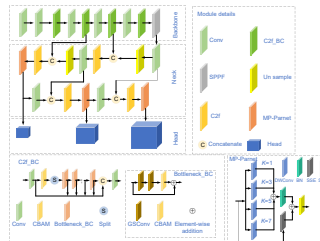
Website

# THI-YOLO: 改进 YOLOv8 的非机动车驾驶员头盔检测

孙光灵<sup>1,2\*</sup>, 王薪博<sup>1</sup>

<sup>1</sup>安徽建筑大学电子与信息工程学院, 安徽 合肥 230601;

<sup>2</sup>安徽建筑大学安徽省古建筑智能感知与高维建模国际联合研究中心, 安徽 合肥 230601



**摘要:** 针对非机动车驾驶员头盔检测任务中常存在复杂背景以及检测场景中存在的检测目标尺寸大小不一的现象, 进而导致检测效率低和误检漏检的问题, 提出了一种面向非机动车驾驶员的头盔检测的 YOLOv8 算法。在 C2f 模块中结合 GSConv 和 CBAM 的优点, 设计 C2f\_BC 模块, 在降低模型参数量的同时, 有效提升了模型的特征提取能力。设计了多核并行感知网络 (MP-Parnet), 提高了模型对不同尺度目标的感知和特征提取能力, 使其能更好地应用到复杂场景中。为缓解复杂场景出现的正负样本不平衡等问题, 在原模型损失函数 CIoU 基础上引入 Focaler-IoU, 引入阈值参数来改进 IoU 损失计算方式, 从而缓解正负样本的不平衡的现象, 有效提升了模型在复杂背景下目标框定位的准确性。实验结果表明, 改进的 YOLOv8n 相较于原模型, 在保持参数量下降的同时, mAP50 和 mAP50 : 95 在自建数据集 Helmet 上分别提升了 2.2% 和 1.9%, 在开源数据集 TWHd 上分别提升了 1.8% 和 1.9%, 说明改进的模型可以更好地应用到非机动车驾驶员的头盔检测场景。

**关键词:** 目标检测; YOLOv8; 注意力机制; 损失函数; 特征提取

**中图分类号:** TP391.41; U495

**文献标志码:** A

孙光灵, 王薪博. THI-YOLO: 改进 YOLOv8 的非机动车驾驶员头盔检测 [J]. 光电工程, 2024, 51(12): 240244

Sun G L, Wang X B. THI-YOLO: Improved non-motorized drivers helmet detection of YOLOv8[J]. *Opto-Electron Eng*, 2024, 51(12): 240244

## THI-YOLO: Improved non-motorized drivers helmet detection of YOLOv8

Sun Guangling<sup>1,2\*</sup>, Wang Xinbo<sup>1</sup>

<sup>1</sup>School of Electronic and Information Engineering, Anhui Jianzhu University, Hefei, Anhui 230601, China;

<sup>2</sup>Anhui Interational Joint Research Center for Intelligent Perception and High-dimensional Modeling of Ancient Buildings, Anhui Jianzhu University, Hefei, Anhui 230601, China

**Abstract:** Aiming at the phenomenon that complex background often exists in non-motorized drivers' helmet detection and the diversity of detection target scales often exists in the detection scene, which in turn leads to the low detection efficiency and misdetection and omission, a YOLOv8 algorithm oriented to the detection of traffic

收稿日期: 2024-10-18; 修回日期: 2024-11-15; 录用日期: 2024-11-16

基金项目: 国家自然科学基金资助项目 (62001004); 中国建设教育协会教育教学科研立项课题 (2023069); 2023 年安徽省住房城乡建设科学技术计划项目 (2023-YF058, 2023-YF113); 安徽省高等学校科学研究重点项目 (2023AH050164); 安徽省高校杰出青年科研项目 (2023AH020022)

\*通信作者: 孙光灵, sunguangling@163.com。

版权所有©2024 中国科学院光电技术研究所

helmets is proposed. Combining the advantages of GSConv and CBAM in the C2f module, the C2f\_BC module is designed to effectively improve the feature extraction capability of the model while reducing the number of model parameters. A multi-core parallel perception network (MP-Parnet) is designed to improve the model's perception and feature extraction ability for multi-scale targets so that it can be better applied to complex scenes. To alleviate the problem of positive and negative sample imbalance in complex scenes, Focaler-IoU is introduced based on the original model's loss function CIoU, and a threshold parameter is introduced to improve the calculation of the IoU loss, thus alleviating the phenomenon of positive and negative sample imbalance, and effectively improves the model's accuracy of target frame localization in complex background. The experimental results show that compared with the original model, the improved YOLOv8n maintains a decrease in the number of parameters while the mAP50 and mAP50: 95 increase by 2.2% and 1.9% on the self-built dataset Helmet, and 1.8% and 1.9% on the open-source dataset TWHD, which suggests that the improved model can be better applied to the helmet detection of non-motorized drivers in the scenario.

**Keywords:** object detection; YOLOv8; attention mechanism; loss function; feature extraction

## 1 引言

随着交通的不断变迁<sup>[1-2]</sup>, 双轮车已经成为我国公民出行的主要交通工具, 根据行业协会和公安部公布的数据显示, 我国电动自行车的保有量已达 3.5 亿辆, 2023 年度全国规模以上企业累计生产电动自行车 4228 万辆, 均居世界第一, 但是公民闯红灯、不佩戴安全头盔等现象仍屡见不鲜, 交通事故发生率也在攀升, 而正确佩戴头盔能够有效地保护头部, 将交通事故的死亡率降低 60%~70%。因此设计一种非机动车骑行人员头盔佩戴的检测方法尤为迫切。现阶段对非机动车驾驶员的头盔检测仍存在一些难题, 比如交通环境复杂、骑行人员密集等现象, 导致检测中经常存在复杂的背景信息以及检测目标尺度丰富等难以检测问题, 因此如何设计更高性能的头盔检测模型需要展开进一步研究。

随着近年来计算机视觉的逐渐兴起, 目标检测算法也在不断创新发展, 因此关于非机动车驾驶员的安全头盔检测也是研究学者关注的重点之一。基于神经网络的目标检测算法主要分为双阶段的 Faster-RCNN (fast region-convolutional neural network)<sup>[3]</sup> 算法、单阶段 SSD (single shot MultiBox detector)<sup>[4]</sup> 和 YOLO (you only look once)<sup>[5]</sup> 算法。双阶段的检测算法在处理图像时会生成候选区域, 检测模型对划分的每个区域进行边界框回归和类别的鉴定。单阶段的目标检测算法会在图像中直接生成预测框并检测出相应的类别, 检测速度相对较快, 适用于实时的目标检测。本文选择 YOLO 算法进行改进, 并应用于交通道路头盔检测的实时任务。

目前, 许多学者在头盔检测领域中做了相应的研究。刘超等人<sup>[6]</sup> 提出了 YOLOv5-MH, 引入空洞卷积和注意力机制模块来减少特征提取的网络参数量并优化检测结果, 但是检测类别过于单一且模型的参数量较大。王茹等人<sup>[7]</sup> 将 MobileViTv3 模块融入到主干网络中, 通过捕获长距离的依赖关系, 并与局部信息相融合, 从而提升模型在安全帽场景的检测精度, 但检测场景较为简单。王鹏飞等人<sup>[8]</sup> 借鉴了双向特征金字塔网络的特征融合思想, 提出了去权重的双向金字塔网络, 加强了特征融合, 减少了卷积过程中信息的丢失, 但在极端环境条件下对于小目标的检测效率不高。周顺勇等人<sup>[9]</sup> 设计了 PCAHead 检测头和 HelmetIoU, 优化了模型处理数据能力, 实现复杂道路场景中更加准确地识别和关注头盔目标, 但模型并未作轻量化处理, 虽然精度有所提升, 但是模型的参数量并未得到优化。

为了应对上述问题, 本文提出了一种针对复杂场景下交通道路头盔的检测算法, 在保证参数量下降的同时提高检测精度。本文工作:

1) 设计了 C2f\_BC 模块, 引入了具有组合分组卷积和空间卷积的思想的 GSConv<sup>[10]</sup> 模块, 并在 C2f 中的 Bottleneck 中引入结合通道和空间的注意力机制 (convolutional block attention module, CBAM)<sup>[11]</sup>, 有效地降低计算的复杂度, 并增强局部和全局特征的提取能力。

2) 设计了并行多尺度特征融合模块 MP-Parnet (parallel multiscale perception networks), 对 Parnet (parallel networks)<sup>[12]</sup> 进行重新设计, 采用不同尺度的

并行深度可分离卷积 (depthwise separable convolution, DWConv)<sup>[13]</sup> 核代替原模块的普通卷积, 有效地适应不同尺度目标的获取能力。

3) 参照 Focaler-IoU<sup>[14]</sup> 的思想, 将原模型的 CIoU 替换为 Focaler-CIoU, 缓解复杂场景下正负样本不平衡的问题, 有效提高了目标框定位的准确性, 进而提高检测精度。

## 2 YOLOv8 模型介绍

YOLOv8 是目前主流的单阶段目标检测算法, 包括三个部分: 主干网络 (backbone)、颈部网络 (neck) 和检测头部 (head)。YOLOv8 根据模型的参数量可分为 YOLOv8n、YOLOv8s、YOLOv8m 和 YOLOv8l。出于模型参数量考虑, 本文选取最轻量化的模型 YOLOv8n 作为基准模型进行实验。

YOLOv8 的主干部分 (backbone) 主要由 Conv 模块、C2f 模块和 SPPF 模块组成。其中 Conv 用于降低特征空间维度并增加特征通道数量; C2f 模块是 YOLOv5<sup>[15]</sup> 模块中 C3 结构的进一步优化, 通过特征分流和融合的方式, 能够提取更加丰富的特征, 并保持较小的计算复杂度; SPPF 是一种空间金字塔池化模块, 通过不同尺度的卷积块增强特征的多尺度感知能力。

颈部 (Neck) 网络主要由 Upsample 模块和 Concat 模块组成。其中 Upsample 是上采样操作, 增加特征图的空间维度以便于不同层次的特征图进行特征融合; Concat 模块用来结合不同尺度的信息, 形成多尺度的特征融合, 使模型可以同时利用来自不同尺度的特征。颈部网络的 C2f 结构和主干网络类似, 用于特征的提取和融合。颈部网络采用了 FPN 的设计理念, 融合不同层次的特征, 以此增强模型的表达能力。

检测头 (Head) 部分用来输出最终的目标预测结果。在 YOLOv5 中, Anchor-based 需要为不同尺度和形状的物体手动设置多个先验框, 设置不当可能会影响检测效果, 但 YOLOv8 中的 Anchor-free 则跳过了这个过程, 直接预测目标的位置和类别。此外, 为了配合 Anchor-free, 以及提升泛化性, 在检测头部中引入了 DFL (distribution focal loss) 损失, 以交叉熵的形式对标签进行优化, 使网络可以更快地聚焦到目标的邻近区域。检测头部的改进有效地降低了计算复杂度, 从而加快了推理速度。

## 3 THI-YOLO 算法

为了应对非机动车驾驶员头盔检测中出现的漏检、误检、遮挡以及复杂检测场景下目标尺度大小不一所带来的检测效率低的问题, 本文提出了 THI-YOLO 模型, 用以提高不同尺度目标的检测性能, 减少误检和漏检问题。THI-YOLO 模型结构如图 1 所示, 本文设计了 C2f-BC 模块, 颈部网络设计了 MP-Parnet 模块, 在损失函数方面引入了 Focaler-CIoU。

主干网络的 C2f 模块中, 在 Bottleneck 中使用了 GSConv 进行卷积, 并且将 CBAM 引入到 Bottleneck 中, 得到 C2f\_BC 模块, 在保证参数量降低的同时有效地增强了主干网络的特征处理能力。在颈部网络特征提取阶段引入改进的 MP-Parnet, 设置不同尺度的卷积特征融合有效地解决了复杂场景中不同尺度目标检测率低的问题。将 Focaler-IoU 引入到原模型的 CIoU 中, 有效提高了目标框定位的准确性, 缓解正负样本不平衡的现象。

### 3.1 C2f\_BC

C2f 模块在计算效率和特征提取方面取得不错的效果, 但也存在一些不足。C2f 的两条路径包括直接路径和 Bottleneck 路径, 虽然残差结构可以有效缓解梯度消失和梯度爆炸问题, 但是残差路径在前后层中存在相似性, 这增加了计算的冗余度, 并且 Bottleneck 的局部卷积特性使得特征图中的全局依赖关系没有被充分利用, 在不断的下采样过程中 C2f 模块对小目标的特征提取效果不佳, 导致检测性能较低。本文设计的 C2f\_BC 有效地提升了模型在复杂场景下的特征提取性能, 同时参数量有所下降。

设计的思路是将 C2f 中的 Bottleneck 层中的普通卷积替换成组合分组卷积和空间卷积的思想的 GSConv, 降低参数量的同时有效地增强特征提取的效率。在 C2f 的起始卷积后和 Bottleneck 的第二个 GSConv 后增加 CBAM, CBAM 融合了空间注意力和通道注意力, 有效缓解了 C2f 模块对远距离特征交互的不足。

#### 3.1.1 GSConv 模块

在标准的卷积模块中, 卷积层的特征图中经常存在大量的冗余信息, 随着网络深度的加深, 计算复杂度也不断增加。GSConv 是一种结合分组卷积 (GhostConv)<sup>[16]</sup> 和 Shuffle 操作 (特征通道混洗, channel

shuffle)<sup>[17]</sup> 的一个卷积模块, 如图 2 所示。将输入特征按照 1 : 1 的比例一分为二, 其中  $C/2$  通道的特征进行普通卷积, 另外  $C/2$  通道的特征渗透到 DWConv 中。Shuffle 操作旨在解决分组卷积中特征通道之间

信息交流不足的问题。Shuffle 操作通过重新排列特征图的通道, 使得不同组之间的特征信息可以进行更充分的交互, 从而提高网络的表达能力。Shuffle 通道混洗操作实现通道间的特征信息融合, 有效地增强

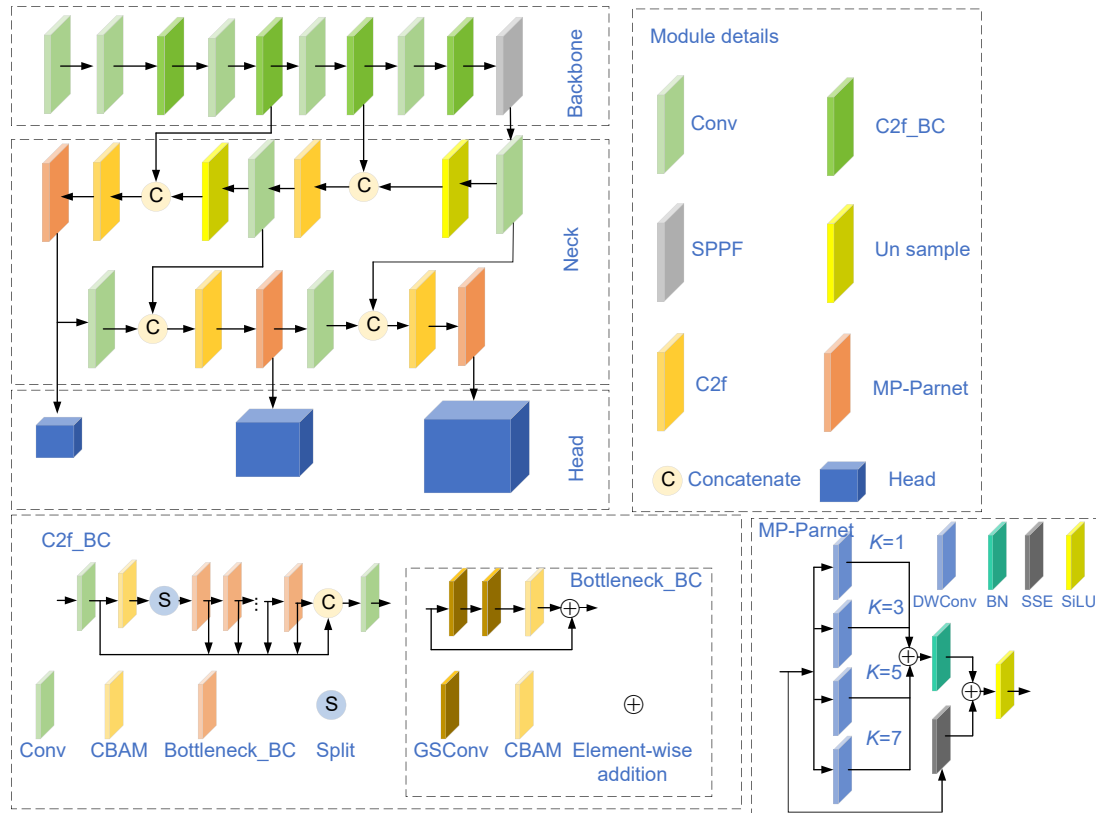


图 1 THI-YOLO 模型  
Fig. 1 THI-YOLO model

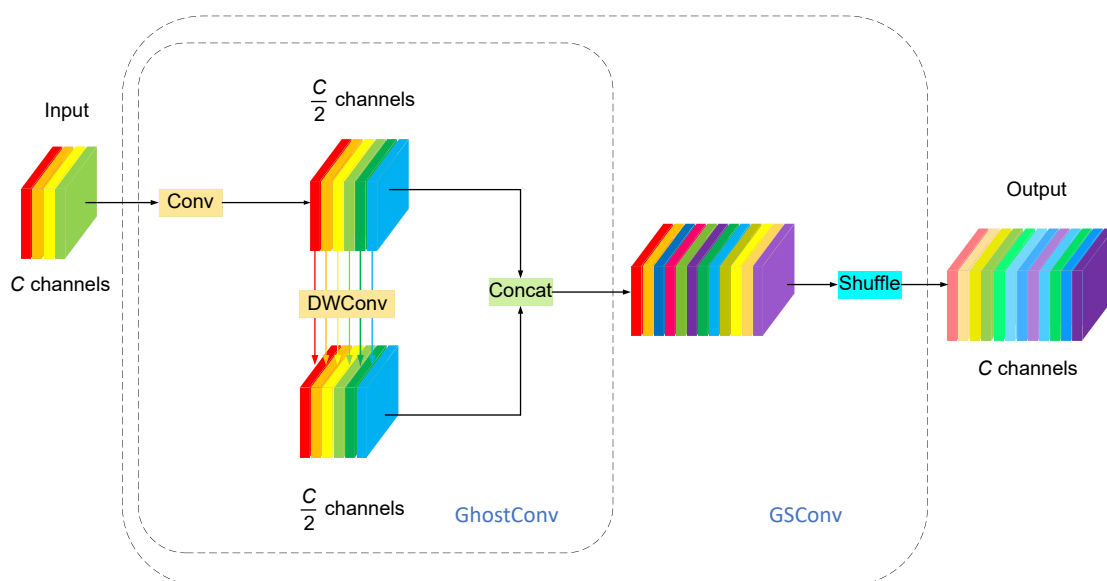


图 2 GSConv 模块结构图  
Fig. 2 Structure of GSConv module

了特征之间的交互。这种设计方式不仅增强了网络的推理效率和精度，而且进一步减少了计算的冗余，提高计算效率。

### 3.1.2 CBAM

非机动车驾驶员的头盔检测任务中经常存在复杂的背景特征，且在不断的下采样过程中很容易丢失关键的信息，为了增强模型对复杂道路场景中的关键信息捕获能力，本文引入了 CBAM，CBAM 结合了通道和空间注意力，使模型能够专注于处理重要的特征区域，同时忽略不相关的特征信息，有效地增强模型的表达能力。CBAM 结构如图 3 所示。

CBAM 主要由两个子模块组成，分别为通道注意力模块 CAM (channel attention module) 和空间注意力模块 SAM (spatial attention module)。通道注意力模块能够增强通道维度上的重要特征，同时舍弃一些不重要的特征。空间注意力模块的作用是找出图像中包含重要特征的位置，并在空间维度上进行注意力权重分配，从而使网络更加专注于重要的空间位置。

通道注意力子模块如图 4 所示，首先将输入特征  $F$  进行全局平均池化 (Avgpool) 和全局最大池化 (Maxpool)，经这两种池化方式处理后分别生成通道

上的全局信息，再将两个池化的结果分别通过共享的多层感知机 (MLP) 进行非线性变换，得到通道注意力权重  $M_c(F)$ ，并利用通道注意力权重  $M_c(F)$  对输入特征图进行加权并调整特征图的通道权重。如式 (1)、式 (2) 所示。

$$M_c(F) = \delta(MLP(Avgpool(F)) + MLP(Maxpool(F))), \quad (1)$$

$$F' = M_c(F) \otimes F, \quad (2)$$

其中： $\otimes$ 表示逐元通道乘法， $\delta$ 表示 Sigmoid 激活函数， $F'$ 表示通道注意力输出模块。

空间注意力子模块如图 5 所示，首先将通道注意力模块的输出特征图  $F'$  进行通道上的最大池化和平均池化，生成压缩后的特征图，再通过卷积操作生成每个空间位置的权重  $M_s(F')$ ，并使用计算得到的空间注意力权重  $M_s(F')$  对通道注意力模块的输出特征图  $F'$  进行加权得到最终经过 CBAM 处理后的特征图  $F''$ 。如式 (3)、式 (4)。

$$M_s(F') = \delta(Conv_{7 \times 7}(Avgpool(F'), Maxpool(F'))), \quad (3)$$

$$F'' = M_s(F') \otimes F'. \quad (4)$$

CBAM 有效地融合了空间和通道的特征，使模

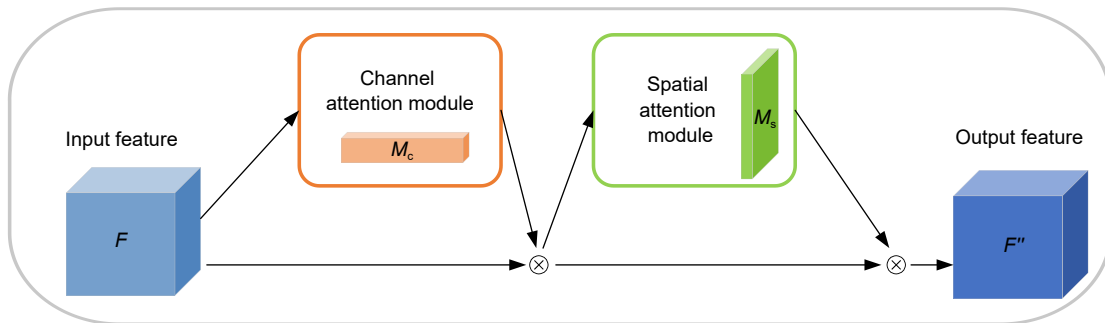


图 3 CBAM 结构图

Fig. 3 CBAM structure

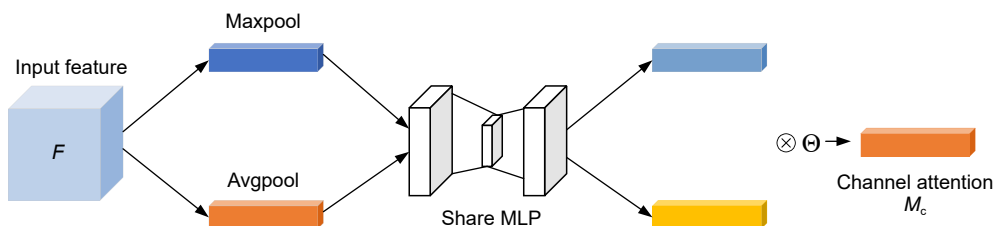


图 4 通道注意力子模块

Fig. 4 Channel attention submodule

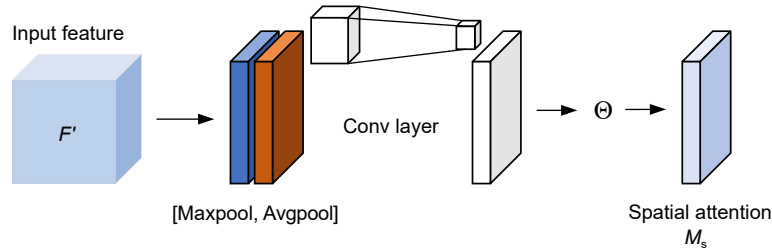


图 5 空间注意力子模块  
Fig. 5 Spatial attention submodule

型能更好地区分背景和目标并有效地捕获到关键信息, 以此提升模型的特征提取性能。

### 3.1.3 C2f\_BC 模块的设计

设计 Bottleneck\_GC 替换 C2f 中的 Bottleneck, 并在 C2f 输入口的卷积层后添加一个 CBAM 得到 C2f\_BC, 如图 6 所示。

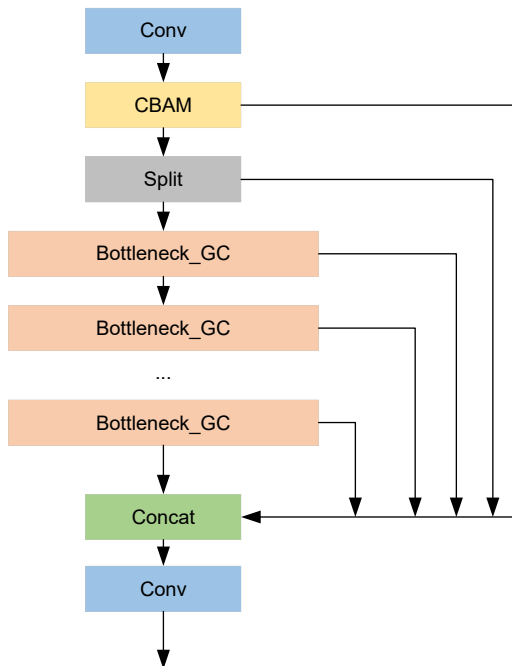


图 6 C2f\_BC 结构图  
Fig. 6 Diagram of the C2f\_BC structure

Bottleneck\_GC 模块的设计是在 Bottleneck 的基础上将普通卷积 (Conv) 替换成 GSConv, 并在第二个 GSConv 后添加 CBAM, 得到 Bottleneck\_GC, 如图 7 所示。

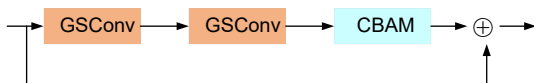


图 7 Bottleneck\_GC 结构图  
Fig. 7 Diagram of the Bottleneck\_GC structure

### 3.2 MP-Parnet 设计

为了增强模型对多种尺度目标的特征提取能力, 使其更好地应用于骑行人员头盔检测的场景, 本文在并行化网络 (parallel networks) 的基础上设计了模块 MP-Parnet, 如图 8 所示。

由于从主干网络到颈部网络, 每一层得到的特征也在不断细化, 在颈部网络加入并行化网络 (Parnet) 可以进一步增强模型的提取特征能力, 虽然  $1 \times 1$  的卷积核和  $3 \times 3$  的卷积核有助于处理应用场景中的多种尺度目标, 但是卷积核的尺度还不够丰富, 无法灵活应用在尺度丰富的检测任务中, 尤其是在复杂场景下, 捕获全局和局部的多尺度特征信息时表现不佳。由于大的卷积核可以获取更大的感受野, 小的卷积核可以更加细化地处理局部的细节特征。因此本文在 Parnet 的基础上将  $1 \times 1$ 、 $3 \times 3$ 、 $5 \times 5$ 、 $7 \times 7$  的并行处理的深度可分离卷积 (DWConv) 替换为原模型的标准卷积 (Conv) 进行提取和融合特征。使用 DWConv 能够将空间维度和通道维度的卷积操作分开, 从而显著减少计算量, 通过设置不同卷积核大小的 DWConv, 不仅可以在不同的空间尺度上提取更加多样化的特征, 而且可以捕捉不同尺度的信息, 使得网络在处理尺度变化大的目标时更具鲁棒性。单纯的多尺度特征提取可能会导致一些不必要的冗余特征, SSE (skip-squeeze-excitation) 模块是基于 SENet (squeeze-and-excitation networks) 设计的, 在多尺度特征融合之前, SSE 模块通过全局平均池化操作为每个空间位置的特征进行加权, 从而过滤掉不重要的区域, 增强与目标相关的重要区域。该模块的引入可以有效增强模型对输入特征的信息感知和特征表达能力。

### 3.3 损失函数优化

在交通道路头盔检测场景中存在背景复杂、尺度目标丰富等情况, 导致负样本的 IoU 非常低。传统 IoU 函数也会计算这些低 IoU 的负样本, 导致训练花

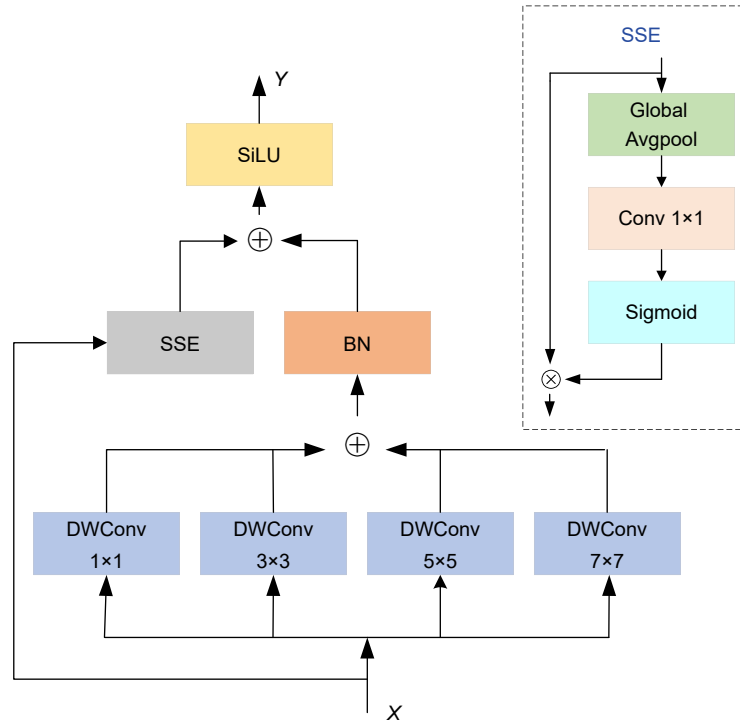


图 8 MP-ParNet 结构图

Fig. 8 Diagram of the MP-ParNet structure

费较大的时间，并且在多尺度目标场景中，小目标的 IoU 较低，大目标的 IoU 较高，传统的 IoU 函数可能会忽视一些具有重要特征的小目标。为了缓解这种正负样本不平衡以及让模型更容易关注到重要特征的小目标，将 Focaler-IoU 引入到原模型使用的 CIoU 中，进一步提升模型的检测性能。Focaler-CIoU 公式如下：

$$L_{Focaler-CIoU} = L_{CIoU} + IoU - (IoU_{Focaler}), \quad (5)$$

其中： $L_{CIoU}$  为 CIoU 损失函数， $IoU$  表示 IoU 损失值， $IoU_{Focaler}$  的定义如下：

$$IoU_{Focaler} = \begin{cases} 0, & IoU < d \\ \frac{IoU - d}{u - d}, & d \leq IoU \leq u \\ 1, & IoU > u \end{cases}, \quad (6)$$

其中， $IoU < d$  表示当  $IoU$  小于  $d$  时， $IoU_{Focaler}$  的值为 0，这种情况反映出预测框和真实框之间的重叠区域过小，忽略这部分  $IoU$  的计算，减少了对冗余样本的计算，有利于模型聚焦在目标样本中； $IoU > u$  表示  $IoU_{Focaler}$  的值为 1，这种情况反映预测框和真实框之间的重叠区域较好，因已经足够接近最终目标，不需要再优化  $IoU$  损失；当  $d \leq IoU \leq u$  时， $IoU_{Focaler}$  根据  $IoU$  插值计算损失，表示预测框和

真实框部分重叠，根据  $IoU$  的值灵活调整阈值  $d$  和  $u$  的值，提升小目标的权重，使模型能够兼顾大小不同尺度的目标获取，提升模型的特征提取能力。

在原模型的 CIoU 的基础上设计 Focaler-CIoU 可以有效地调节正负样本不平衡以及多目标样本的检测问题。

## 4 实验与结果分析

### 4.1 数据集介绍

现阶段对骑行人员的头盔检测的数据集应用场景较为简单，检测类别较为单一，因此本文采用网络爬虫以及在合肥市街道实景拍摄的方式获取实验图片，使用 Labelimg 脚本工具对图片进行人工标注，制成 Helmet 数据集。图 9 为 Labelimg 脚本的标注界面，从标注脚本界面可以看出，由于远景目标和近景目标的尺寸存在明显差异，所以目标标签的尺度也十分丰富。图 10 展示了该图例不同类别尺寸分布的散点图，从中可以定量看出检测目标尺寸大小的差异性。Helmet 数据集类别包括：骑行人员 (driver)、佩带头盔 (helmet) 和未佩带头盔 (no helmet)。数据集共 6500 张，按照训练集、验证集和测试集 8 : 1 : 1 的比例进行实验。处理后的类别数量如表 1 所示。



双轮车佩戴头盔检测数据集 (Two wheeler helmet dataset, TWHD) 为开源数据集, 用来进一步验证本文提出方法的泛化能力。TWHD 数据集共 5448 张, 类别包括: 双轮车和驾驶员整体 (two\_wheeler)、佩戴头盔 (helmet) 和未佩戴头盔 (without\_helmet)。具体类别数量如表 2 所示。

### 4.2 实验环境

本次实验是在 Windows 10 操作系统下进行的, CPU 型号为 i7-12700F, GPU 型号为 NVIDIA GeForce RTX 3090 (24 G), 训练的框架为 pytorch 2.0.0、CUDA11.8、Python 3.9。具体超参数配置如表 3 所示。



图 9 Labeling 脚本界面  
Fig. 9 Labeling script interface

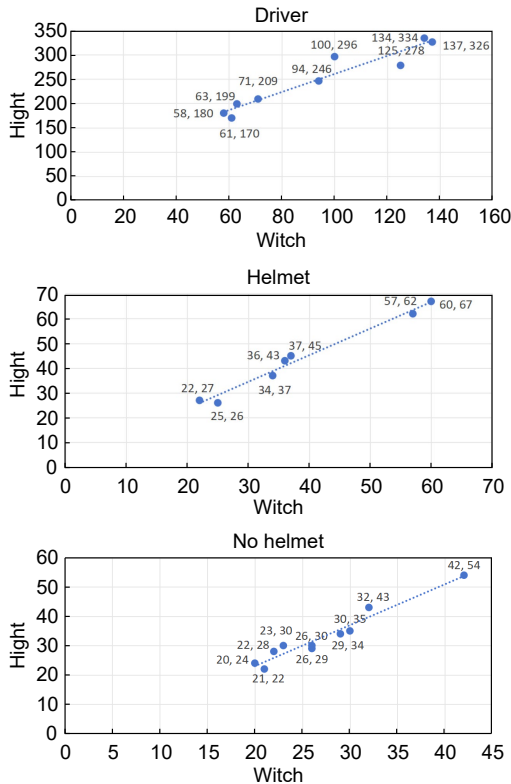


图 10 不同类别长宽分布散点图

Fig. 10 Scatterplots of length and width distribution of different categories

表 1 Helmet 数据集处理后各目标数量

Table 1 Number of target after Helmet dataset processing

Class	Driver	Helmet	No helmet
Target number	9027	7743	3184

表 2 TWHD 数据集处理后各目标数量

Table 2 Number of target after TWHD dataset processing

Class	Two_wheeler	Helmet	Without_helmet
Target number	13790	11742	6895

表 3 实验参数配置

Table 3 Configuration of experimental parameters

Key parameter	Parameter value
Epoch	200
lr0	0.01
lmgz	640
Batch	8
Momentum	0.937

### 4.3 评价指标

本文实验采用平均精度 (mean average precision, mAP)、召回率 (recall,  $R$ ) 以及模型的参数量 (parameters) 作为评价指标。

mAP用来评估模型在检测任务中的整体性能, 综合了精确度 (precision,  $P$ ) 和召回率两个指标, 精确度和召回率的计算公式如下:

$$P = \frac{TP}{TP + FP}, \quad (7)$$

$$R = \frac{TP}{TP + FN}, \quad (8)$$

其中:  $TP$ 为正确检验到的正样本,  $FP$ 为错误检测为正样本的负样本,  $FN$ 表示没有检测到的负样本。 $AP$  (average precision) 是通过在不同的  $R$  阈值下计算出  $P$ , 通过绘制曲线  $P-R$ , 计算  $P-R$  曲线所围成的面积, 并运用积分求得  $AP$  值。 $mAP$  为对所有类别的  $AP$  求和取平均值。计算公式如下:

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

$$mAP = \frac{1}{N} \sum_{i=1}^N AP_i, \quad (10)$$

其中:  $P(R)$  为 0 到 1 阈值内的  $P-R$  函数曲线, 代表检测的类别个数,  $AP_i$  是第  $i$  个类别的精度。

根据 IoU 的阈值的不同, 将  $mAP$  指标分为

$mAP50$  和  $mAP50 : 95$ ,  $mAP50$  为 IoU 阈值为 0.5 时的平均检测精度,  $mAP50 : 95$  则是 IoU 阈值在 0.5~0.95, 并以 0.05 为间隔步长计算出来的平均检测精度。

### 4.4 损失函数分析

#### 4.4.1 阈值设定

阈值  $d$  和  $u$  根据具体任务来确定, 为了验证阈值对实验结果的影响, 从而确定较优的阈值选取, 本文选取了 25 组阈值  $[d, u]$  进行实验对比, 如表 4 和表 5 所示, 实验结果表明, 在当前头盔检测任务中, 当阈值  $[d, u]$  设定为  $[0.3, 0.8]$  时的  $mAP50$  和  $mAP50 : 95$  的效果最佳。后续对比实验的损失函数均使用 Focaler-CIoU。

#### 4.4.2 损失函数对比

为了进一步验证损失函数的优势和泛化能力, 在 YOLOv8、YOLOv5 均使用 Focaler-CIoU 进行对比, 如表 6 所示。实验结果表明, 相较于原模型, 使用 Focaler-CIoU 损失后  $mAP$  均得到提高, 验证了本文损失的优势和泛化性。

图 11 是在 YOLOv8 原模型的基础上, 对比改进损失前后损失变化曲线, 改进后模型的边界框损失值明显下降, 对骑行人员头盔检测任务的数据拟合效果有明显提升。

表 4 不同阈值下  $mAP50$  值  
Table 4  $mAP50$  values at different thresholds

$d$	$u$				
	1	0.9	0.8	0.7	0.6
0	0.842	0.840	0.840	0.841	0.842
0.1	0.842	0.842	0.838	0.840	0.842
0.2	0.840	0.840	0.839	0.842	0.840
0.3	0.841	0.842	<b>0.843</b>	0.842	0.838
0.4	0.839	0.841	0.842	0.838	0.840

表 5 不同阈值下  $mAP50 : 95$  值  
Table 5  $mAP50 : 95$  values at different thresholds

$d$	$u$				
	1	0.9	0.8	0.7	0.6
0	0.681	0.682	0.684	0.681	0.681
0.1	0.684	0.685	0.682	0.682	0.682
0.2	0.682	0.684	0.686	0.686	0.679
0.3	0.685	0.685	<b>0.690</b>	0.686	0.678
0.4	0.683	0.682	0.688	0.680	0.679

表 6 改进前后 mAP 对比

Table 6 Comparison of mAP before and after improvement

Algorithm	Pre-improvement loss		Post-improvement loss	
	mAP50	mAP50 : 95	mAP50	mAP50 : 95
YOLOv5n	0.830	0.652	<b>0.832 (+0.2%)</b>	<b>0.656 (+0.4%)</b>
YOLOv8n	0.839	0.686	<b>0.843 (+0.4%)</b>	<b>0.690 (+0.4%)</b>

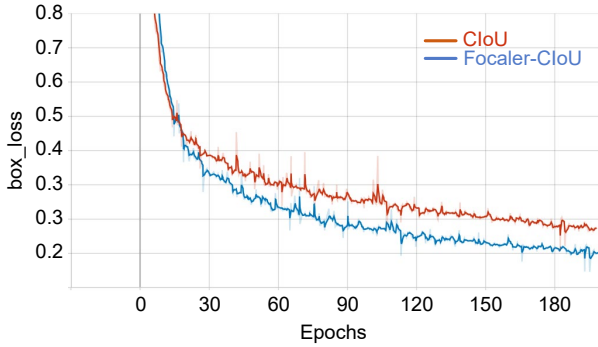


图 11 改进前后 box\_loss 对比

Fig. 11 Comparison of box\_loss before and after improvement

#### 4.5 C2f\_BC 模块分析

为了验证 GSConv 在 C2f 中的有效性, 在原模型的基础上将主干网络中的 C2f 中的普通卷积 (Conv) 替换成 GSConv, 并将其与原模型进行对比, 如表 7 所示。数据表明, 用 GSConv 卷积替换 Conv 后, mAP50 和 mAP50 : 95 分别提升了 0.6% 和 0.3%, 且参数量降低了 7.0%。

为了验证 CBAM 注意力的有效性, 在 C2f 模块引入 GSConv 的基础上进行对比实验。将 CBAM 分别设计在 Bottleneck 的第一个 GSConv 后和第二个 GSConv 后进行对比。如表 8 所示, C2f\_B1 和 C2f\_B2 分别表示在第一个和第二个 GSConv 后添加 CBAM。实验表明将 CBAM 添加在第一个 GSconv 后效果不佳,

添加在第二个 GSConv 后, 虽然参数量略微提升, 但是 mAP50 和 mAP50 : 95 分别又提升了 0.4% 和 0.3%。

为了验证 C2f\_BC 中引入 CBAM 的优势, 以引入 GSConv 的 C2f 模块为基础, 在特征提取阶段的相同位置分别加入不同注意力模块, 包括 ECA<sup>[18]</sup> (efficient channel attention)、SE<sup>[19]</sup> (squeeze-and excitation)、GAM<sup>[20]</sup> (global attention mechanism) 和 EMA<sup>[21]</sup> (expectation-maximization attention)。由表 9 所示。经对比, 引入 ECA 和 SE 注意力时, 评价指标 mAP50 : 95 分别上升了 0.1% 和 0.2%, 提升效果不明显。引入 GAM 注意力后, mAP50 和 mAP50 : 95 均提升了 0.4%, 但模块的参数量相较于第一组实验增加了 18.4%。经对比发现, 引入 CBAM 之后虽然参数量增加了 3.1%, 但 mAP50 和 mAP50 : 95 分别提升了 0.4% 和 0.3%。相对于其他注意力模块, 在第一组实验的基础上引入 CBAM 在检测精度和参数量之间实现了良好的平衡。

#### 4.6 MP-Parnet 分析

为了验证 MP-Parnet 模块在原模型的优势以及此模块的泛化能力, 分别在 YOLOv5n、YOLOv6n、YOLOv8n、YOLOv10n<sup>[22]</sup> 和 RT-DETR (r18)<sup>[23]</sup> 的颈部网络特征的提取阶段分别加入 MP-Parnet 进行实验对

表 7 GSConv 和普通卷积对比

Table 7 Comparison of GSConv and Conv

Convolutional type	mAP50	mAP50 : 95	Params/10 <sup>6</sup>
Conv	0.843	0.690	3.15
<b>GSConv</b>	<b>0.849 (+0.6%)</b>	<b>0.693 (+0.3%)</b>	<b>2.93 (-7.0%)</b>

表 8 在不同位置引入 CBAM

Table 8 Introducing CBAM at different convolutional layers

Module	mAP50	mAP50 : 95	Params/10 <sup>6</sup>
C2f+GSConv	0.849	0.693	2.93
C2f_B1	0.849	0.691	3.02
<b>C2f_B2</b>	<b>0.853 (+0.4%)</b>	<b>0.696 (+0.3%)</b>	3.02 (+3.1%)

比, 如表 10 所示。实验结果表明, 将 MP-Parnet 作用在以上网络中, mAP50 和 mAP50 : 95 均得到了提升, 证明 MP-Parnet 模块在本文模型的优势, 同时证明此模块具有较强的泛化能力。

#### 4.7 消融实验对比

##### 4.7.1 消融实验

为了验证 C2f\_BC 模块和 MP-Parnet 对模型改进的有效性, 以 YOLOv8n 为基准模型分别在 Helmet 和 TWHD 数据集上进行消融实验。如表 11 所示, 第 1 组实验为 YOLOv8n 基准模型; 第 2 组实验是在 YOLOv8n 的基础上引入损失函数 Focaler-CIoU, 较于基准模型, mAP50 和 mAP50 : 95 在 Helmet 数据集上分别提升了 0.4% 和 0.5%, 在 TWHD 数据集上分别提升了 0.5% 和 0.3%, 说明损失函数 Focaler-CIoU 可以通过平衡正负样本改善精度值; 第 3 组实

验是在第 2 组实验的基础上将主干网络 (Backbone) 的 C2f 替换为 C2f\_BC, 较于第 2 组实验, mAP50 和 mAP50 : 95 在 Helmet 数据集上分别提升了 1.0% 和 0.5%, 在 TWHD 数据集上分别提升了 1.1% 和 1.4%, 且参数量降低了 4.1%, 计算量降低了 5.7%, 说明 C2f\_BC 模块可以降低计算冗余的同时增强模型的特征提取能力; 第 4 组实验是在第 2 组实验的基础上在颈部网络 (Neck) 引入 MP-Parnet 模块, 较于第 2 组实验, 虽然在计算量和参数量方面分别上升了 5.0% 和 3.4%, 但是 mAP50 和 mAP50 : 95 在 Helmet 数据集上分别提升了 1.3% 和 0.7%, 且在 TWHD 数据集上分别提升了 0.9% 和 1.2%, 说明多核并行感知网络 (MP-Parnet) 在处理目标尺度丰富的场景中展现出了优异的性能; 第 5 组实验是将损失函数 Focaler-CIoU、C2f\_BC 模块和 MP-Parnet 模块全部加入基准模型得

表 9 不同注意力之间精度对比

Table 9 Comparison of mAP among different attentions

No.	Module	Params/10 <sup>6</sup>	mAP50	mAP50 : 95
1	C2f+GSCConv	2.93	0.849	0.693
2	1+ECA	2.93	0.848	0.694
3	1+SE	2.98	0.849	0.695
4	1+GAM	3.47	0.853	0.697
5	1+EMA	2.94	0.852	0.692
6	1+CBAM	3.02	0.853	0.696

表 10 在不同网络颈部加入 MP-Parnet

Table 10 Adding MP-Parnet to different network necks

Module	Pre-Improvement		Post-Improvement	
	mAP50	mAP50 : 95	mAP50	mAP50 : 95
YOLOv5n	0.832	0.656	<b>0.840 (+0.8%)</b>	<b>0.674 (+1.8%)</b>
YOLOv6n	0.822	0.663	<b>0.841 (+1.9%)</b>	<b>0.671 (+0.8%)</b>
YOLOv8n	0.843	0.690	<b>0.856 (+1.3%)</b>	<b>0.698 (+0.8%)</b>
YOLOv10n	0.837	0.685	<b>0.844 (+0.7%)</b>	<b>0.690 (+0.5%)</b>
RT-DETR (r18)	0.802	0.651	<b>0.829 (+2.7%)</b>	<b>0.674 (+2.3%)</b>

表 11 消融实验

Table 11 Comparison of ablation experiments

No.	YOLOv8n	Focaler-CIoU	C2f_BC	MP-Parnet	Helmet		TWHD		Params/10 <sup>6</sup>	GFLOPs
					mAP50	mAP50 : 95	mAP50	mAP50 : 95		
1	√				0.839	0.686	0.725	0.437	3.15	8.7
2	√	√			0.843	0.691	0.730	0.440	3.15	8.7
3	√	√	√		0.853	0.696	0.741	0.454	3.02	8.2
4	√	√		√	0.856	0.698	0.739	0.452	3.31	9.0
5	√	√	√	√	<b>0.861</b>	<b>0.705</b>	<b>0.743</b>	<b>0.456</b>	3.12	8.4

到最终模型 THI-YOLO, 较于基准模型, mAP50 和 mAP50 : 95 在 Helmet 数据集上分别提升了 2.2% 和 1.9%, 在 TWH D 数据集上分别提升了 1.8% 和 1.9%, 且参数量降低了 0.9%, 计算量降低了 3.4%, 说明 THI-YOLO 算法在非机动车驾驶员的安全头盔检查任务中可以实现更高的精度。

#### 4.7.2 对比实验

为了进一步验证本文模型的优越性, 将改进后的模型在 Helmet 和 TWH D 数据集上分别与 YOLOv5n、YOLOv5s、YOLOv6n、YOLOv8s、YOLOv8n-ghost<sup>[24]</sup>、YOLOv10n 和 RT-DETR (r18) 等当前的主流算法进行实验对比。如表 12 所示。实验结果表明, 本文的方法相较于以上几种方法, 在保持较小参数量和计算量的同时取得了较好的 mAP。

#### 4.8 实验检测效果展示

为了更加直观验证改进模型的优势, 采用 Helmet 和 TWH D 的测试集对比 THI-YOLO 模型和 YOLOv8n 模型的检测效果, 如图 12 所示。通过对比, 在非机动车驾驶员的头盔检测场景中, 原始的 YOLOv8n 模型无法定位到远景的小目标, 且在复杂背景的干扰下出现了误检现象, 如图 12 (a, b) 所示, 说明原模型缺乏对尺度小的目标的特征提取能力, 且在复杂场景中对上下文信息的识别能力较差, THI-YOLO 模型展现出对局部细节特征捕获的优势。图 12 (c) 展现出在不同尺度的密集目标中, 相较于原模型, THI-YOLO 模型整体表现出更高的检测精度。实验结果表明, 本文模型可以有效地提高目标的检测精度, 且在一定程度上避免了误检和漏检现象的发生, 体现出本文模型的优越性。

## 5 结束语

综上所述, 本文提供了一种应用于非机动车驾驶员的头盔检测的算法, 在 YOLOv8n 的基础上设计了 C2f\_BC 模块, 有效地降低了计算的复杂度, 增强对特征的提取能力, 采用并行处理不同卷积核大小的 DWConv 对 Parnet 进行重新设计, 使得检测器更有效地捕获不同尺度的目标。实验结果表明, 改进后的模型在保证参数量和计算量下降的同时, mAP50 和 mAP50 : 95 在自建 Helmet 数据集分别提升了 2.2% 和 1.9%, 在开源数据集 TWH D 上分别提升了 1.8% 和 1.9%, 有效地提升了模型对复杂场景的特征提取能力以及对不同尺度信息的捕获能力。

在未来阶段, 面向非机动车驾驶员的头盔检测还有很大的提升空间。虽然改进模型相较于基准模型在精度上有所提升, 但是参数量的优化不明显, 未来工作主要考虑用剪枝或者蒸馏技术对模型进行进一步的轻量化。

**利益冲突:** 所有作者声明无利益冲突

## 参考文献

- [1] Li B, Hou D Z, Zhang J S, et al. Study on conception and mechanism of intelligent vehicle-infrastructure cooperation[J]. *J Highw Transp Res Dev*, 2020, 37 (10): 134-141. 李斌, 侯德藻, 张纪升, 等. 论智能车路协同的概念与机理[J]. *公路交通科技*, 2020, 37 (10): 134-141.
- [2] Cen Y Q, Song X H, Wang D Z, et al. Establishment of technology system of smart expressway[J]. *J Highw Transp Res Dev*, 2020, 37 (7): 111-121. 岑晏青, 宋向辉, 王东柱, 等. 智慧高速公路技术体系构建[J]. *公路交通科技*, 2020, 37 (7): 111-121.
- [3] Ren S, He K, Girshick R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[C]//

表 12 与其他模型对比

Table 12 Comparison with other models

Module	Helmet		TWH D		Params/10 <sup>6</sup>	GFLOPs
	mAP50	mAP50: 95	mAP50	mAP50: 95		
YOLOv5n	0.832	0.656	0.719	0.430	2.5	7.1
YOLOv5s	0.848	0.691	0.724	0.446	9.11	23.8
YOLOv6n	0.822	0.663	0.713	0.432	4.23	11.8
YOLOv8s	0.859	0.700	0.738	0.453	11.12	28.4
YOLOv8n-ghost	0.834	0.655	0.720	0.429	1.71	5.0
YOLOv10n	0.837	0.685	0.729	0.441	2.69	8.2
RT-DETR (r18)	0.802	0.651	0.696	0.417	19.9	57.0
<b>Ours</b>	<b>0.861</b>	<b>0.705</b>	<b>0.743</b>	<b>0.456</b>	<b>3.12</b>	<b>8.4</b>



图 12 定性检测结果展示。(a) 漏检情况; (b) 误检情况; (c) 骑行人员密集场景检测效果对比

Fig. 12 Display of qualitative experiment results. (a) Cases of object leakage; (b) Cases of object misdetections; (c) Comparison of detection effects for dense scenes of drivers

- Proceedings of the 28th International Conference on Neural Information Processing Systems*, 2015, 1: 91–99.
- [4] Liu W, Anguelov D, Erhan D, et al. SSD: single shot MultiBox detector[C]//*Proceedings of the 14th European Conference on Computer Vision (ECCV)*, 2016: 21–37. [https://doi.org/10.1007/978-3-319-46448-0\\_2](https://doi.org/10.1007/978-3-319-46448-0_2).
- [5] Carion N, Massa F, Synnaeve G, et al. End-to-end object detection with transformers[C]//*Proceedings of the 16th European Conference on Computer Vision*, 2020: 213–229. [https://doi.org/10.1007/978-3-030-58452-8\\_13](https://doi.org/10.1007/978-3-030-58452-8_13).
- [6] Liu C, Gao J. An improved YOLOv5 electric vehicle helmet wearing detection method[J]. *Software Guide*, 2024, 23 (6): 143–149.  
刘超, 高健. 一种改进的 YOLOv5 电动车头盔佩戴检测方法[J]. *软件导刊*, 2024, 23 (6): 143–149.
- [7] Wang R, Liu D M, Zhang J. Wear-YOLO: research on detection methods of safety equipment for power personnel in substations[J]. *Comput Eng Appl*, 2024, 60 (9): 111–121.  
王茹, 刘大明, 张健. Wear-YOLO: 变电站电力人员安全装备检测方法研究[J]. *计算机工程与应用*, 2024, 60 (9): 111–121.
- [8] Wang P F, Huang H M, Wang M Q. Complex road target detection algorithm based on improved YOLOv5[J]. *Comput Eng Appl*, 2022, 58 (17): 81–92.  
王鹏飞, 黄汉明, 王梦琪. 改进 YOLOv5 的复杂道路目标检测算法[J]. *计算机工程与应用*, 2022, 58 (17): 81–92.
- [9] Zhou S Y, Peng Z Y, Zhang H L, et al. Helmet-YOLO: a higher-accuracy road safety helmet detection algorithm[J]. *Comput Eng Appl*, 2024, .  
周顺勇, 彭梓洋, 张航领, 等. Helmet-YOLO: 一种更高精度的道路安全头盔检测算法[J]. *计算机工程与应用*, 2024, .
- [10] Li H L, Li J, Wei H B, et al. Slim-neck by GSConv: a better design paradigm of detector architectures for autonomous vehicles[Z]. arXiv: 2206.02424, 2022. <https://doi.org/10.48550/arXiv.2206.02424>.
- [11] Woo S, Park J, Lee J Y, et al. CBAM: convolutional block attention module[C]//*Proceedings of the 15th European Conference on Computer Vision (ECCV)*, 2018: 3–19. [https://doi.org/10.1007/978-3-030-01234-2\\_1](https://doi.org/10.1007/978-3-030-01234-2_1).
- [12] Goyal A, Bochkovskiy A, Deng J, et al. Non-deep networks[C]//*Proceedings of the 36th International Conference on Neural Information Processing Systems*, 2022: 492.
- [13] Chollet F. Xception: deep learning with depthwise separable convolutions[C]//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017: 1251–1258. <https://doi.org/10.1109/CVPR.2017.195>.
- [14] Zhang H, Zhang S J. Focaler-IoU: more focused intersection over union loss[Z]. arXiv: 2401.10525, 2024. <https://doi.org/10.48550/arXiv.2401.10525>.
- [15] Chen X, Peng D L, Gu Y. Real-time object detection for UAV images based on improved YOLOv5s[J]. *Opto-Electron Eng*, 2022, 49 (3): 210372.  
陈旭, 彭冬亮, 谷雨. 基于改进 YOLOv5s 的无人机图像实时目标检测[J]. *光电工程*, 2022, 49 (3): 210372.
- [16] Han K, Wang Y H, Tian Q, et al. GhostNet: more features from cheap operations[C]//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020: 1577–1586. <https://doi.org/10.1109/CVPR42600.2020.00165>.
- [17] Zhang X Y, Zhou X Y, Lin M X, et al. ShuffleNet: an extremely efficient convolutional neural network for mobile devices[C]//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2018: 6848–6856. <https://doi.org/10.1109/CVPR.2018.00716>.
- [18] Zhang H, Xu C, Zhang S J. Inner-IoU: more effective intersection over union loss with auxiliary bounding box[Z]. arXiv: 2311.02877, 2023. <https://doi.org/10.48550/arXiv.2311.02877>.
- [19] Hu J, Shen L, Sun G. Squeeze-and-excitation networks[C]//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2018: 7132–7141. <https://doi.org/10.1109/CVPR.2018.00745>.
- [20] Liu Y C, Shao Z R, Hoffmann N. Global attention mechanism: retain information to enhance channel-spatial interactions[Z]. arXiv: 2112.05561, 2021. <https://doi.org/10.48550/arXiv.2112.05561>.
- [21] Ouyang D L, He S, Zhang G Z, et al. Efficient multi-scale attention module with cross-spatial learning[C]//*ICASSP 2023-IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2023: 1–5. <https://doi.org/10.1109/ICASSP49357.2023.10096516>.
- [22] Wang A, Chen H, Liu L H, et al. YOLOv10: real-time end-to-end object detection[Z]. arXiv: 2405.14458, 2024. <https://doi.org/10.48550/arXiv.2405.14458>.
- [23] Lv W Y, Xu S L, Zhao Y A, et al. DETRs beat YOLOs on real-time object detection[Z]. arXiv: 2304.08069, 2023. <https://doi.org/10.48550/arXiv.2304.08069>.
- [24] Xiong E J, Zhang R F, Liu Y H, et al. Ghost-YOLOv8 detection algorithm for traffic signs[J]. *Comput Eng Appl*, 2023, 59 (20): 200–207.  
熊恩杰, 张荣芬, 刘宇红, 等. 面向交通标志的 Ghost-YOLOv8 检测算法[J]. *计算机工程与应用*, 2023, 59 (20): 200–207.

## 作者简介



【通信作者】孙光灵 (1976-), 男, 硕士, 副教授, 主要研究方向为机器学习、图像处理。

E-mail: [sunguangling@163.com](mailto:sunguangling@163.com)



王薪博 (2001-), 男, 硕士研究生, 主要研究方向为图像处理、目标检测。

E-mail: [tongzi83715@163.com](mailto:tongzi83715@163.com)

# THI-YOLO: Improved non-motorized drivers helmet detection of YOLOv8

Sun Guangling<sup>1,2\*</sup>, Wang Xinbo<sup>1</sup>

**Overview:** With the gradual emergence of computer vision in recent years, target detection algorithms are also constantly innovating and developing, and the detection of safety helmets on two-wheeled vehicles is also one of the focuses of research scholars. Two-wheeled vehicles have become the main means of transportation for citizens to travel at this stage, but the phenomena of citizens running red lights and not wearing safety helmets are still common, so it is especially urgent to design a helmet-wearing detection method for cyclists. At the present stage of safety helmet detection, there are still some difficulties, such as complex background information, the detection target exists in different scales of the diversity of changes, so the design of higher performance helmet detection algorithms needs to carry out further research. Aiming at the phenomenon that complex background often exists in non-motorized drivers' helmet detection and the diversity of detection target scales often exists in the detection scene, which in turn leads to low detection efficiency and misdetection and omission, a YOLOv8 algorithm oriented to traffic helmet detection is proposed. The alteration points of this paper mainly contain the following two blocks. The first is the C2f\_BC module, the GSConv module with the idea of combining group convolution and spatial convolution is introduced, and the attention mechanism combining channel and space (convolutional block attention module, CBAM) is introduced in Bottleneck in C2f. To effectively reduce the computational complexity and enhance the extraction of local and global features, we designed the parallel multiscale feature fusion module MP-Parnet (parallel multiscale perception networks) and redesigned the Parnet (parallel networks) by using the parallel depth-separable feature fusion module with different scales. The second is a parallel depthwise separable convolution (DWConv) kernel for different scales. It is used instead of the ordinary convolution of the original module, which effectively adapts to the acquisition ability of different scales of targets. 3. Focaler-IoU is introduced into the original model, and Focaler-CIoU is designed, which effectively enhances the performance of the detection model in both classification and detection. detection performance. The experimental results show that compared with the original model, the improved YOLOv8n maintains a decrease in the number of parameters while the mAP50 and mAP50: 95 increase by 2.2% and 1.9% on the self-built dataset Helmet, and 1.8% and 1.9% on the open-source dataset TWHHD, which suggests that the improved model can be better applied to the helmet detection of non-motorized drivers in the scenario.

Sun G L, Wang X B. THI-YOLO: Improved non-motorized drivers helmet detection of YOLOv8[J]. *Opto-Electron Eng*, 2024, 51(12): 240244; DOI: [10.12086/oe.2024.240244](https://doi.org/10.12086/oe.2024.240244)



扫描二维码, 获取PDF全文