DOI:10.19652/j.cnki.femt.2305586

改进YOLOv5s算法的电动车头盔检测研究*

侯恩翔[1] 张 旭[1] 刘 罡[2] 张秀再[1]
(1.南京信息工程大学电子信息工程学院 南京 210044;
2.无锡学院电子信息工程学院 无锡 214105)

摘 要:针对电动车头盔佩戴检测存在遮挡、车辆密集以及一车多人的复杂场景下出现的漏检、误检问题,在YOLOv5s的基础上,提出了一种应用于电动车头盔佩戴检测的改进算法。设计了一种由递归门控卷积改进的GBC3模块,替换网络主干和特征融合层(feature pyramid networks, FPN)中的C3模块,加强邻间特征的空间信息交互,提高网络的特征提取和特征融合能力;其次在主干和特征融合网络添加无参注意力机制(SimAM),以调整特征图中不同区域的注意力权重,对重要目标施加更多关注;最后引入调整超参后的WIOU损失函数,优化预测框回归,提高模型的目标定位能力。在自制电动车头盔数据集上的实验结果表明,改进模型在仅增加较少参数的前提下,其平均精度均值(mAP)达到97.3%,较YOLOv5s提高了3.2%,并且检测速度为87.1 fps,改善了误检和漏检的问题,同时仍具有较高的实时性,更适用于电动车驾乘者的头盔佩戴检测。

关键词:电动车头盔检测;递归门控卷积;空间交互;注意力机制
中图分类号:TP391.4 **文献标识码**:A **国家标准学科分类代码**:520.20

Improved YOLOv5s algorithm for electric bike helmet wearing detection

Hou Enxiang[1] Zhang Xu[1] Liu Gang[2] Zhang Xiuzai[1]
(1. School of Electronics and Information Engineering, Nanjing University of Information Science and Technology, Nanjing 210044, China;
2. School of Electronics and Information Engineering, Wuxi University, Wuxi 214105, China)

Abstract: Aiming at the leakage and misdetection problems of electric vehicle helmet wearing detection in the presence of occlusion, dense vehicles and the complex scene of multiple people in one vehicle, an improved algorithm applied to electric vehicle helmet wearing detection is proposed on the basis of YOLOv5s. A GBC3 module improved by recursive gating convolution is designed to replace the C3 module in the network backbone and feature fusion layer (feature pyramid networks, FPN), strengthen the spatial information interaction of neighbor features, and improve the feature extraction and feature fusion capabilities of the network. Secondly, non-parametric attention mechanism (a simple, parameter-free attention module for convolutional neural networks, SimAM) is added to the backbone and feature fusion network to adjust the attention weight of different regions in the feature map and pay more attention to important targets. Finally, the WIOU loss function is introduced to optimize the prediction box regression and improve the target localization ability of the model. The experimental results on the self-made electric vehicle helmet dataset show that the mAP of the improved model reaches 97.3% under the premise of only adding fewer parameters, which is 3.2% points higher than that of YOLOv5s, and the detection speed is 87.1 fps, which improves the problem of false detection and missed detection, and still has high real-time performance, which is more suitable for helmet wearing detection of electric vehicle drivers.

Keywords: electric scooter helmet detection; recursive gated convolution; spatial interaction; attention mechanism

收稿日期:2023-09-23
* **基金项目**:国家自然科学基金(62204172)项目资助

0 引 言

目前，电动车出行的方式因其节能环保、体积小巧、速度轻快、成本较低的特点，成为广大市民短途出行选择较多的交通工具。但市民的一些不安全驾驶习惯，引发了许多交通事故，其中的大部分事故是因为电动车驾乘人员未将头盔佩戴。据相关研究发现，未戴头盔时头部受伤的概率是佩戴头盔时的2.5倍，而致命伤的发生率则是佩戴头盔时的1.5倍。在2023年7月1日全国实行的电动车新规中，政府也明确规定电动车驾乘者骑行时必须规范佩戴符合国家标准的头盔才可上路。所以对头盔佩戴进行检测、减少相关安全事故的发生尤为必要。

有研究人员使用基于传感器识别和基于图像处理的目标检测方法[1]对安全帽的佩戴检测进行研究，但这些传统目标检测算法的效率和精度较低，不适用于实时性要求高的复杂场景中。随着近年来深度学习的发展，许多学者已将基于深度学习的目标检测算法应用到工业领域中对安全帽的佩戴检测[2-3]，检测速度和准确度大幅提高，对工人生命安全保障起了一定的促进作用。类似地，为了更好地保护电动车驾乘者的安全和提高交警执法效率，研究基于深度学习的目标检测算法对驾乘者的头盔佩戴情况进行监测，具有很大的现实意义。

传统的目标检测方法是通过滑窗操作进行区域选择，手动设计特征提取器，例如SIFT[4]、Harr等，然后用支持向量机进行分类来确定目标在图像中是否存在。而基于深度学习的目标检测算法是根据边界回归框生成的区别，分为基于区域选择的两阶段算法与基于回归的一阶段算法。两阶段算法主要有R-CNN系列[5-7]、R-FCN[8]等算法，一阶段算法主要有YOLO系列[9-12]、SSD[13]、RetinaNet[14]等算法。两阶段算法通常拥有较高的检测精度，但生成候选框时引入了额外的运算量，检测实时性较低。然而一阶段算法通过直接在输入图像上进行目标检测，其检测速度较快。由于头盔佩戴检测所需的实时性要求较高，所以近年来大多数学者选择一阶段算法进行改进，使其更适用于头盔的佩戴检测。朱硕等[15]对YOLOv5模型进行改进，引入卷积注意力模块[16](convolutional block attention module，CBAM)和(simple online and realtime tracking，SORT)目标追踪算法，提高了对电动车头盔佩戴的检测速度和检测准确性，能够同时实现追踪和识别。但对密集场景下目标的检测能力较差，准确率较低。刘雪纯等[17]针对密集目标检测的检测准确率不高的问题，提出了一种YOLOv5-Q的改进算法，在原网络80×80的特征图上进行2倍上采样操作，并融合原模型的3层特征信息，改进成四尺度检测，提高了密集目标的检测精度，但总的平均精度有所下降，并且在目标流量较大的场景下，易出现漏检和误检的情况。张瑞芳等[18]提出一种改进YOLOv5s的头盔检测算法，使用软池化[19]替换空间金字塔池化(spatial pyramid pooling，SPP)结构中的最大池化，并将加权双向金字塔[20](bidirectional feature pyramid network，Bi-FPN)网络与坐标注意力[21](coordinate attention，CA)进行结合，以增强不同层级的特征传递。张子涵等[22]使用Kmeans++算法聚类生成锚框，采用上采样算子(content-aware reassembly of feature，CARAFE)[23]，同时引入坐标注意力机制CA，最后利用非极大值抑制算法DIOU-NMS[24]进行模型的输出后处理，增大了感受野，充分利用了特征语义信息，改善了对密集目标的检测能力，同时保证了其轻量化，提高了算法的检测精度。但上述方法均未考虑到实际场景中的小目标检测。针对这一问题，朱周华等[25]通过引入卷积注意力模块CBAM和CA注意力模块，采用改进的DIOU-NMS，并增加多尺度特征融合检测，提高了头盔小目标的检测效果。谢博轩等[26]通过添加高效通道注意力机制(Efficient Channel Attention Network，ECA[27])，以及Bi-FPN模块，并使用Alpha-CIOU Loss[28]，以平衡不同层级的特征重要性和增强定位性能，提升了小目标的检测效果。吕志轩等[29]曾提出一种改进基于YOLOX网络的检测方法，在模型预测端添加分支注意力模块，又提出一种新的余弦退火算法(CDWR)，减少了模型收敛时间，改善了对小目标的检测能力。但对存在遮挡场景下的目标检测却依然较弱，有一定的漏检现象。张鑫等[30]基于YOLOv5s进行改进，增加了浅层检测尺度以及上采样特征融合结构，并引入CBAM注意力模块，有效的提高了在遮挡情况下的目标物检测。纪超等[31]基于YOLOv7算法进行改进，提出多分支深度卷积，增加深度可分离卷积层来增强特征提取的完备性；然后加入多通道交互注意力(multimodal interaction attention，MIA)，并结合ECA注意力机制，提高了模型对遮挡目标的识别精度。上述改进算法虽然具有一定的贡献，但检测性能仍然不佳，存在较多的漏检、误检情况，难以满足在车辆种类繁多、车辆密集且有遮挡和一车多人等复杂道路场景下的实时性检测。

针对以上问题，本文基于YOLOv5s模型进行改进，提出了一种实时高效且能改善上述复杂道路场景下头盔漏检、误检问题的电动车头盔佩戴检测算法。

1 YOLOv5s模型

YOLOv5算法是一种非常高效的一阶段目标检测算法，检测速度快并且精度高。而YOLOv5s是YOLOv5系列最小的模型，考虑到模型大小、速度、参数量以及便于部署在嵌入式设备上，本文选用YOLOv5s作为检测电动车驾乘者是否佩戴头盔的基础模型，以便更好地满足道路上的检测实时性要求。

YOLOv5s模型的框架由输入端(input)、主干网络(backbone)、颈部(neck)网络和检测头(head)4部分构成。输入端部分在训练阶段将图像进行宽高大小为640×640的自适应缩放，然后使用Mosaic数据增强方法，随机抽取

4 张图片进行组合拼接，丰富了检测目标的背景和数据集。主干网络的作用是用来提取输入图像的特征，沿用了类似于 YOLOv3 的 Darknet 网络结构同时采用 CSP 结构提高网络性能，设计成高效的 CSPDarknet53 结构，主要由 CBS、C3、SPPF 模块组成。颈部网络由 FPN 和 PAN 结构组成，实现了多尺度特征融合[32]。检测头部使用 3 个不同尺度大小的分类器来预测大、中、小规格的类别目标。YOLOv5s 模型结构如图 1 所示。

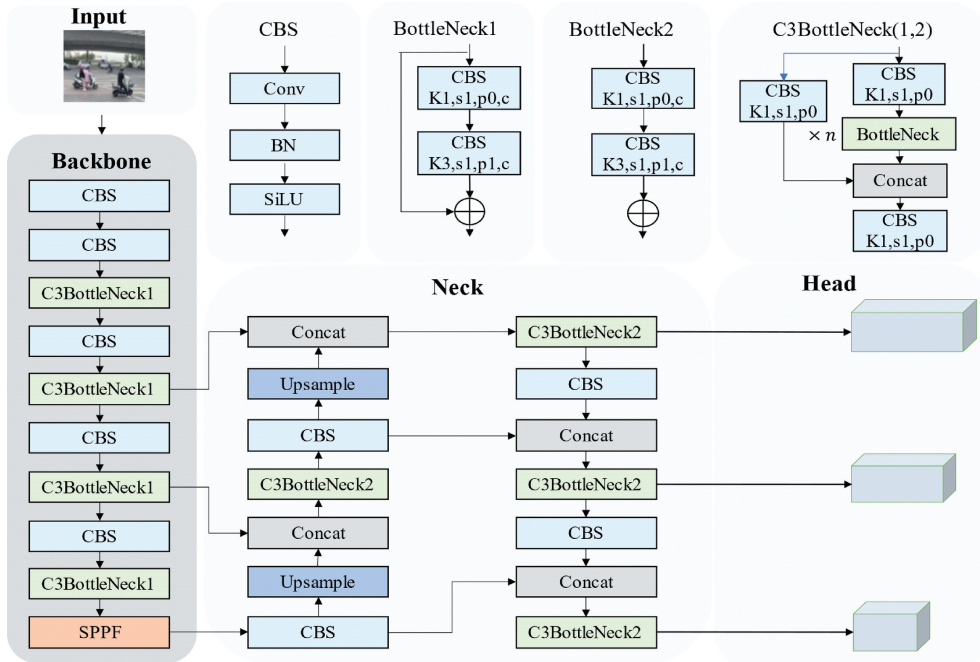

图 1　YOLOv5s 算法模型结构

2　改进 YOLOv5s 算法

本文基于 YOLOv5s 算法，提出了一种电动车驾乘者头盔佩戴检测方法。首先在 YOLOv5s 对主干网络和颈部 FPN 特征融合层的 C3 模块改进为 GBC3 模块增强特征间的空间交互能力，进一步提升网络对上下文语义信息的理解能力，GBC3 模块由 3 个递归门控卷积（g^nConv）和 GBlock 构成。

主干网络的浅层特征提取处添加 SimAM 注意力机制，以调整模型对关键区域的注意力权重，给予重点目标更多关注，提升头盔佩戴的检测性能。最后将 CIOU 损失函数改进为 WIOU 损失函数，进一步提升了模型的定位性能和检测精度。改进 YOLOv5s 算法模型结构如图 2 所示。

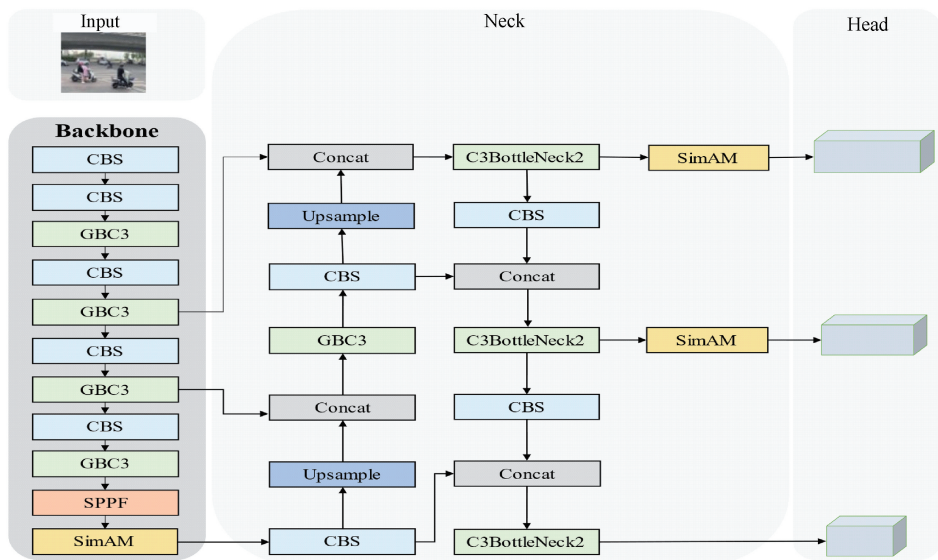

图 2　改进 YOLOv5s 算法模型结构

2.1 GBC3 模块

YOLOv5s 网络结构中的 C3 模块，采用由普通卷积构成的残差模块，但是这种模块未能充分融合多尺度特征信息，缺乏实现高阶空间交互的能力，未能有效地解决对头盔目标的漏检、误检问题。因此引入由递归门控卷积(g^nConv)构成的 GBC3 模块，对 C3 模块中的残差结构和标准卷积块进行改进，以增强邻近特征的空间信息交互和上下文语义信息的理解，获取目标更多的细节信息，提升模型对头盔目标的检测能力。

GBC3 模块主要由 GBlock、g^nConv 组成。GBlock 是采用 Transformer 的块设计进行构建，包括多层感知机(MLP)和层归一化(Layer Norm)。根据原 C3 模块中的结构，将 C3 中的 3 个标准卷积替换为 g^nConv，然后用 N 个改进的 GBlock 模块替换残差模块。GBC3 模块如图 3 所示。

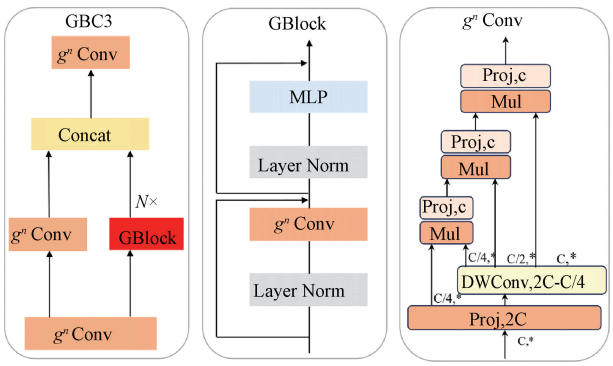

图 3 GBC3 模块结构

GBC3 模块中主要部分为 g^nConv，此卷积可实现长期与高阶空间交互，拥有与自注意力相似的输入自适应空间混合的能力。输入特征通过递归门控卷积，首先是进入一个卷积层，将输入通道数 C 调整为原来的 2 倍，并获取到一组 p_0 与 $\{q_k\}_{k=0}^{n-1}$ 的投影特征，如式(1)所示。

$$[p_0^{HW\times C_0}, q_0^{HW\times C_0}, \cdots, q_{n-1}^{HW\times C_{n-1}}] = \varphi_{in}(x) \in \boldsymbol{R}^{HW\times(C_0+\sum_{0\leqslant k\leqslant n-1}C_k)} \tag{1}$$

然后进行门控卷积的递归操作，特征 q_k 进入卷积核大小为 7×7 的深度可分离卷积，然后和每一块 q_0 的特征与前一块 p_k 的特征进行逐元素相乘，再乘于 $1/\alpha$ 的比例因子对输出进行缩放，以实现当前特征和邻近特征的信息交互，得到输出特征 p_{k+1}。最后再通过大小为 1×1 的卷积，保持输出与输入通道数相同。

$$p_{k+1} = f_k(q_k)\odot d_k(p_k)/\alpha \quad k=0,1,\cdots,n-1 \tag{2}$$

式中：$\{f_k\}$ 代表一组深度卷积层；$\odot$ 代表逐元素相乘。$\{d_k\}$ 代表匹配不同顺序维度的卷积操作；缩放因子 $1/\alpha$ 用于网络的稳定训练。

另外，为避免在进行高阶交互时引入较多的计算量，把每阶内的通道维度定义为：

$$C_k = \frac{C}{2^{n-k-1}}, \quad 0\leqslant k\leqslant n-1 \tag{3}$$

改进后的 GBC3 模块，既提高了模型对特征的提取和表达能力，又增强了特征融合能力，使得目标在经过多层卷积后，细节信息丢失较少。

2.2 SimAM 注意力机制

在拍摄图像过程中，由于成像条件、视角等因素，电动车驾乘者的头盔的分辨率会有所降低，其像素占比较小，容易影响检测效果。同时由于道路背景复杂，存在较多背景物体对模型识别头盔目标造成干扰。在进行多次卷积后也会导致目标细节信息的丢失。因此引入注意力机制使模型更加关注图像中驾乘者的头部区域特征，提高模型对头盔目标的感知能力，进而提升模型检测的准确性。

现有的注意力机制只能在通道域或空间域调整特征的权重，限制了模型学习注意力权重时的灵活性，且会增加参数量。所以选择能同时注意通道和空间域，并获取到 3-D 权重的 SimAM 注意力机制[33]，在不增加额外参数的同时，调整网络在特征图上不同区域的注意力权重，给予驾乘者头部区域更多关注，提升头盔佩戴的检测性能。

SimAM 是建立在视觉神经科学和空间抑制理论的基础上的模块，拥有丰富信息的神经元，在处理视觉任务时比周围神经元表现出更显著的活跃性，并对周围神经元产生空间抑制现象，在检测任务中这些承载着更多重要信息的神经元应赋予更高的权重。基于上述理论定义了一个最小能量函数，来得出头盔特征图中每个神经元所占权重。SimAM 注意力机制的神经元最小能量函数 e_t^* 和重要参数如下：

$$e_t^* = \frac{4(\hat{\sigma}^2+\lambda)}{(t-\hat{\mu})^2+2\hat{\sigma}2+2\lambda} \tag{4}$$

$$\hat{\mu} = \frac{1}{M}\sum_{n=1}^{M}x_n \tag{5}$$

$$\hat{\sigma}^2 = \frac{1}{M}\sum_{n}^{M-1}(x_n-\hat{\mu})^2 \tag{6}$$

式中：t 为输入特征在单通道内的目标神经元；n 表示在空间维度的索引；$M=H\times W$ 为神经元在当前通道的总数量。式(4)表明了在能量 e_t^* 越低时，神经元 t 和它周围的神经元之间的差别就越显著，并且对视觉信息处理也更重要。所以，每一个神经元的注意力重要性可用 $1/e_t^*$ 来表示。最终优化后的输出结果 $\tilde{X}$ 如式(7)所示，E 代表头盔特征图中神经元 e_t^* 值的集合。

$$\tilde{X} = \mathrm{sigmoid}\left(\frac{1}{E}\right)\odot X \tag{7}$$

其中，E 把全部的 e_t^* 在跨通道和空间维度进行分组，引入单峰函数 Sigmoid 来限制 E 值过大。SimAM 注意力机制如图 4 所示。

2.3 边界框回归损失函数

YOLOv5s 算法采用 CIOU 损失函数，通过使用边界

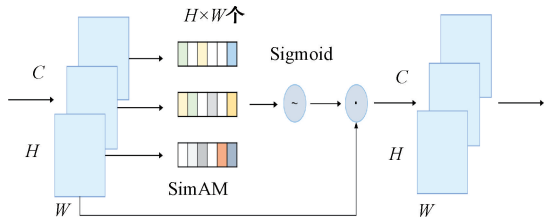

图4 SimAM注意力

框的重叠区域、长宽比以及中心点距离作为惩罚项，但未考虑到训练集中高、低质量样本的竞争问题。并且目前的目标检测研究大多数都假设训练数据中的样本是高质量的，侧重于增强边界框回归损失的拟合能力，但自制头盔训练数据集中难以避免会包含一些低质量样本，倘若盲目的增强低质量样本的边界框回归，将会损害模型的定位性能。

针对以上问题，本文使用具有动态非单调聚焦机制的WIOU损失函数[34]替换CIOU损失函数，以提升模型的定位性能和检测精度。由于距离和长宽比等几何因素会加剧对低质量示例的惩罚，为了削弱这种惩罚，并在训练过程中减少干预，由此提出了WIOUv1，如式(8)所示。

$$L_{WIOUv1}=R_{WIOU}LIOU \tag{8}$$

$$R_{WIOU}=\exp\left(\frac{(A-A_{mt})^2+(B-B_{mt})^2}{(W_m^2+H_m^2)^*}\right) \tag{9}$$

其中，惩罚项R_{WIOU}会扩大普通质量锚框的L_{IOU}，而L_{IOU}则会降低高质量锚框的R_{WIOU}，式(9)中W_m与H_m为最小闭合矩形框的宽高尺寸(图5)。为避免$\mathcal{R}_{WIOU}$产生引起阻碍收敛的梯度，把W_m和H_m从计算式中分离出来，上标*代表该操作。

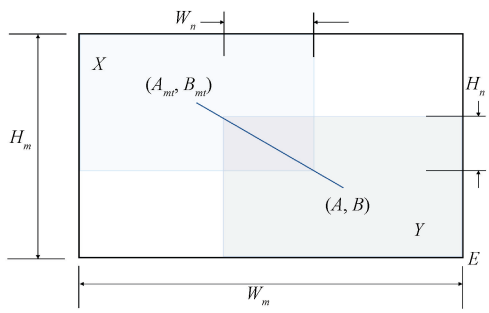

图5 最小闭合矩形框图E

动态非单调聚焦机制锚框的离群度定义为：

$$\beta=\frac{L_{IOU}^*}{\overline{L_{IOU}}}\in[1,+\infty) \tag{10}$$

式中：L_{IOU}^*为梯度增益；$\overline{L_{IOU}}$是带有动量m的指数加权移动平均值，解决了训练后期收敛慢的问题。离群度小代表锚框是高质量的，离群度大意味着锚框为低质量的。然后用β构造一个非单调聚焦系数τ并将其应用于WIOUv1：

$$L_{WIOUv3}=\tau L_{WIOUv1},\quad \tau=\frac{\beta}{\delta\alpha^{\beta-\delta}} \tag{11}$$

其中，α、β为超参，本文经过大量实验得出，采用$\alpha=1.8$，$\beta=4$时，改进模型的检测效果最佳。

$\overline{L_{IOU}}$是动态的，所以最终选用的WIOUv3的锚框质量分级标准也是动态变化的，可以根据当前情况的每个时刻，自适应地调整最符合的梯度增益分配方式，给低质量锚框分以小梯度增益来减少产生有害梯度，给高质量的锚框分以较小的梯度增益，来使边界回归框集中于普通质量的锚框上，以提升模型的定位性能，提高电动车头盔检测的准确性。

3 实验及结果分析

3.1 实验环境介绍

本文对改进模型进行训练与测试的实验环境平台基于Ubuntu(20.04)操作系统，具体环境配置如表1所示。

表1 实验环境平台配置

配置名称	版本参数
GPU	NVIDIA GeForce RTX 3090
CPU	Intel(R) Xeon(R) Platinum 8358P CPU @ 2.60 GHz
深度学习框架	PyTorch1.11.0
编程语言	Python3.8
CUDA	Cuda11.3

3.2 实验数据集

1)数据集来源

目前电动车驾乘者头盔佩戴仍未有公共数据集，因此本文所使用的数据集为自制头盔数据集，此数据集一部分来源于南京的地铁站附近，和一些街道交通路口的实地进行拍摄，共有2 672张，其他来源于网络爬虫所得，有408张。另外为了丰富本数据集的路况背景，还融合了来自bike helmet dataset的122张自行车骑行图片，最后再将所有数据放在一起，共有3 202张。图片的像素大小为768×1 024，图片尺寸大小为10.67 cm×14.22 cm，除了加入的122张自行车图片，其余3 080张图片均至少含有一个电动车驾驶者，并且有118张图片存在电动车驾乘者佩戴、遮阳帽等未按规定佩戴好头盔的情况，存在遮挡场景的有586张图片，车辆密集的有447张，一车多人的情景有359张，同时存在密集和遮挡的情况有432张，同时存在遮挡、车辆密集以及一车多人的复杂场景图片有221张。

2)数据集标注与划分

为有效地把驾乘者驾驶电动车与未驾驶状态下的电动车、骑行者驾驶自行车、行人区分开，在数据集标注时将会把电动车驾乘者佩戴头盔的情况标注为helmet，类别为0。其次将电动车与驾乘者一起标注为motor，类别为1；最后将未佩戴头盔的情况标注为head，类别为2；自行车、自行车驾驶员和行人不进行标注。数据集标注工具为

Labelimg，标注完成后将图片输出为 VOC 格式的 xml 文件。将训练集与测试集按照 7∶2∶1 的比例进行划分，其中训练集 2 242 张，验证集 640 张，测试集 320 张。

3.3 训练策略与超参数设置

本文在电动车头盔佩戴检测训练中，采取迁移学习的方式，加载 YOLOv5s 的官方预训练权重，然后将自制数据集在改进的 YOLOv5s 网络模型迭代 200 轮，batch-size 设置为 32，采用 SGD 优化器，并使用 Mosic 增强的策略。算法的部分超参数设置如表 2 所示。

表 2 算法的部分超参数

超参数名称	数值
lr0	0.01
lrf	0.1
momentum	0.937
batch-size	32
epochs	200
workers	6

3.4 算法评估指标

为了准确评估改进后 YOLOv5s 模型检验电动车驾乘者头盔佩戴状况的准确性，主要选择准确率(precision，P)、召回率(recall，R)、参数量(Params)、帧率和平均精度均值(mean average precision，mAP)作为算法评估指标。P 为在总检测样本中，预测为正样本的比例；R 为正确预测正样本的数量占实际正样本数量的比例；参数量表示算法模型中需要学习的可调整参数的总量；帧率代表每秒处理的图像帧数，其值越高表示推理速度越快；mAP 为全部类别下的 AP 平均值，AP 表示每一个召回率点下准确率的平均值。

$$P = \frac{TP}{TP + FP} \tag{12}$$

$$R = \frac{TP}{TP + FN} \tag{13}$$

$$AP = \int_0^1 P(r)\mathrm{d}(r) \tag{14}$$

$$mAP = \frac{1}{n_c}\sum_{i=1}^{n_c} AP_i \tag{15}$$

式中：TP 表示正样本被模型正确预测的数量；FP 表示模型误检的负样本数；FN 表示漏检的正样本数；n_c 为总类别数。

3.5 实验结果与分析

1)实验结果对比

对改进前后的 YOLOv5s 模型训练，将训练结束后的实验数据得到 PR 曲线，如图 6 所示。从图 6 可以看出，改进 YOLOv5s 模型对各类目标的检测精度均有显著提升，mAP 提高了 3.2%。其中，改进模型在检测 motor 类时尤为准确，达到了 99.2%，helmet 类的 AP 值提升了 1.5%，head 类 AP 值提升最高，表明改进模型较大地提高了对驾乘者是否戴头盔的检测性能。

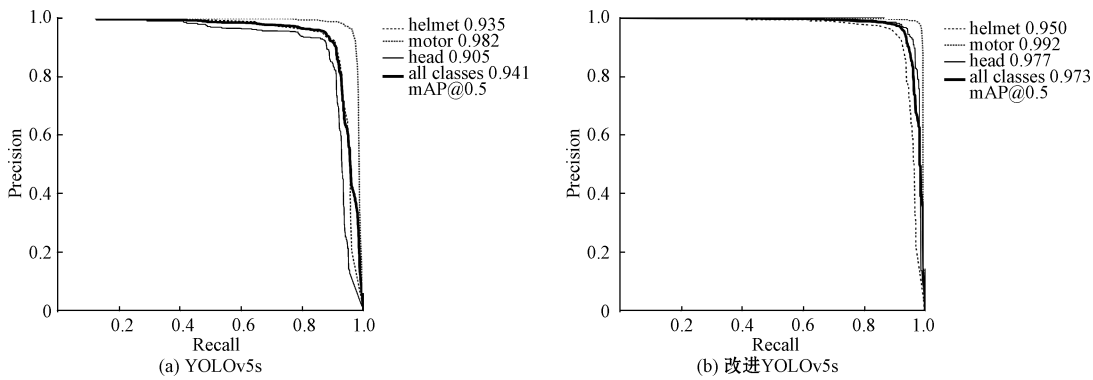

图 6 PR 曲线对比

2)消融实验

为了验证改进 YOLOv5s 算法对头盔的检测性能，将每种改进策略在相同配置环境中用自制电动车头盔数据集进行训练和测试。如表 3 所示，改进算法 1 为添加 GBC3 模块，改进算法 2 为引入注意力机制 SimAM，改进算法 3 为替换 WIOU 损失函数。

表 3 消融实验

算法	SimAM	GBC3	WIOU	P/%	R/%	mAP/%	帧率/fps
YOLOv5s				92.8	90.6	94.1	90.0
改进算法 1		√		95.7	93.4	96.4	88.4
改进算法 2	√			93.9	91.3	95.2	89.2
改进算法 3			√	92.6	91.0	94.8	89.6
本文算法	√	√	√	96.3	94.8	97.3	87.1

由表3看出，改进算法1对模型的检测性能改善较大，mAP提高了2.3%，表明在原算法改进的GBC3模块在增加较少的参数量下，增强了空间信息的交互和上下文语义信息，使得特征信息丰富融合。改进算法2的mAP比原YOLOv5s算法提高了1.1%，这表明添加SimAM后，模型能够更加关注驾乘者头部区域，提高了识别小目标的能力。改进算法3的mAP提升了0.7%，表明WIOU损失函数能够较为有效的平衡自制数据集中锚框的质量水平，提高模型对目标边界框回归的能力，提升了定位性能。实验结果综合分析，几种改进策略的推理速度在降低较少的情况下，均提升了YOLOv5s的检测精度。且经综合改进后，本文算法的mAP达到了97.3%，比原YOLOv5s算法高出3.2%，召回率和准确率分别提高了3.5%、4.2%，检测效果最佳。

3)与其他算法的对比实验

为进一步验证本文改进算法对电动车头盔佩戴检测的优势和有效性，在相同实验条件下，在自制电动车头盔数据集上将YOLOv4、PP-YOLO、YOLOX-s、YOLOv7、YOLOv5s等8种一阶段目标检测算法和改进YOLOv5s算法进行对比试验，以mAP、GFLOPs、Params、帧率为评价指标。实验对比结果如表4所示。

表4 算法模型对比实验

算法	mAP/%	GFLOPs	Param /($\times10^6$)	帧率 /fps
YOLOv4	93.5	57.9	62.9	66.8
YOLOv5m	94.9	51.2	20.8	79.2
PP-YOLO[35]	94.5	45.3	45.2	81.3
PP-YOLOv2[36]	95.8	116.2	54.8	83.5
YOLOX-s[37]	95.4	25.6	8.92	88.6
YOLOv7	96.6	110.3	52.3	87.4
YOLOv5s	94.1	15.8	7.02	90.0
改进YOLOv5s	97.3	17.4	7.68	87.1

由表4可见，在自制电动车头盔数据集上的对比实验结果中，改进YOLOv5s模型的mAP有着显著的提升；与YOLOv4、YOLOv5m、PP-YOLO、PP-YOLOv2算法的检测结果对比，改进YOLOv5s模型的检测精度有较大提高，推理速度更快，且计算量和参数量远低于这5个模型；与YOLOX-s模型相比，虽然改进YOLOv5s模型推理速度略低，但mAP相比其提高了1.9%；与推理速度相当的YOLOv7模型相比，改进的YOLOv5s模型的mAP提高了0.7%，参数量和计算量远远低于YOLOv7；与原YOLOv5s模型相比较，改进YOLOv5s模型在推理速度略微下降的情况下，其mAP提高了3.2%，达到了97.3%。综合各算法的复杂度和实际检测性能来看，改进YOLOv5s算法模型在这些算法中拥有较良好的表现，能满足实时检测的需要，验证了改进模型对头盔佩戴状态检测的实际有效性。

4)实际道路场景检测

为了检验改进的YOLOv5s模型对电动车头盔检测的实际效果，在真实的道路场景进行测试，并与YOLOv5s模型的测试结果进行对比，实验对比结果如图7所示。

图7(a)为一车多人的场景下的检测结果，YOLOv5s未将没戴头盔的坐乘者检测出来，而改进YOLOv5s并未出现漏检情况；图7(b)为存在遮挡场景下的检测结果，YOLOv5s模型未将中间右侧电动车及其驾乘者的头盔佩戴情况检测出来，而改进模型均能正确检测；图7(c)为存在遮挡且车辆密集的场景下的检测结果，可以看出，YOLOv5s模型未检测出白衣男左侧的电动车，并且将右前方驾驶者的草帽误检为头盔，改进模型则将漏检电动车以及误检的头盔佩戴状态正确检测出来。图7(d)为车辆种类复杂且目标较小的路口场景下的检测结果，YOLOv5s模型仅检测到远距离的两个电动车，而改进模型将佩戴头盔情况也检测出来了。

经上述结果分析，原因是添加GBC3模块能够有效增强邻近目标特征信息的交互和上下文语义信息的理解，从

(a) 一车多人的场景

(b) 存在遮挡的场景

(c) 存在遮挡且车辆密集的场景

(d) 车辆种类复杂且目标较小的路口场景

图 7　原模型与改进模型在不同道路场景的实验结果对比

而提升模型对头部区域遮挡时的检测能力；替换 WIOU 损失函数使得边界框回归更为精准，提升了电动车和头盔等目标的定位性能；同时引入 SimAM 注意力机制，减少了头部区域的信息丢失，更精确地捕捉重要目标的细节信息，改善了模型对复杂路口场景下远距离目标的检测能力。

4　结　论

针对电动车驾乘者的头盔佩戴在有遮挡、车流量大的复杂道路场景下，出现的漏检、误检以及检测精度低的问题，本文基于 YOLOv5s 模型进行了一系列改进。首先用改进 GBC3 模块替换主干和特征融合层的 C3 模块，以增强空间信息交互能力和上下文语义理解，使深浅特征信息丰富融合。其次引入 SimAM 注意力机制，抑制无关背景的干扰，给予头盔目标更多关注。最后，使用 WIOU 损失函数替换 CIOU 损失函数，调整自制数据集的难易样本权重，使模型的收敛效果更好，预测框更为精准。在自制头盔数据集上进行实验，结果表明与 YOLOv5s 模型及其他主流模型相比，改进后的 YOLOv5s 算法具有更好的检测性能。同时在有遮挡和密集车辆的复杂道路场景下，漏检和误检情况更少，模型鲁棒性更好。未来的研究中会考虑丰富数据集，补充摩托车驾驶者头盔检测，并进一步优化网络模型，使其在雨、雾、雪等恶劣的天气情况下仍能具有良好的检测表现。

参 考 文 献

[1] 李政谦，刘晖. 基于深度学习的安全帽佩戴检测算法综述[J]. 计算机应用与软件，2022，39(6)：194-202.

[2] 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.

[3] 高腾，张先武，李柏. 深度学习在安全帽佩戴检测中的应用研究综述[J]. 计算机工程与应用，2023，59(6)：13-29.

[4] LOWE D G. Distinctive image features from scale-invariant keypoints[J]. International Journal of Computer Vision，2004，60(2)：91-110.

[5] LU M Q，HU Y C，LU X B. Dilated light-head R-CNN using tri-center loss for driving behavior recognition[J]. Image and Vision Computing，2019，90(C)：108-121.

[6] REN S，HE K，GIRSHICK R，et al. Faster R-CNN：Towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence，2017，39(6)：1137-1149.

[7] HE K，GKIOXARI G，DOLLAR P，et al. Mask R-CNN[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence，2020，42(2)：386-397.

[8] DAI J F，LI Y，HE K M，et al. R-FCN：Object detection via region-based fully convolutional networkers[C]. Proceedings Systems. Newyork：Curran Associates Inc.，2016：379-387.

[9] 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.

[10] REDMON J，FARHADI A. YOLO9000：Better，faster，stronger[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition，2017：7263-7271.

[11] REDMON J，FARHADI A. YOLOv3：An incremental improvement[J]. ArXiv Preprint，2018，ArXiv：1804.02767.

[12] BOCHKOVSKIY A，WANG C Y，LIAO H Y M.

YOLOv4: Optimal speed and accuracy of object detection[J]. ArXiv Preprint, 2020, ArXiv: 2004.10934.
[13] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot multibox detector[C]. European Conference on Computer Vision. Springer, 2016:21-37.
[14] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection[C]. Proceedings of the IEEE International Conference on Computer Vision, 2017:2980-2988.
[15] 朱硕,黄剑翔,汪宗洋,等.基于深度学习的非机动车头盔佩戴检测方法研究[J].电子测量技术,2022,45(22):120-127.
[16] WOO S, PARK J, LEE J Y, et al. CBAM: Convolutional block attention module[C]. European Conference on Computer Vision,2018:3-19.
[17] 刘雪纯,刘大铭,刘若晨.改进的轻量级安全帽佩戴检测算法[J].液晶与显示,2023,38(7):964-974.
[18] 张瑞芳,董凤,程小辉.改进 YOLOv5s 算法在非机动车头盔佩戴检测中的应用[J].河南科技大学学报(自然科学版),2023,44(1):44-53,7.
[19] STERGIOU A, POPPE R, KALLIATAKIS G. Refining activation downsampling with SoftPool [C]. 2021 IEEE/CVF International Conference on Computer Vision (ICCV). IEEE, 2021:10337-10346.
[20] TAN M, PANG R, LE Q V. EfficientDet: S-calable and efficient object detection[C]. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE,2020:10781-10790.
[21] HOU Q, ZHOU D, FENG J. Coordinate attention for efficient mobile network design[C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021:13713-13722.
[22] 张子涵,袁栋,张经炜,等.改进 YOLOv5s 的非机动车头盔佩戴检测方法[J].软件,2023,44(3):37-43.
[23] WANG J Q, CHEN K, XU R, et al. Carafe: Content-aware reassembly of features[C]. Proceedings of the IEEE International Conference on Computer Vision. IEEE,2019:3007-3016.
[24] ZHENG Z H, WANG P, LIU W, et al. Distance-IoU loss: Faster and better learning for bounding box regression[C]. Proceedings of the AAAI Conference on Artificial Intelligence,2020:12993-13000.
[25] 朱周华,齐琦.基于改进 YOLOv5s 电动车头盔的自动检测与识别[J].计算机应用,2023,43(4):1291-1296.
[26] 谢溥轩,崔金荣,赵敏.基于改进 YOLOv5 的电动车头盔佩戴检测算法[J].计算机科学,2023,50(S1):420-425.
[27] WANG Q, WU B, ZHU P, et al. ECA-Net: Efficient channel attention for deep convolutional neural networks[C]. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. IEEE, 2020:11531-11539.
[28] HE J, ERFANI S, MA X, et al. Alpha-IoU: A family of power intersection over union losses for bounding box regression[J]. Advances in Neural Information Processing Systems, 2021, 34:20230-20242.
[29] 吕志轩,魏霞,马志钢.改进 YOLOX 的轻量级安全帽检测方法[J].计算机工程与应用,2023,59(1):61-71.
[30] 张鑫,周顺勇.改进 YOLOv5s 的摩托车头盔佩戴检测算法[J].四川轻化工大学学报(自然科学版),2023,36(3):50-58.
[31] 纪超,侯威,高鸣江,等.基于 MBDC 和双重注意力的变电站人员穿戴检测[J].电子测量与仪器学报,2023,37(6):247-255.
[32] 闫钧华,张琨,施天俊,等.融合多层级特征的遥感图像地面弱小目标检测[J].仪器仪表学报,2022,43(3):221-229.
[33] YANG L, ZHANG R Y, LI L, et al. SimAM: A simple, parameter-free attention module for convoluteonal neural networks[C]. Proceedings of the 38th International Conference on Machine Learning, 2021:11863-11874.
[34] TONG Z, CHEN Y, XU Z, et al. Wise-IoU: Bounding box regression loss with dynamic focusing mechanism [J]. ArXiv Preprint, 2023, ArXiv: 2301.10051.
[35] LONG X, DENG K, WANG G, et al. PP-YOLO: An effective and efficient implementation of object detector[J]. ArXiv Preprint, 2020, ArXiv: 2007.12099.
[36] HUANG X, WANG X, LV W, et al. PP-YOLOv2: A practical object detector[J]. ArXiv Preprint, 2021, ArXiv:2104.10419.
[37] GE Z, LIU S, WANG F, et al. YOLOx: Exceeding yolo series in 2021 [J]. ArXiv Preprint, 2021, ArXiv:2107.08430.

作者简介

侯恩翔,硕士研究生,主要研究方向为深度学习、目标检测。

张旭,硕士研究生,主要研究方向为机器视觉、人工智能。

刘罡(通信作者),硕士,高级工程师,硕士生导师,主要研究方向为深度学习、机器视觉、移动通信、物联网等。E-mail:mrliug@163.com

张秀再,博士,教授,硕士生导师,主要研究方向为气象通信技术、量子通信技术、机器学习、电子测量与控制技术等。