

文章编号:1000-1638(2025)01-0077-11

DOI:10.13484/j.nmgdxxbzk.20250109

基于深度学习的道路小目标检测改进算法^{*}

关浩¹,关宇²,王宇峰³,张宏³

(1. 内蒙古自治区交通运输事业发展中心,呼和浩特 010020;

2. 呼和浩特市交通运输服务中心,呼和浩特 010010; 3. 内蒙古大学交通学院,呼和浩特 010070)

摘要:针对现有道路目标检测技术在小目标识别方面的挑战,如精度低、误检率高和漏检问题,提出一种基于YOLOv5的目标检测改进方法(GCE-YOLOv5),旨在提高小目标检测能力。首先,在主干网络上集成全局注意力机制;其次,在头部网络中增加小目标检测层;最后,将原来的损失函数改为高效交并比损失函数。在KITTI数据集上评估了GCE-YOLOv5算法,并将性能与YOLOv5算法进行了比较。结果表明,GCE-YOLOv5将检测 $mAP@0.5$ 提高了1.6%, P 提高了1.5%。与YOLOv5算法相比,GCE-YOLOv5模型检测道路小目标的精度更高,有效改善了小目标漏检问题,具有更好的检测效果。模型检测速度满足实时性的要求,适合集成到自动驾驶系统中,用于提升车辆在复杂道路环境中的感知能力。

关键词:智能交通;小目标检测;YOLOv5;IoU;GAM

中图分类号:U495;TP391.4 **文献标志码:**A

2014年以前,目标检测多采用机器学习的传统检测方法,具有代表性的方法有维奥拉-琼斯目标检测框架(Viola-Jones object detection framework)、方向梯度直方图检测器(Histogram of oriented gradients detector, HOG)以及可变形部件模型(Deformable part model, DPM)^[1]。2014年以后,伴随深度学习的发展,以卷积神经网络为基本框架的目标检测方法成为主流。Girshick等^[2]提出了区域卷积神经网络(Region convolutional neural network, R-CNN)算法,该算法选择性地产生候选区域,利用神经网络提取特征,使用支持向量机进行分类和预测^[3]。但R-CNN算法区域选择效率低下,导致计算成本大量增加。基于R-CNN算法,研究人员陆续提出了检测效果更好的Fast-RCNN和Faster-RCNN算法,减少了模型的计算量。

在R-CNN算法之后,Redmon等^[4]提出了速度更快的你只看一次(You only look once, YOLO)算法。YOLO算法从回归问题的角度看待目标检测任务,用单个神经网络对目标的类别及位置信息进行预测^[5]。虽然YOLO算法检测速度优异,但存在定位精度不高、困难样本处理能力弱、目标类别不平衡等缺点,尤其在小目标检测上。2016年,Liu等^[6]提出了单步多框目标检测(Single shot multibox detector, SSD)算法,该算法对特征图进行多尺度检测,使其能够有效地检测不同尺寸的目标,提高了检测的准确性,尤其适用于密集目标检测,但是在整体感知和实时性检测方面不如YOLO

^{*} 收稿日期:2024-04-10;修回日期:2024-08-19

基金项目:国家自然科学基金项目(62362053);内蒙古高等学校青年科技英才支持计划项目(NJYT23060);“内蒙古科技成果转移转化示范区、高校协同创新基地和高校创业实训基地”建设项目(超级算力项目)(21300-231510);2024年度基础研究与应用基础研究(2024-规-基-29)

作者简介:关浩(1967-),男(满族),内蒙古呼和浩特人,副高级经济师。主要从事公路建设、养护检查、监测、交通运输经济研究。E-mail:zhang2001hong@126.com

通信作者:张宏(1981-),女,内蒙古呼和浩特人,教授,博士。主要从事智能网联汽车研究。E-mail:zh-hong@foxmail.com

算法^[7]。

“小目标”有两种定义:绝对小目标和相对较小的目标。绝对小目标是指像素尺寸小于图像中的 32×32 像素的目标,相对较小的目标则指特定目标尺寸小于原始图像尺寸一定百分比的目标^[8-9]。本文小目标的定义采用前者,即绝对小目标。石春鹤等^[10]基于YOLOv3算法,优化损失函数,增加特征提取层数,以提高小目标检测性能;赖颖等^[11]在YOLOv4网络中添加新的特征层,以加强检测目标特征的多尺度特征融合能力;胡倩等^[12]在YOLOv5的骨干网络中引入了Conv-SPD网络模块代替传统的跨步卷积,缓解特征信息丢失现象,提高小目标检测的准确性。小目标由于信息量不足、目标遮挡和类别混淆等原因,现有目标检测方法在处理小目标时可能会忽略关键特征信息,导致检测精度不高。

针对道路小目标检测精度低,存在误检和漏检现象,本文提出一种基于YOLOv5的单阶段目标检测算法(GCE-YOLOv5)。首先在主干网络上集成全局注意力机制(Global attention mechanism, GAM)模块,提高主干网络对道路小目标特征信息的关注度;其次,增加小目标检测层,优化特征信息,减少目标检测的误差率;最后,将原先的损失函数改为高效交并比(Efficient intersection over union, EIou)损失函数,增加网络的定位精度,降低对有障碍物遮挡目标的漏检率。在KITTI公开数据集上评估模型改进效果。

1 YOLOv5 模型优化与改进

1.1 YOLOv5 检测模型

与YOLOv4相比,YOLOv5是新一代的单阶段检测方法,模型更小,训练速度更快,多用于实时检测。本文采用YOLOv5 v6.2作为基础架构,模型主要包括主干网络、颈部网络和头部网络三部分。主干网络采用CSPDarknet53对图像进行特征提取;颈部网络在主干网络之后加入了一些额外的卷积层,用于进一步提取特征并进行特征融合;头部网络有一系列的卷积层和池化层,用于生成检测目标的类别、位置和置信度等信息^[13]。为了提高YOLOv5对小目标的检测能力,本文对YOLOv5算法做了如下改进:将GAM注意力模块集成到主干网络中,头部网络中增加小目标检测层,并引入EIou损失函数。

1.2 集成 GAM 注意力模块

在车道目标检测中,道路上的车辆、行人等目标通常具有不同的尺寸、形状和位置。同时在检测过程中存在遮挡、多目标重叠等复杂情况。通过引入注意力机制,检测算法可以根据目标的重要性和特征的相关性,动态调整对不同区域的关注程度,从而更好地捕捉目标的特征并准确地进行检测,减少不必要的干扰,提高检测算法的鲁棒性和泛化能力^[14]。

常用的注意力机制包括空间注意力、通道注意力和混合注意力机制,不同的注意力机制适用于不同的检测任务^[15-16]。然而,由于空间位置和通道信息的分离,使得对图像特征提取不足。GAM注意力机制^[17]在减少信息损失的同时放大全局跨维度交互作用,以此来增强通道和空间之间的交互,从而保留并利用跨维度的重要信息。GAM通过采用顺序的通道-空间注意力机制和对子模块的重新设计,减少了信息缩减并放大了全局维度交互特征。GAM注意力机制运算过程如图1所示,其公式为

$$F_2 = M_C(F_1), F_3 = M_S(F_2),$$

其中 M_C 代表通道注意力, M_S 代表空间注意力, F_1 为输入特征, F_2 为中间状态, F_3 为输出特征。

1.2.1 通道注意力子模块

GAM中的通道注意力子模块首先通过一个3D排列来重组特征映射的维度,然后通过一个两层的多层感知器(MLP)来学习通道间的依赖关系,并放大跨维度的交互作用。通道注意力子模块如图2所示。

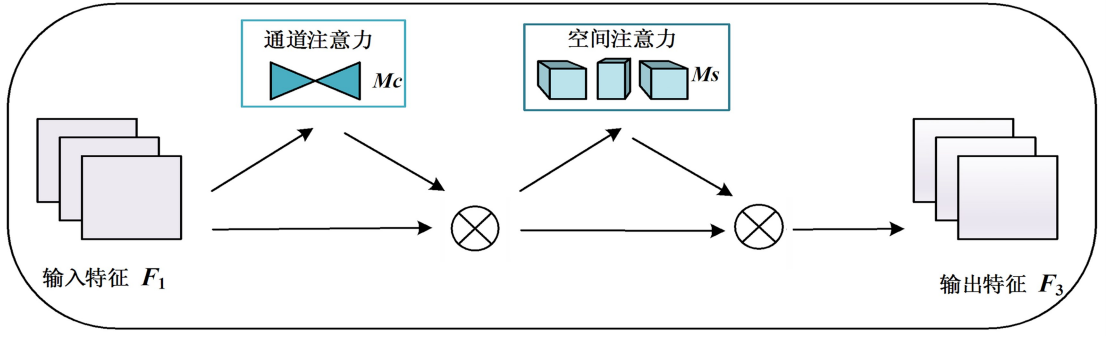


图 1 GAM 注意力机制运算过程

Fig. 1 Calculation process of global attention mechanism

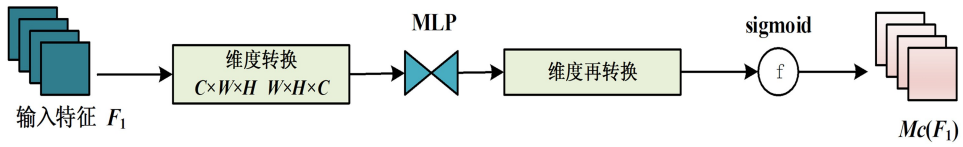


图 2 通道注意力子模块

Fig. 2 Sub-module of channel attention

1. 2. 2 空间注意力子模块

空间注意力子模块则关注特征图片的空间维度,通过卷积层来融合空间信息,并采用组卷积和通道混洗来防止参数的增加。空间注意力子模块如图 3 所示。

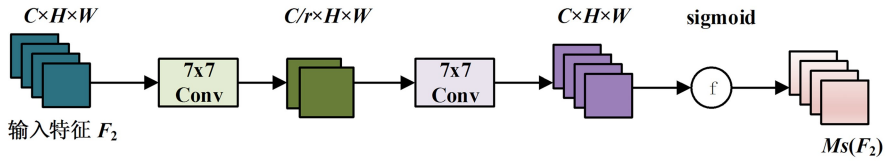


图 3 空间注意力子模块

Fig. 3 Sub-module of spatial attention

本文将 GAM 集成到 YOLOv5 主干网络第十一层,对上一层 C3 模块所提取到的空间以及通道特征信息在全局维度上进行交互与学习,增加有效特征的权重,抑制无效特征的权重,聚焦包含重要信息的目标区域。

1. 3 检测层的优化

YOLOv5 模型存在对小目标的漏检、错检的问题。由于 YOLOv5 的下采样乘数较大,使得图片远处小于下采样乘数的小目标被忽略,因此提出增加一个小目标检测层来识别远处的小目标^[18-19]。

由于 YOLO 算法采用单一尺度的特征图进行检测,模型对小尺寸目标的检测效果不佳,导致小目标检测效果不理想^[20],具体过程如图 4 红框所示。因此,在原网络基础上增加尺度为 160×160 的小目标检测层(完善了颈部网络的特征融合功能,并增加了额外的检测头),提升对远距离目标的检测效果。

首先将主干网络第六层(C3)的 80×80 尺度特征层在颈部网络 Concat 模块中张量拼接后继续向上堆叠,经 C3 模块、Conv 卷积及上采样处理(Upsample)后得到深层次语义特征层;然后与第四层的浅层次融合小目标语义特征以及位置信息,最后经 C3 模块多尺度融合以及跨通道信息传递后,将特征信息加载到额外检测头用于解码检测。在颈部网络将不同尺度层的语义信息特征融合,将特征信息通过网络传递至各层,从而加强网络的特征融合能力,提升小目标的检测精度。

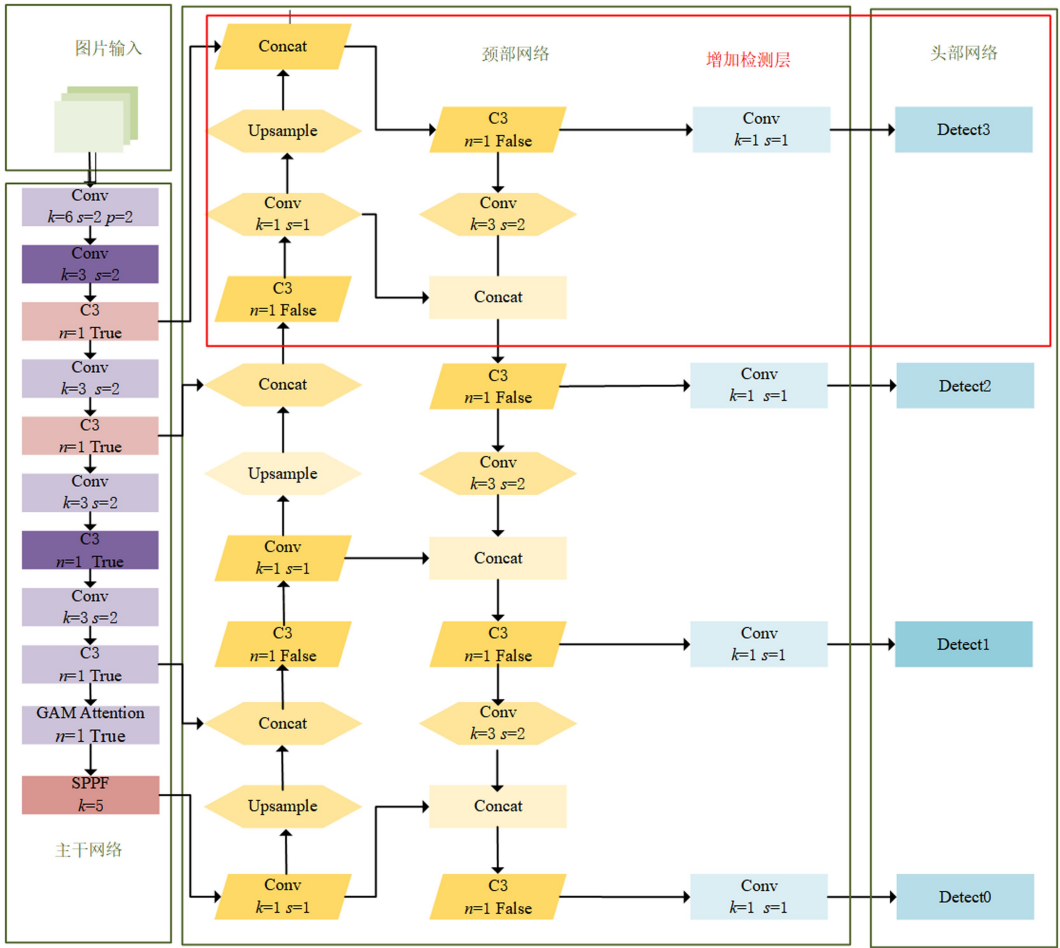


图4 添加小目标检测层后的 YOLOv5 模型

Fig. 4 Structural diagram of YOLOv5 network with added small object detection layer

1.4 损失函数优化

在 YOLOv5 算法中, IoU (Intersection over union) 是一个重要组成部分, 用于衡量预测框与真实框之间的重叠程度^[21]。IoU 是指预测框 B 与真实框 A 之间的交并比, 能够更直观地反映边界框的定位精度。图形表达如图 5 所示, 计算公式为

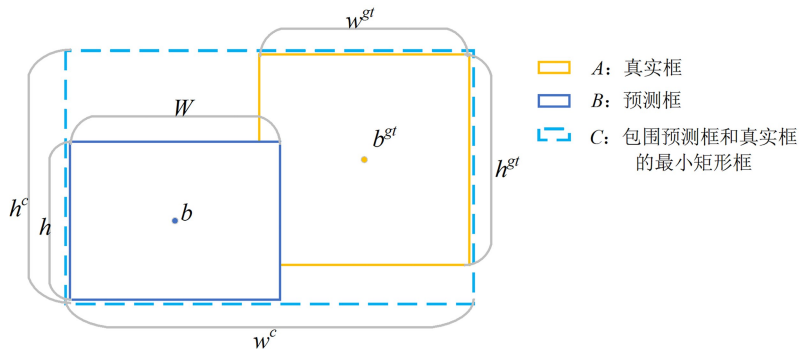


图5 预测框与真实框关系

Fig. 5 Relationship between prediction bounding box and ground truth

$$\text{IoU} = \frac{A \cap B}{A \cup B},$$

YOLOv5 算法早期版本使用 GIoU (Generalized IoU) 作为损失函数, 计算公式为

$$\text{GIoU} = \text{IoU} - \frac{C - (A \cup B)}{C},$$

式中, C 表示包围预测框和真实框的最小矩形。GIoU 不仅能够衡量预测框与真实框之间的重合程

度,还能反映两者之间的相对位置和重叠方式。这使得 GIoU 在目标检测任务中更具鲁棒性,特别是在预测框与真实框部分相交或不相交的情况下。由于 GIoU 可以直接用作优化的目标函数,因此在训练过程中能够更有效地调整预测框的位置和大小,提高模型的检测性能。然而,GIoU 过于关注最小矩形框,可能导致预测框与真实框之间的重叠区域较小。当预测框与真实框之间无交集时,GIoU 损失会增大预测框的面积,使其与真实框重合,这样使得其收敛速度较慢。

CIoU 在 IoU 的基础上加入了边界框回归和角度回归的信息,使它能够更好地衡量预测框与真实框之间的相似度。CIoU 计算公式为

$$\text{CIoU} = \text{IoU} - \left(\frac{\rho^2}{c^2} + \alpha v \right),$$

式中 ρ 是预测框与真实框中心点的欧氏距离, c 为预测框和真实框外接矩形的对角线距离, v 是宽高比的惩罚项, α 是一个可调的参数,用于平衡几何距离和宽高比惩罚的权重。 v 和 α 公式为

$$v = \frac{4}{\pi^2} \left(\tan^{-1} \frac{w^{\text{gt}}}{h^{\text{gt}}} - \tan^{-1} \frac{w}{h} \right)^2, \alpha = \frac{v}{(1 - \text{IoU}) + v},$$

式中 w^{gt} 和 h^{gt} 表示真实框的宽和高, w 和 h 表示预测框的宽和高。

CIoU 损失函数的优势在于通过引入修正因子,使得损失对于不同形状的目标框更具鲁棒性。虽然 CIoU 在很多情况下表现优异,但其需要计算更多的参数,计算过程相对复杂,增加了计算成本。为了简化计算步骤和减少不必要的计算量,实现更快速的收敛过程,EIoU(Efficient IoU)^[22]应运而生。EIoU 计算公式为

$$\text{EIoU} = \text{IoU} - \left(\frac{\rho^2}{c^2} + \frac{(\rho_w)^2}{(w_c)^2} + \frac{(\rho_h)^2}{(h_c)^2} \right),$$

式中, w_c 、 h_c 分别表示预测框和真实框的外接矩形的宽、高, ρ_w 、 ρ_c 分别表示预测框与真实框宽差值、高差值。可以看出,EIoU 损失包括 IoU 损失、距离损失、方向损失。EIoU 在 CIoU 的基础上将预测框和真实框的纵横比的影响因子拆开,分别计算预测框和真实框的长和宽,解决了 CIoU 收敛速度慢及定位不准确的问题。通过优化 EIoU 值,模型可以学习如何更好地调整预测框的位置和大小,以更准确地匹配真实框。这有助于提升模型的检测精度和召回率,从而改善道路目标检测的整体性能。

2 实验

YOLOv5 算法提供了 5 种不同比例的模型:YOLOv5n、YOLOv5s、YOLOv5m、YOLOv5l 和 YOLOv5x。每种比例模型都有不同的深度和宽度^[23]。5 种比例模型的结构相同,只是规模和复杂性有所不同,为保证自动驾驶的准确率和实时性,采用 YOLOv5s 模型对小目标的识别能力进行了实验。实验的环境参数配置如表 1 所示,训练轮数 200 轮,其余参数按照 YOLOv5 v6.2 初始值设置。

表 1 实验环境参数配置

Table 1 Parameter configuration of experimental environment

配置项	参数	配置项	参数
CPU	Core(TM)i5-7300HQ	深度学习框架	Pytorch 1.12.0
操作系统	Win10	GPU	Nvidia RTX 4060
CUDA	12.30		

2.1 数据集

KITTI 数据集是在多种不同场景中采集的真实世界图像^[24],数据集中类别多为汽车与行人,存在较多的小目标以及不同程度的被遮挡目标。KITTI 数据集共有 7481 张图片,按照 8:1:1 的比例划分为训练集、测试集和验证集。数据集中部分图像如图 6 所示。

2.2 评估标准

目标检测算法有 3 个性能评估指标： P (精确率)、 R (召回率)和 mAP (平均精度)。 P 、 R 和 mAP 的公式分别为

$$P = \frac{X_{TP}}{X_{TP} + X_{FP}} \times 100\%, R = \frac{X_{TP}}{X_{TP} + X_{FN}} \times 100\%, mAP = \frac{\sum_{i=1}^c AP_i}{c} \times 100\%,$$

式中, X_{TP} 是目标检测正确的数量, X_{FP} 是错误检测到的数量, X_{FN} 是未检测到的数量, c 是类别数量, AP 是单个目标类别的平均精度, 其公式为

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



图 6 KITTI 数据集中随机选择的图像

Fig. 6 Images randomly selected from the KITTI datasets

2.3 实验结果分析

2.3.1 添加注意力机制实验

本文提出了 G-YOLOv5 模型, 在 YOLOv5s 算法的主干网络中添加 GAM 注意力机制。为了验证添加 GAM 注意力机制能否提升模型检测小目标的效果, 在 KITTI 数据集上对比了 YOLOv5 算法和添加不同注意力机制的 YOLOv5 算法, 实验结果如表 2 所示。

表 2 添加不同注意力机制的结果

Table 2 Results with adding different attention mechanisms

网络模型	$P/\%$	$R/\%$	$mAP@0.5/\%$	$GFLOPs$
YOLOv5	93.2	89.6	93.2	16.1
YOLOv5-SE ^[25]	93.6	89.8	93.4	20.8
YOLOv5-CBAM ^[26]	93.3	89.5	93.9	16.5
G-YOLOv5	93.8	90.4	94.5	16.2

由表 2 可知,原始的 YOLOv5 算法的 $mAP@0.5$ 为 93.2%,在 YOLOv5s 算法上分别添加 SE 注意力机制、CBAM 注意力机制和 GAM 注意力机制的 $mAP@0.5$ 为 93.4%、93.9% 和 94.5%,相比于原始的 YOLOv5 算法, $mAP@0.5$ 分别提升了 0.2%、0.7% 和 1.3%。相比于添加另外两种注意力机制,在 YOLOv5 算法的主干网络中添加 GAM 注意力机制对于提升道路目标的识别效果最优。图 7 为检测效果对比图,左侧为 YOLOv5 算法检测效果,右侧为 G-YOLOv5 模型检测效果。通过对比,可以看出改进后的 G-YOLOv5 检测道路小目标的置信度更高,GAM 注意力机制在提升检测精度方面也有显著效果,且对模型的计算效率影响较小。



图 7 改进后模型在 KITTI 数据集上效果对比图

Fig. 7 Effect comparison diagrams of the improved model on the KITTI

小目标在图像中占据较小的空间,易于被忽略或误判。GAM 注意力机制可以帮助模型更有效地关注和析图像中的重要信息,从而提高道路小目标的检测准确性。在复杂的道路交通场景中,存在大量的冗余信息可能会干扰小目标的检测。通过注意力机制,模型可以有选择地关注与小目标相关的信息,从而减少冗余信息的影响,提高检测的效率和准确性。因此,本文提出的 G-YOLOv5 模型对于小目标的检测效果优于原始的 YOLOv5 算法。

2.3.2 添加小目标检测层实验

本文实验在 YOLOv5s 模型中添加一层小目标检测层,将改进后的模型命名为 C-YOLOv5。为了验证添加小目标检测层的效果,在 KITTI 数据集上对比了 YOLOv5 模型和 C-YOLOv5 模型,实验结果如表 3 所示。

表 3 添加小目标检测层的对比效果

Table 3 Comparison effects with adding small object detection layer

网络模型	$P/\%$	$R/\%$	$mAP@0.5/\%$	$GFLOPs$
YOLOv5	93.2	89.6	93.2	16.1
C-YOLOv5	93.6	90.2	94.5	16.6

由表 3 可以看出,通过添加小目标检测层,YOLOv5 模型的计算量增加了 0.5, $mAP@0.5$ 提高了 0.7%,改进后的 C-YOLOv5 模型在一定程度上提升了对道路小目标的检测效果。检测效果对比如图 8 所示,上侧为 YOLOv5 算法检测效果,下侧为 C-YOLOv5 模型检测效果,可以看出改进后的 C-YOLOv5 改善了 YOLOv5 的漏检现象。



图 8 KITTI 数据集上添加小目标层前后模型的效果对比图

Fig. 8 Effects comparison diagrams before and after adding small object layers on KITTI

2.3.3 更换损失函数实验

为了验证损失函数对于模型检测性能的影响,在 KITTI 数据集上进行评估,将 YOLOv5s 模型的损失函数更换为 EIoU,将改进后的模型命名为 E-YOLOv5,并与 YOLOv5 模型的检测效果进行对比,实验结果如表 4 所示,对比效果如图 9 所示。

表 4 更换损失函数的对比效果

Table 4 Comparison effect of changing loss function

网络模型	$P/\%$	$R/\%$	$mAP@0.5/\%$	$GFLOPs$
YOLOv5	93.2	89.6	93.2	16.1
E-YOLOv5	94.7	90.3	94.8	16.1



图 9 KITTI 数据集上更换损失函数前后模型的效果对比图

Fig. 9 Effects comparison diagrams before and after changing the loss function on KITTI

由表 4 可以看出,与使用 CIoU 损失函数的 YOLOv5 模型相比,使用 EIoU 损失函数的 E-YOLOv5 模型计算量未增加, $mAP@0.5$ 提升了 1.6%。因此,更换 EIoU 损失函数可以提升 YOLOv5 模型检测道路小目标的性能。由图 9 可知,相比于左侧 YOLOv5 模型的检测效果,右侧 E-YOLOv5 模型在 KITTI 数据集上检测的车辆置信度得分更高。相比传统的 CIoU, EIoU 对于目标形状和大小的变化更具有鲁棒性,能够更好地适应各种尺寸和形状的目标检测任务。EIoU 通过最小化预测框和真实框中心点之间的距离,提高边界框的定位准确性。通过惩罚宽度和高度的差异,确保预测框在形状上更接近真实框。因此,本文提出的 E-YOLOv5 模型在 KITTI 数据集上的道路目标检测精度更高。

2.3.4 消融实验

为了验证本文提出的 3 种改进方法同时作用在 YOLOv5 算法上的效果,以 YOLOv5s 模型为基础,集成注意力模块、增加小目标检测层以及更换损失函数,将最后得到的算法命名为 GCE-YOLOv5,选择 KITTI 数据集进行验证,实验结果如表 5 所示。

表 5 消融实验结果
Table 5 Ablation experiment results

网络模型	$P/\%$	$R/\%$	$mAP@0.5/\%$	$GFLOPs$
YOLOv5	93.2	89.6	93.2	16.1
YOLOv5+GAM 注意力机制	93.8	90.4	94.5	16.2
YOLOv5+GAM 注意力机制+小目标检测层	94.3	90.6	94.6	16.6
GCE-YOLOv5 (YOLOv5+GAM 注意力机制+小目标检测层+EIoU 损失)	94.7	90.9	94.8	16.7

由表 5 可知,相比于原始的 YOLOv5 算法,添加 GAM 注意力机制、添加小目标检测层和更换 EIoU 损失之后得到的模型计算量增加 0.6, $mAP@0.5$ 提升了 1.6%,精确率提高了 1.5%。在 KITTI 数据集上,GCE-YOLOv5 模型检测道路小目标的效果更好,这是因为 GCE-YOLOv5 模型有效地解决了小目标漏检和误检问题,提高了对小目标的识别和定位能力。

2.3.5 对比实验

为了进一步证明本文所提算法的有效性以及优越性,将本文提出的 GCE-YOLOv5 与 YOLOv3、YOLOv4、YOLOv5、Faster R-CNN、SSD 等经典模型在 KITTI 数据集上进行实验对比。所有实验均在相同参数设置下进行,实验结果如表 6 所示。

表 6 对比实验结果
Table 6 Comparison experiment results

网络模型	$P/\%$	$R/\%$	$mAP@0.5/\%$	$GFLOPs$
YOLOv3	78.3	68.3	87.5	54.7
YOLOv4	71.8	73.0	89.3	20.6
YOLOv5	93.2	89.6	93.2	16.1
Faster R-CNN	68.6	84.6	88.7	194.3
SSD	86.8	72.5	77.2	62.3
GCE-YOLOv5	94.7	90.9	94.8	16.7

由表 6 可知,本文提出的 GCE-YOLOv5 模型与基准模型相比,在 P 、 R 及 $mAP@0.5$ 等评价指标上均有较大提升。与 YOLOv3、YOLOv4、YOLOv5、Faster R-CNN、SSD 相比,本文的 GCE-

YOLOv5 模型精确率分别提升 16.4%、22.9%、1.5%、26.1%、7.9%，召回率分别提升 22.6%、17.9%、1.3%、6.3%、18.4%，在 $mAP@0.5$ 上分别提升 7.3%、5.5%、1.6%、6.1%、17.6%。说明本文提出的 GCE-YOLOv5 模型能够有效改善道路小目标的漏检情况，提升小目标检测精度。

3 结论

针对目前深度学习在检测道路小目标的过程中存在的精度低、漏检和误检问题，本文提出了一种基于 YOLOv5 的目标检测改进算法(GCE-YOLOv5)。在复杂的道路交通场景中，存在大量的冗余信息可能会干扰道路小目标的检测。通过添加 GAM 注意力机制，减少冗余信息的影响，使模型更有效地关注和分析图像中的重要信息，从而提升小目标检测的效率和准确性。添加小目标检测层，专门针对尺寸较小的目标进行优化，提高对小目标的识别和定位能力；将损失函数更换为 EIoU 损失函数，有效地预防了梯度消失的问题，加速了预测框的回归过程。EIoU 将损失函数与包含预测框和真实框的最小封闭框的尺寸相结合，从而使得损失更加敏感于目标的尺寸和位置，解决了 CIoU 收敛速度慢及定位不准确的问题。本文提出的 GCE-YOLOv5 模型的计算量增加 0.6， $mAP@0.5$ 提升了 1.6%，精确率提高了 1.5%。与 YOLOv5 算法相比，GCE-YOLOv5 模型检测精度更高，有效解决漏检现象，具有更好的鲁棒性。

虽然基于深度学习的道路小目标检测算法已经取得了一定的发展，但是在面对复杂的道路场景时依然存在如恶劣天气条件及遮挡物的挑战。未来可以着重于研究如何有效地应对天气条件、遮挡物等因素对道路小目标检测造成的干扰，提高算法在复杂场景中的鲁棒性。

参考文献：

- [1] 郭庆梅,刘宁波,王中训,等.基于深度学习的目标检测算法综述[J].探测与控制学报,2023,45(6):10-20,26.
- [2] GIRSHICK R,DONAHUE J,DARRELL T,et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Columbus,United States of America:IEEE,2014:580-587.
- [3] 周睿.基于注意力机制的高分辨率多任务网络研究[D].哈尔滨:哈尔滨工业大学,2022.
- [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,United States of America:IEEE,2016:779-788.
- [5] 柴宏伟.人车目标检测跟踪研究[D].武汉:华中科技大学,2022.
- [6] LIU W,ANGUELOV D,ERHAN D,et al. SSD:Single shot multiBox detector[C]//Computer Vision-ECCV 2016:14th European Conference,Amsterdam,The Netherlands:Springer International Publishing,2016:21-37.
- [7] 张阳婷,黄德启,王东伟,等.基于深度学习的目标检测算法研究与应用综述[J].计算机工程与应用,2023,59(18):1-13.
- [8] 曹健,陈怡梅,李海生,等.基于深度学习的道路小目标检测综述[J].计算机工程,2023,49(10):1-12.
- [9] 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.
- [10] 石春鹤,张浩楠.基于改进 YOLOv3 的小尺度车辆目标检测算法[J].沈阳大学学报(自然科学版),2024,36(3):213-220,273.
- [11] 赖颖,巨志勇,叶雨新.基于改进 YOLOv4 的车辆检测算法[J/OL].电子科技,1-8[2024-07-18].<https://doi.org/10.16180/j.cnki.issn1007-7820.2025.01.011>.
- [12] 胡倩,皮建勇,胡伟超,等.基于改进 YOLOv5 的密集行人检测算法研究[J/OL].计算机工程,1-12[2024-07-18].<https://doi.org/10.19678/j.issn.1000-3428.0068753>.
- [13] 袁磊,唐海,陈彦蓉,等.改进 YOLOv5 的复杂环境道路目标检测方法[J].计算机工程与应用,2023,59(16):212-222.
- [14] SOYDANER D. Attention mechanism in neural networks:Where it comes and where it goes[J]. Neural Computing and Applications,2022,34(16):13371-13385.
- [15] BRAUWERS G,FRASINCAR F. A general survey on attention mechanisms in deep learning[J]. IEEE Transac-

- tions on Knowledge and Data Engineering, 2021, 35(4): 3279-3298.
- [16] NIU Z Y, ZHONG G Q, YU H. A review on the attention mechanism of deep learning[J]. Neurocomputing, 2021, 452: 48-62.
- [17] 刘辉, 杨照华, 吴云, 等. 基于全局注意力机制的单像素成像图像增强方法[J]. 空间控制技术与应用, 2023, 49(6): 68-76.
- [18] 牛为华, 殷苗苗. 基于改进 YOLOv5 的道路小目标检测算法[J]. 传感技术学报, 2023, 36(1): 36-44.
- [19] 黄治翔, 张艺骞. 基于深度学习的目标检测综述[J]. 科技资讯, 2023, 21(24): 13-16.
- [20] SUN R Y, LI D W, LIANG S Y, et al. The global landscape of neural networks: An overview[J]. IEEE Signal Processing Magazine, 2020, 37(5): 95-108.
- [21] 刘辉, 刘鑫满, 刘大东. 面向复杂道路目标检测的 YOLOv5 算法优化研究[J]. 计算机工程与应用, 2023, 59(18): 207-217.
- [22] ZHANG Y F, REN W Q, ZHANG Z, et al. Focal and efficient IOU loss for accurate bounding box regression [J]. Neurocomputing, 2022, 506: 146-157.
- [23] 韩佳彤, 张宏娜, 李召波, 等. 基于 YOLOv5 网络模型的市政道路检测识别[J]. 内蒙古大学学报(自然科学版), 2021, 52(5): 514-519.
- [24] GEIGER A, LENZ P, STILLER C, et al. Vision meets robotics: The KITTI dataset[J]. International Journal of Robotics Research, 2013, 32(11): 1231-1237.
- [25] 罗舜, 于娟. 改进多尺度网络的行人目标检测算法[J]. 福州大学学报, 2022, 50(5): 587-594.
- [26] WOO S, PARK J, LEE J Y, et al. CBAM: Convolutional block attention module[C]//Computer Vision-ECCV 2018. Munich, Germany: Springer International Publisher, 2018: 3-19.

Improved Algorithm for Small Road Object Detection Based on Deep Learning

GUAN Hao¹, GUAN Yu², WANG Yufeng³, ZHANG Hong³

(1. Inner Mongolia Autonomous Region Transportation Development Center, Hohhot 010020, China;

2. Hohhot City Transportation Service Center, Hohhot 010010, China;

3. Institute of Transportation, Inner Mongolia University, Hohhot 010070, China)

Abstract: To cope with the challenges of existing techniques in small road object recognition, such as low accuracy, high false detection rate and missed detection, a YOLOv5-based object detection improvement method (GCE-YOLOv5) is proposed, aiming to improve the small object detection capability. Firstly, the Global Attention Mechanism (GAM) module is integrated on the backbone network; secondly, a small object detection layer is added to the header network; and finally, the original loss function is changed to Efficient Intersection over Union (EIoU) loss. We evaluated the performance of the GCE-YOLOv5 algorithm on the KITTI datasets, and compared it with the YOLOv5 algorithm. The results show that the GCE-YOLOv5 algorithm improves $mAP@0.5$ by 1.6% and P by 1.5%. Compared with the YOLOv5 algorithm, the GCE-YOLOv5 model has higher accuracy in detecting small road object, which effectively improves the problem of missed detection of small object and has better detection performance. The GCE-YOLOv5 model detection speed meets the real-time requirements, which make it suitable for being integration into autonomous driving systems to enhance vehicle perception capacity in complex road environments.

Key words: intelligent transportation; small object detection; YOLOv5; EIoU; GAM