

文章编号:1007-2780(2026)06-0878-16

双目视觉下基于 GGW-YOLOv8 的车辆 测距及识别技术

程 瑶, 唐青玲*, 刘云阳, 唐清涛, 李 伟, 刘 鹏
(重庆理工大学 机械工程学院, 重庆 400054)

摘要:针对传统目标检测和测距技术存在测距传感器昂贵、复杂路段检测精度低、检测速率慢的问题,本文提出一种基于 GGW-YOLOv8 的车辆前车距离测量及目标识别技术。该算法对 YOLOv8n 进行改进,在中目标和大目标检测头前引入全局注意力机制(GAM),通过减少信息损失和放大大局交互提高模型对车辆目标的特征提取能力;在颈端网络引入 GSConv 轻量化卷积模块,通过深度可分离卷积和混洗操作减少模型参数量和计算量;用 WIoU_v3 损失函数替换原有的 CIoU 损失函数,通过非单调动态聚焦机制的梯度增益分配策略提升模型的泛化能力和网络边界框的回归精度;搭建基于 SGBM 半全局立体匹配算法的双目视觉测距系统进行车辆测距,开发融合车辆目标检测与测距的上位机系统,并将其移植部署到嵌入式边缘设备,开展实车搭载检测。实验结果表明,相较于 YOLOv8n 模型,本文改进模型参数量下降 0.3M,计算量下降 0.2G,在自制数据集和公共数据集上,检测精度分别提升 0.8% 和 4.4%;在校内封闭路段,测距结果平均误差的绝对值为 104 mm,平均相对误差为 1.49%;在真实道路环境,平均误差的绝对值为 307.4 mm,平均相对误差为 3.29%,平均每帧耗时 30.12 ms,可满足实际测距场景下的需求。

关 键 词:车辆目标检测;YOLOv8;双目测距;嵌入式移植部署

中图分类号:TP391.4 **文献标识码:**A **doi:**10.37188/CJLCD.2026-0075 **CSTR:**32172.14.CJLCD.2026-0075

Vehicle measurement and recognition technology based on GGW-YOLOv8 in binocular vision

CHENG Yao, TANG Qingling*, LIU Yunyang, TANG Qingtao, LI Wei, LIU Peng

(College of Mechanical Engineering, Chongqing University of Technology, Chongqing 400054, China)

Abstract: Aiming at the problems of expensive and complex distance measurement sensors, low detection accuracy in complex road sections, and slow detection efficiency in traditional target detection and distance

收稿日期:2026-04-29;**修订日期:**2026-05-17.

基金项目:国家自然科学基金(No. 52375526);重庆市科委基础与前沿研究一般项目(No. cstc2016jcyjA0497);重庆市科委基础研究与前沿探索专项(No. CSTB2022NSCQ-MSX0950);重庆市研究生科研创新项目(No. CYS25735);重庆理工大学研究生创新项目(No. gzlxcx20253081, No. gzlxcx20253082, No. gzlxcx20252012) Supported by National Natural Science Foundation of China(No. 52375526); General Project of Basic and Frontier Research of Chongqing Science and Technology Commission(No. cstc2016jcyjA0497); Chongqing Science and Technology Commission Research and Frontier Exploration Project(No. CSTB2022NSCQ-MSX0950); Chongqing Graduate Research Innovation Project(No. CYS25735); Graduate Innovation Project of Chongqing University of Technology(No. gzlxcx20253081, No. gzlxcx20253082, No. gzlxcx20252012)

*通信联系人, E-mail: 631239340@qq.com

measurement technologies, a vehicle front distance measurement and target recognition technology based on GGW-YOLOv8 is proposed. This algorithm improves YOLOv8n by introducing GAM global attention mechanism before medium target and large target detection head, and improves the feature extraction ability of vehicle target by reducing information loss and amplifying global interaction. The GSConv lightweight convolutional module is introduced in the neck end network to reduce the number of model parameters and the calculation amount by deep separable convolution and shuffle operation. The WIoU_v3 loss function is used to replace the CIoU loss function, and the generalization ability and regression accuracy of the network boundary box are improved by the gradient gain distribution strategy of non-monotone dynamic focusing mechanism. The binocular vision ranging system based on SGBM semi-global stereo matching algorithm was built for vehicle ranging, and the upper computer system integrating vehicle target detection and ranging was developed, which is further transplanted and deployed to the embedded edge device for real vehicle on-board detection. The experimental results show that compared with the YOLOv8n model, the improved model has a 0.3M reduction in the number of parameters and 0.2G reduction in the amount of computation, and achieves accuracy improvements of 0.8% and 4.4% on the self-made dataset and the public dataset, respectively. In the closed section of the school, the absolute value of the average error is 104 mm, and the average relative error is 1.49%; in the actual road environment, the absolute value of the average error is 307.4 mm, the average relative error is 3.29%, and the average time per frame is 30.12 ms, which can meet the needs of actual ranging scenarios.

Key words: vehicle target detection; YOLOv8; binocular ranging; embedded migration deployment

1 引言

车辆检测与测距技术作为当下自动驾驶技术的研究重点,对减少交通事故发生率、提高道路交通安全具有重要的现实意义^[1]。传统目标检测算法的特征提取主要是指从采集的图像数据中人工提取目标车辆特征,如边缘检测、纹理特征等。常用的特征提取方法包括HOG特征、SIFT特征和SURF特征等。分类器则是依托提取到的特征对目标车辆进行分类和识别。传统的车辆目标检测算法虽然计算速度较快,但在更高精度和复杂场景下的稳定性和准确性较差。类似地,在双目视觉应用中,传统方法也面临弱纹理环境下特征提取困难、定位精度下降等问题^[2-5]。

近年来,随着机器视觉的快速发展,基于深度学习的目标检测算法被广泛应用于图像分类、目标检测等领域。与传统缺陷检测技术相比,基于深度学习的目标检测算法不仅检测精度高、计算速度快,模型的泛化能力也更强。基于深度学习的车辆目标检测算法包括以R-CNN^[6]、Fast R-CNN^[7]为代表的分类算法和以SSD^[8]、YOLO系列为代表的回归算法两大类。其中,计算速度更快的回归算

法显然更适用于实时性更高的车辆目标检测场景。典型的基于回归的目标检测算法有Redmon等人^[9]在2016年提出的YOLOv1算法,该算法将输入图像划分成网格送入检测网络,利用网络预测目标边界框,但其对小目标的物体检测效果较差。同年,Liu等人^[10]在YOLOv1的基础上提出了SSD算法,该算法引进金字塔模型,同样采用回归思想直接一步得到检测结果。后续研究学者在YOLOv1和SSD的基础上又相继提出YOLO9000^[11]、YOLOv3^[12]、YOLOX^[13]和DSSD^[14]等基于回归检测的改进算法,在实时性和检测精度上表现更佳。

YOLO算法已更新迭代到YOLOv3^[15-16]、YOLOv5、YOLOX、YOLOv7^[17]等多个版本。YOLOv8是由Ultralytics在2023年开发的最新版本,虽然在此前的基础上进行了主干网络的优化,但在颈端网络的卷积层略显冗余,且计算量太大、推理速度过慢。针对YOLOv8在特定场景下检测精度不足的问题,已有研究通过引入注意力机制、轻量化卷积模块等策略进行改进,取得了良好效果^[18]。

车辆检测与测距是自动驾驶环境感知的关

键技术,直接影响行车安全与辅助驾驶系统的可靠性。传统方法依赖人工设计特征,在复杂光照、遮挡及远距离场景下鲁棒性差、精度低。基于深度学习的YOLO系列算法凭借高效的检测框架,在实时性与精度之间取得较好平衡,但YOLOv8n仍存在颈端结构冗余、计算开销大、中远距离车辆易漏检、嵌入式边缘设备部署困难等问题。针对以上问题,本文提出一种基于GGW-YOLOv8的车辆目标检测算法。该算法通过注意力增强、结构轻量化与损失函数优化,提升复杂道路下车辆检测精度与推理效率;结合双目立体视觉构建测距系统,可实时检测前方车辆的位置、类别和距离信息。在此基础上,本文将整个系统移植部署到嵌入式开发平台Jetson Xavier NX上并进行实车搭载实验,实现了30 FPS级实时检测与测距,基本满足实际场景的应用,可为辅助驾驶前向防撞提供可靠技术支撑。

2 车辆测距及识别检测原理

为了实现对前方目标车辆的实时类别检测和测距,需要将改进的目标检测网络与双目立体测距进行融合。而目标检测与双目测距属于两种任务类型,既相互联系又表达不同的功能。基于YOLOv8n改进的目标检测网络可以准确获取目标车辆的2D检测框,并通过特征提取识别出车辆种类并给出置信度;基于SGBM半全局立体匹配的双目视觉测距模块则获取目标车辆的3D位置信息和视差信息。两者融合的目的是在目标车辆的检测框上同时输出类别、置信度及距离信息。改进的YOLOv8n网络已完成目标检测框输出类别以及置信度的任务,因此融合任务只需在2D检测框上叠加车辆的距离信息。具体流程如图1所示。

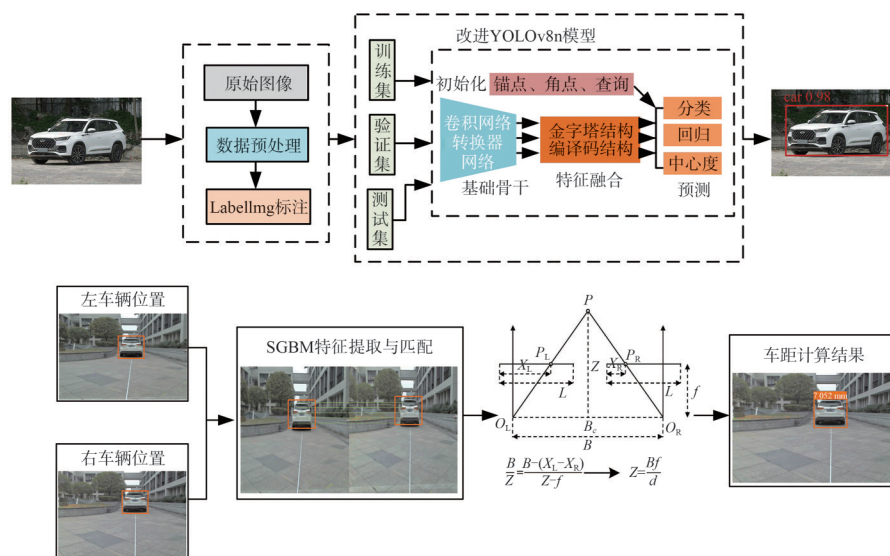


图1 基于双目视觉的车辆检测与测距系统框图

Fig. 1 Flow chart of vehicle detection and ranging system based on binocular vision

利用双目相机采集前方车辆图像,使用改进的YOLOv8n检测网络对左视图的目标车辆进行检测和识别,得出目标车辆的检测框位置、类别以及置信度信息。利用标定好的双目相机参数对左右视图进行立体校正,通过SGBM半全局的立体匹配算法得到左右视图的视差图,再利用目标框内每个像素点对应的视差信息测距并求平均值。最后,在输出图像上显示检测的车辆类别、置信度以及距离等信息。

3 基于GGW-YOLOv8的车辆检测与测距算法

3.1 车辆目标检测算法

针对目前辅助驾驶系统中车辆目标检测算法YOLOv8n模型参数量大、计算复杂度高、推理速度慢以及在复杂道路场景检测精度不高、容易漏检的问题,现对目标检测网络进行改进,本文改进后的YOLOv8网络结构如图2所示。

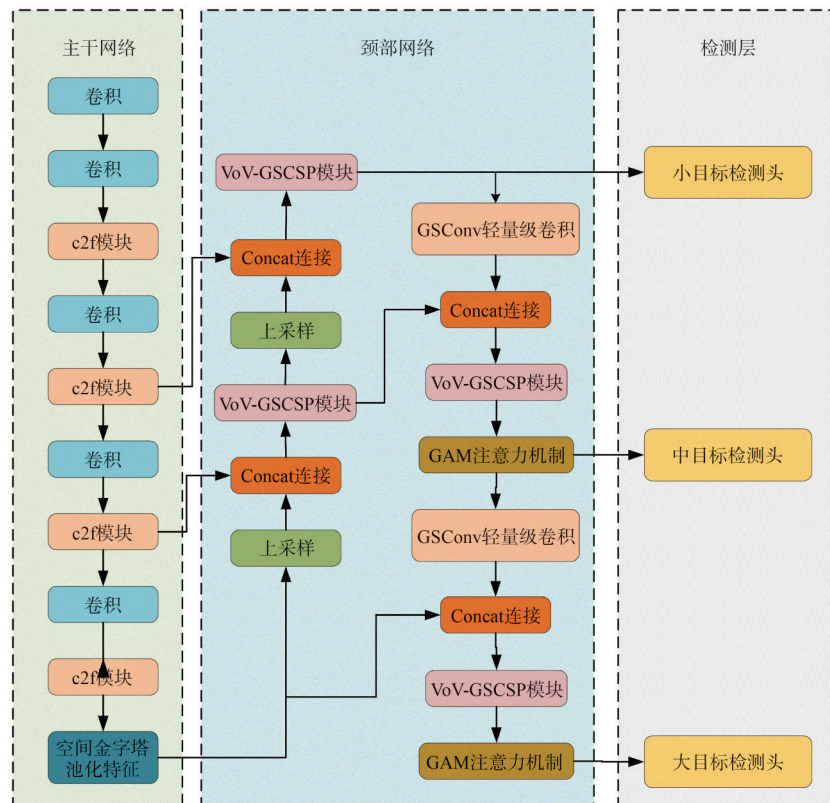


图2 改进的YOLOv8网络结构图

Fig. 2 Architecture diagram of improved YOLOv8 network

本文改进工作如下:(1)在颈部网络(Neck)对应的中目标检测头P4和大目标检测头P5前引入全局注意力机制(Global Attention Mechanism, GAM),增强模型复杂光照下对中大目标的特征提取能力;(2)在Neck网络使用GSConv轻量化卷积引入的VoV-GSCSP模块替换原有的4个C2f模块,在保证检测性能的同时降低颈端网络冗余的参数量和计算量,与加入全局注意力机制相互平衡,保障嵌入式实时性;(3)使用WIoU_v3损失函数替代原有回归损失函数CIoU,在保证高质量锚框检测效果的同时分配相当的梯度增益给低质量锚框以此减少有害梯度影响,从而有

效提高收敛速度和边界框的回归精度。

3.1.1 全局注意力机制

由于本文针对的目标检测对象是各种车辆,在数据集的图像中占据大部分像素,故为了提高模型对车辆等中大型目标的特征表达能力,本文在Neck端网络的中目标检测头和大目标检测头前引入全局注意力机制^[19]来提升网络模型的性能。常见的注意力机制模块有CBAM^[20]、SE^[21]等,GAM在此基础上不仅引入了3D排列和多层感知的通道注意力模块,同时还引入了卷积空间注意力模块,其结构如图3所示。

通道注意力子模块通过3D排列跨越3个维

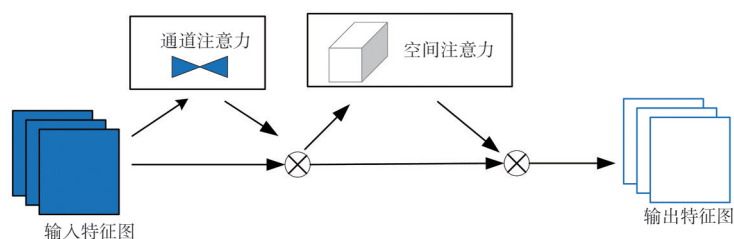


图3 全局注意力机制结构

Fig. 3 Structure of GAM

度来保留特征信息,并使用多层感知机 MLP 来放大跨维度的通道与空间的依赖关系^[22],其通道注意力子模块如图 4 所示。

空间注意力子模块使用两个卷积层进行空间信息融合,通过去池化操作减少信息对检测结果产生的负面影响,如图 5 所示。

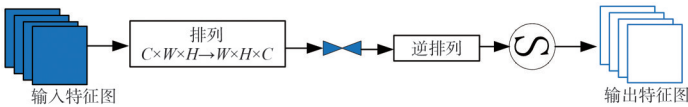


图 4 通道注意力子模块
Fig. 4 Channel attention submodule

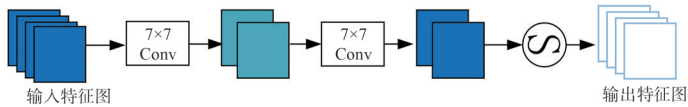


图 5 空间注意力子模块
Fig. 5 Spatial attention submodule

3.1.2 GSConv 轻量化卷积网络

YOLOv8 其三通道的卷积和全内容聚合导致模型的参数量过大,推理速度过慢。针对以上

问题,本文引入 Li Hulin 提出的基于 Slim-Neck 设计的 GSConv 轻量化卷积模块^[23],其结构如图 6 所示。

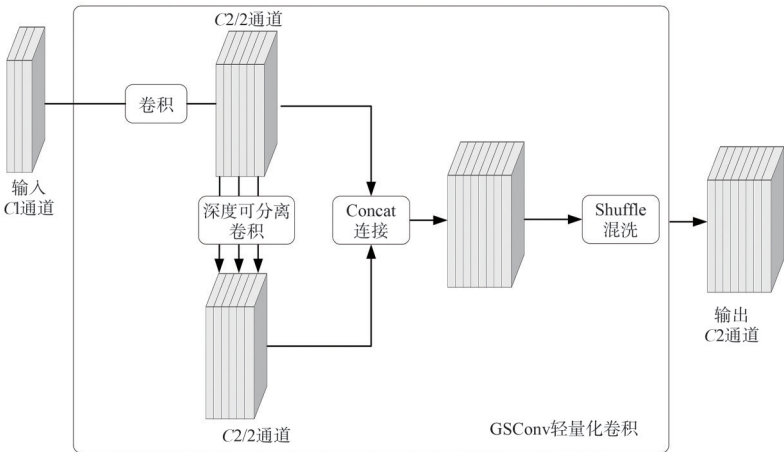


图 6 GSConv 轻量卷积结构图
Fig. 6 Structure diagram of GSConv lightweight convolution

前文为加强模型对车辆等中大型目标的特征表达能力,引入了全局注意力机制,但该模块会导致参数量和计算量有所上升。考虑到所设计的网络模型需移植到计算能力更弱的嵌入式移动设备,本文设计 VoV-GSCSP 复合模块代替原 C2f 模块^[24]。该模块采用一次性聚合方法,核心是对输入特征的多尺度信息、跨通道特征进行单次并行融合,具体实现为:将输入特征拆分为直接传递支路与 GSConv 处理支路,通过 GSConv 完成多尺度特征提取后,经特征混洗操

作增强跨通道关联,最终通过单次 Concat 操作完成特征整合。该方法可进一步降低模型参数量和计算量,其网络模块结构如图 7 所示。

3.1.3 WIoU_v3 损失函数

YOLOv8 使用的回归损失函数为 CIOU loss + DFL, CIOU 没有考虑样本难易的平衡问题,忽略了样本数据集中高质量样本和低质量样本的平衡,损失函数中几何因素的惩罚机制会加重对低质量样本的标注,导致最后低质量样本的检测效果不佳。为此,本文引入 2023 年 Tong 等人^[25]提出的

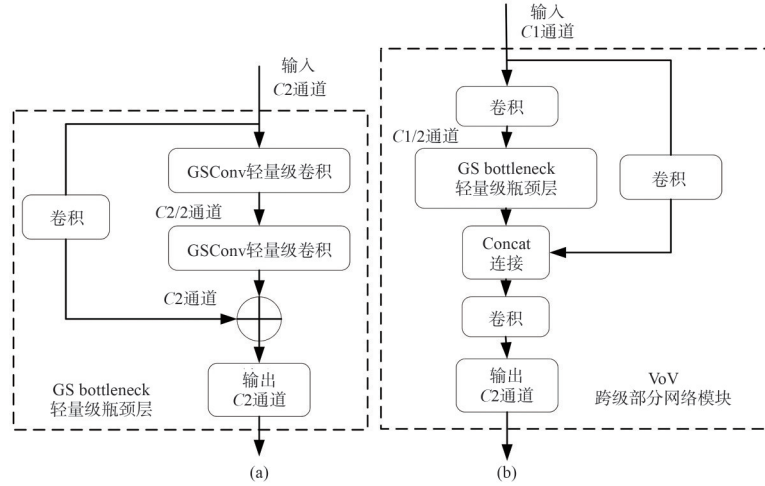


图7 GSConv轻量卷积(a)与VoV跨级网络模块(b)

Fig. 7 GSConv lightweight convolution (a) and VoV cross-stage network module (b)

最新 WIoU (Wise-IoU) 损失函数。该损失函数按结构分为 v1、v2 和 v3 三种。WIoU_v3 构建了一种针对交叉熵的非单调动态聚焦机制。该机制利用梯度增益的分配策略在不同阶段做出最适合当前增益的选择,使模型的损失函数 WIoU_v3 不仅聚焦于高质量图像的锚框,同时也聚焦于普通质量以及低质量的锚框,从而提高检测器的性能,增强模型的泛化能力。WIoU_v3 的定义如式(1)~(3)所示:

$$L_{\text{WIoU}_v3} = \gamma \cdot R_{\text{WIoU}} \cdot L_{\text{IoU}}, \quad (1)$$

$$R_{\text{WIoU}} = \exp \frac{(x - x_{\text{gt}})^2 + (y - y_{\text{gt}})^2}{(W_g^2 + H_g^2)}, \quad (2)$$

$$\gamma = \frac{\beta}{\delta \alpha^{\beta - \delta}}. \quad (3)$$

式中: R_{WIoU} 表示真实框和预测框中心点的归一化距离, L_{IoU} 表示边界框 IoU 的损失值, W_g 表示最小预测框的宽, H_g 表示最小预测框的高, α 表示常数, β 表示离群度, δ 为可调节的超参数, 非单调聚焦系数 γ 随着训练轮数的变化而变化。WIoU_v3 利用非单调动态聚焦机制的梯度分配策略在早期能很好地将梯度增益分配给高质量锚框, 在中后期则将梯度增益分配给低质量锚框, 以此减少低质量图像的有害梯度。因此 WIoU_v3 的分配策略是一个动态的过程, 能在训练过程中的每一时刻做出符合当前梯度增益分配的最佳策略。该损失函数不仅能使模型获得更好的泛化能力, 同时也提升了网络边界框的回归精度。

3.2 车辆前车测距算法

为进一步验证本文车辆测距系统设计的合理性, 搭建了基于双目视觉测距实验平台对前方车辆进行静态测距实验验证, 并将搭建的实验平台移植到真实道路行驶的车辆中对前方动态行驶的车辆进行动态测距实验验证。

本文选用 SGBM 半全局立体匹配算法^[26]对双目视图的匹配图像进行视差信息获取, 该算法包括预处理、代价计算、代价聚合和后处理 4 个步骤。SGBM 算法的总体流程如图 8 所示。

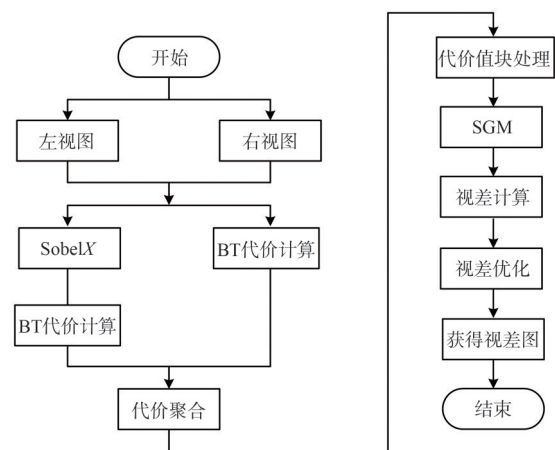


图8 SGBM算法流程图

Fig. 8 Flowchart of SGBM algorithm

4 模型验证与结果分析

实验环境配置如下: 操作系统为 Windows10

64 位、CPU 为 Intel i5 13600F、GPU 为 NVIDIA GeForce RTX4070、内存为 32G、开发环境为 Pytorch2.01、Python3.9、cuda11.8,通过配置的软件环境对改进模型进行训练和测试。

训练阶段超参数配置:输入图像大小为 640×640 ;公共数据集 BDD100K 的迭代轮数为 300 轮;自制数据集的迭代轮数为 200 轮;批量通道数大小均为 32;初始学习率为 0.01,最终学习率为 0.2,动量因子为 0.937,权重衰减系数为 0.0005;为防止过拟合,优化器选择 SGD 优化器。

4.1 实验数据集

本文使用的数据集为两类,一类为 BDD100K 公共数据集。BDD100K 数据集是由加州大学伯克利分校发布的用于自动驾驶领域研究的大规模、多样化的数据集^[27]。本文从该数据集 70 000 张原始图像中选择了车辆样本数量至少为 2 的

10 720 张图片进行训练,同时剔除车辆目标信息以外的其他标签信息。数据集比例按照 8:1:1 分别划分训练集、测试集、验证集,其中训练集 8 576 张,验证集 1 072 张,测试集 1 072 张。数据集的种类包含 Car、Bus、Truck、Motorcycle 四类。

另一类为自制数据集。该数据集是由相机在车辆行驶过程中拍摄的真实道路视频,对原视频中的道路车辆等样本进行抽帧取样,共得到 8 216 张图片。数据集同样按照 8:1:1 的比例划分成训练集、测试集、验证集,其中训练集 6 972 张,验证集 622 张,测试集 622 张。数据集的种类同样包含 Car、Bus、Truck、Motorcycle 四类。Car 主要包括小轿车、面包车、小型越野车等;Bus 主要包括公交车、中巴车、大巴车等;Truck 主要包括卡车、货运车等;Motorcycle 主要包括摩托车和电动车等。各类车辆样本示意图如图 9 所示。



图9 数据集各类车辆样本示意图
Fig. 9 Schematic diagram of various vehicle samples in the dataset

4.2 基于公共数据集的实验结果分析

本文采用精确率(Precision, P)、召回率(Recall, R)、平均精度(AP)、mAP 指标评价目标检测算法性能。这 4 类评价指标均由 TP、TN、FP、FN 计算。TP 表示正确识别的车辆目标;

TN 表示正确检测出非车辆的其他样本情况;FP 表示实际不是车辆检测结果却是车辆,属于误检的情况;FN 表示实际是车辆但检测结果不是车辆,属于漏检的情况。

由图 10 可以看出,随着训练轮数的迭代进

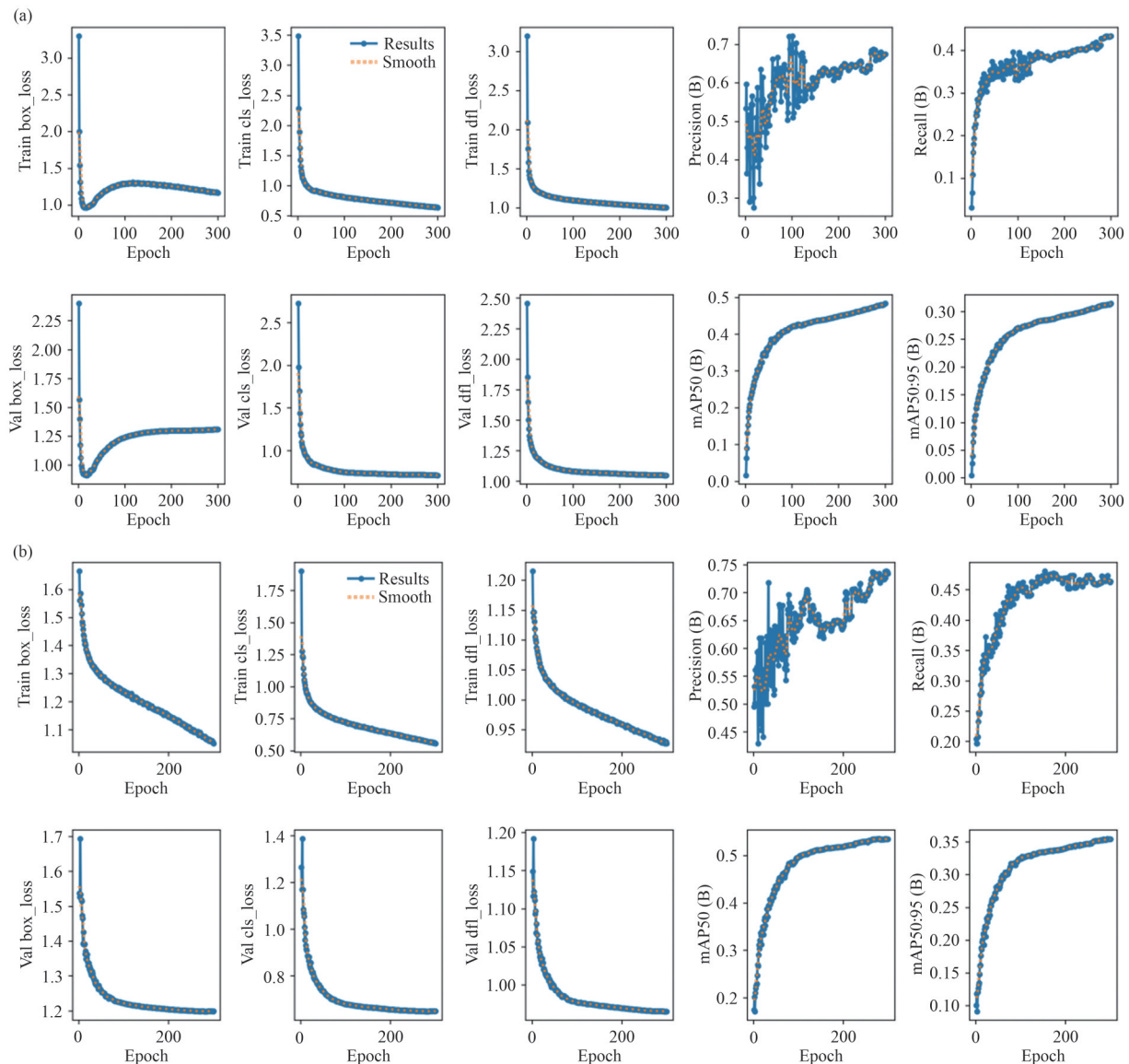


图10 模型训练对比图1。(a) YOLOv8n; (b)改进 YOLOv8n。

Fig. 10 Comparison diagram of model training 1. (a) YOLOv8n; (b) Improved YOLOv8n.

行,改进网络模型的训练集和验证集的损失函数曲线均呈现下降趋势,并在迭代轮数为300轮时趋于稳定。未改进的YOLOv8n模型在训练集和验证集定位损失函数处出现了先下降后上升的情况,分析原因可能是基线模型复杂度过高,导致模型在训练初期损失函数下降较快。但随着训练的进行,模型开始出现过拟合现象,损失函数上升。改进损失函数后的YOLOv8n模型在损失函数下降阶段则避免了这一问题,且下降曲线趋于平滑。从训练对比图还可以看出改进后的YOLOv8n模型在精确率和召回率上均有所提升,改进的YOLOv8n模型 $mAP@0.5$ 为53.6,

$mAP@0.5:0.95$ 为35.5,相较于未改进YOLOv8n在 $mAP@0.5$ 上提升4.4个百分点,在 $mAP@0.5:0.95$ 上提升2.8个百分点。各样本车辆模型精度对比如表1所示。

图11为两个模型在公共数据集BDD100K测试集上的检测结果对比。基线模型YOLOv8n可以在复杂环境下对前方的车辆实施检测并能识别出不同的类别,但对远距离车辆则出现漏检情况。改进后的YOLOv8n相较于基线模型YOLOv8n,在各类车辆样本的检测精度均有所提升,且能有效检测出更远距离车辆,总体检测效果有所改进。

表 1 BDD100K 公共数据集上各类别精度对比
Tab. 1 Precision comparison on BDD100K public dataset

模型	Precision/%				mAP@0.5/%	mAP@0.5:0.95/%
	Car	Bus	Motorcycle	Truck		
YOLOv8n	69.0	53.8	24.7	49.2	49.2	32.7
改进 YOLOv8n	70.6	56.4	33.7	53.8	53.6	35.5



图 11 模型检测对比图 1
Fig. 11 Comparison of model detection 1

4.3 基于自制数据集的实验结果分析

由模型训练对比图 12 可以看出,随着迭代的进行,训练集和验证集的损失函数曲线均呈现下降趋势,并在迭代轮数为 200 轮时趋于稳定。未改进的 YOLOv8n 模型与公共数据集上的损失函数一致,在训练集和验证集定位损失函数处出现先下降后上升的情况,分析原因可能是自制数据集由于抽帧相隔帧数较小出现大量特征相似的车辆数据图像,导致模型在训练初期损失函数下降较快。但随着训练的进行,模型开始出现过拟合现象,损失函数有所上升。为解决上述问题,改进损失函数后的 YOLOv8n 模型在损失函数下降阶段将梯度增益有效分配给训练过程的每个图像,避免了损失函数先下降后上升的过拟合现象,故下降曲线趋于平滑。从模型对比图还可看出,改进后的 YOLOv8n 模型精确率和召回率均有所提升,改进后的 YOLOv8n 的 mAP@0.5 为 84.5, mAP@0.5:0.95 为 66.7,相较于未改进 YOLO8n 在 mAP@0.5

上提升了 0.8 个百分点,在 mAP@0.5:0.95 上提升了 1.1 个百分点。模型训练的 *P-R* 曲线如图 13 所示,各样本车辆模型精度对比如表 2 所示。

由表 2 可知,改进的 YOLOv8n 模型相比于基线模型在 Bus、Motorcycles、Truck 三个类别的精度、mAP@0.5 和 mAP@0.5:0.95 均有提升。其中,Car 类别的精度相较于基线模型没有提升,主要原因是在基线模型训练过程中,检测精度已达到实际应用需求,改进后的模型旨在增强中大型车辆的特征提取能力,使中大型车辆的检测精度得到提升。图 14 为两个模型在自制数据集的测试集上的检测结果对比。

如图 14 所示,基线模型 YOLOv8n 和改进 YOLOv8n 模型在自制数据集上均有较好的检测效果,但改进后的 YOLOv8n 由于引入了全局注意力机制,对 Bus、Truck 类中大型目标车辆的检测精度都有所提升,对更远距离的车辆目标也具备更好的检测能力。

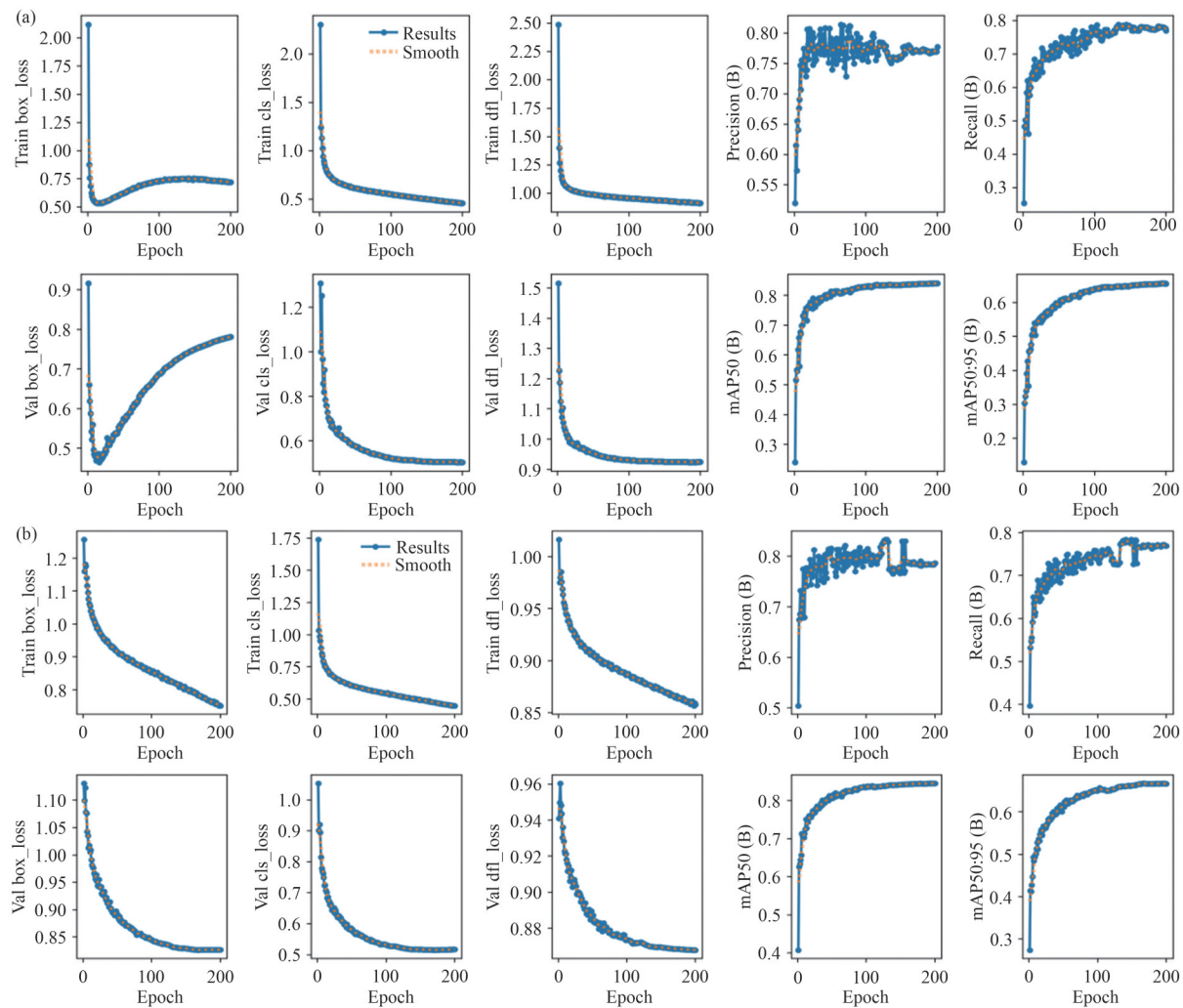


图12 模型训练对比图2。(a)YOLOv8n;(b)改进YOLOv8n。

Fig. 12 Comparison of model training 2. (a)YOLOv8n; (b) Improved YOLOv8n.

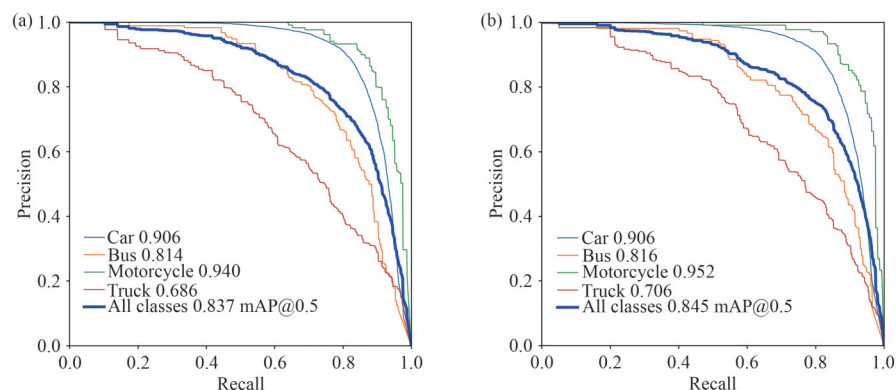


图13 自制数据集的训练P-R曲线。(a) YOLOv8n;(b)改进YOLOv8n。

Fig. 13 P-R curves of self-made dataset training. (a) YOLOv8n; (b) Improved YOLOv8n.

4.4 整体改进效果分析

在相同的软硬件配置和超参数设置环境下,分别与不同的改进子模块进行组合,通过消融实验

观察各个子模块的改进方式对基线网络模型的影响。以自制数据集为例,消融实验结果如表3所示。

表 2 模型在自制数据集上各类别精度对比

Tab. 2 Precision comparison of the models on self-made dataset

模型	Precision/ %				mAP@0. 5/ %	mAP@0. 5:0. 95/ %
	Car	Bus	Motorcycle	Truck		
YOLOv8n	90. 6	81. 4	94. 0	68. 6	83. 7	65. 6
改进 YOLOv8n	90. 6	81. 6	95. 2	70. 6	84. 5	66. 7



图 14 模型检测对比图 2

Fig. 14 Comparison of model detection 2

表 3 模型在自制数据集上的消融实验

Tab. 3 Ablation experiments of the models on self-made dataset

算法模型	mAP@0. 5/ %	mAP@0. 5:0. 95/ %	Parameters/M	FLOPs/G
YOLOv8n(Baseline)	83. 7	65. 6	3. 2	8. 1
YOLOv8n-Slim	84. 0	66. 1	2. 9	7. 7
YOLOv8n-GAM	83. 8	65. 8	3. 3	8. 2
YOLOv8n-WIoU	83. 8	65. 7	3. 1	8. 1
YOLOv8n-Slim-GAM	84. 3	66. 4	3. 1	8. 0
YOLOv8n-Slim-WIoU	84. 1	66. 2	3. 1	8. 0
YOLOv8n-Slim-GAM-WIoU	84. 5	66. 7	3. 0	7. 9

注:粗体数字表示最优值

实验结果表明,改进模型将原模型 Neck 层中传统卷积模块用 GSConv 卷积模块替换后,参数量和计算量都有所下降,且 mAP@0. 5 和 mAP@0. 5:0. 95 分别有 0. 3% 和 0. 5% 的提升;在注意力机制方面,由于在检测层的中目标检测头和大目标检测头的前面分别引入全局注意力机制,使模型在精度方面有所提升,但在参数量和计算量有所增加;在损失函数方面,用 WIoU_v3 替换原模型的 CIoU 损失函数,在精度上有所提升,但不会改变模型的参数量和计算量;将 Slim 网络分别与 GAM 注意力机制和损失函数 WIoU_v3 组合,模型在精度、参数量以及计算量上都有一定的提

升,证明基于 Slim-Neck 网络的设计范式与该注意力机制和损失函数组合可较好提升模型性能;改进后的 YOLOv8n 网络结合了 Slim-Neck 的轻量化卷积、GAM 注意力机制和 WIoU_v3 损失函数,其精度 $mAP@0.5$ 和 $mAP@0.5:0.95$ 分别提升 0.8% 和 1.1%,参数量下降 0.3M,浮点运算次数下降 0.2G,整体性能较基线网络 YOLOv8n 有

所提升,验证了各改进策略的有效性和互补性。

为验证改进模型性能提升的有效性和泛化性,在保持实验环境基本不变的情况下,使用公共数据集 BDD100K 和自制数据集将改进的网络模型与目前一些主流的目标检测算法,包括 YOLO 系列、SSD、RTDETR 等进行性能对比,实验结果如表 4 所示。

表 4 改进算法与其他目标检测算法的实验结果对比

Tab. 4 Comparison of experimental results between the improved algorithm and other target detection algorithms

算法模型	BDD100K					自制数据集				
	mAP50	mAP50:95	Parameters/ FLOPs/		Epochs	mAP50	mAP50:95	Parameters/ FLOPs/		Epochs
			M	G				M	G	
YOLOv5s	54.1	36.9	7.0	15.8	300	84.9	67.5	6.9	15.7	200
SSD(VGG)	52.8	34.2	25.9	61.8	300	83.6	62.9	25.4	61.3	200
YOLOv7-tiny	51.5	30.7	5.8	13.4	300	82.2	61.4	5.4	12.9	200
YOLOv9c	55.5	37.7	25.3	102.3	300	85.2	66.9	25.3	102.3	200
RTDETR-L	36.4	21.7	32.0	103.7	300	68.3	50.1	32.0	103.4	200
YOLOv8n-MobileNet	49.7	28.9	2.6	5.9	300	78.9	52.9	2.4	5.6	200
YOLOv8n(基线)	49.2	32.7	3.3	8.2	300	83.7	65.6	3.2	8.1	200
Ours(改进)	53.6	35.5	3.1	8.0	300	84.5	66.7	3.0	7.9	200

注:粗体数字表示最优值

实验结果表明,本文提出的改进模型与其他两阶段的算法相比,在公共数据集和自制数据集上的各评价指标都有较好提升,实现了检测精度与轻量化部署能力的良好平衡,综合性能优于同量级轻量化模型,大幅压缩了参数量与计算量,完全适配车载嵌入式平台的实时性需求。

以 BDD100K 公共数据集为例,本改进模型的 $mAP@0.5$ 达 53.6%、 $mAP@0.5:0.95$ 达 35.5%,检测精度虽不及 YOLOv5s 网络模型,但 YOLOv5s 作为 s 量级的网络模型相比于本改进的 n 小量级模型,其参数量和计算量成倍增加,推理速度慢于本改进模型,综合性能存在一定差距。SSD 作为经典的两阶段算法,其检测精度 $mAP@0.5$ 和 $mAP@0.5:0.95$ 较改进后的模型分别低 0.8% 和 1.3%,虽然相差较小,但改进模型的参数量和计算量降低超 87%,整体性能更优。与 YOLOv9c、RTDETR-L 等大模型相比,本文模型在参数量、计算量显著降低的基础上, $mAP@0.5$ 分别高出 1.9%、17.2%, $mAP@0.5:0.95$ 分别高出 2.2%、13.8%,综合性能优势显著。

同时,本改进模型相较于轻型网络 YOLOv7-

tiny 在 $mAP@0.5$ 上高出 2.1%, $mAP@0.5:0.95$ 高出 4.8%,但 YOLOv7-tiny 的参数量和计算量均高于本模型,本文方法在更轻量化的基础上实现了检测精度的全面提升。YOLOv8n-MobileNet 作为 YOLOv8n 替换主干网络的轻量化模型,虽然在参数量和计算量上分别下降 16.1% 和 26.3%,但 $mAP@0.5$ 和 $mAP@0.5:0.95$ 分别下降 7.3% 和 15.5%。本文模型仅通过小幅增加参数量与计算量,即可弥补轻量化带来的精度损失,综合性能更适配车载复杂场景的检测需求;相比于基线模型 YOLOv8n,本改进模型在公共数据集和自制数据集上 $mAP@0.5$ 和 $mAP@0.5:0.95$ 均得到全面提升,同时进一步降低了模型复杂度:两项指标在公共数据集上分别提升 4.4% 和 2.8%;在自制数据集上分别提升 0.8% 和 1.1%。参数量和计算量也有所下降:在公共数据集上分别下降 6.5% 和 2.5%,在自制数据集上分别下降 6.7% 和 2.5%。在不增加部署成本的前提下,本改进模型有效提升了车辆检测能力,完全满足车载嵌入式平台的实时部署要求。综上所述,本改进模型在保证轻量化的同时能较

好平衡检测精度,相比于其他目标检测算法性能有所提升。在车载嵌入式部署场景中,高参数量和高计算量会导致模型推理速度下降,难以满足 30 FPS 以上的实时检测需求。相比之下,本文改进的 GGW-YOLOv8 在精度损失不足 1% 的前提下,大幅压缩了模型参数量和计算量,更符合车载辅助驾驶系统对实时性与轻量化的硬性要求,实现了精度与部署成本的最优平衡。

4.5 测距实验与结果分析

本文搭建的基于双目视觉的车辆测距平台主要由双目相机、支架、目标车辆、上位机、米尺等构成。

为了进一步验证本测距系统在真实道路场景中的测距精度以及实时性,本文将搭建的实验平台移植到真实道路行驶的车辆中并对前方动态行驶的车辆进行测距实验;模拟双动态行驶过程中前车距离的实验测试。实验中,以激光测距仪测得数值为参考的实际距离,双目相机测得的数值为实验距离。激光测距仪的测量范围为 0.02~120 m,测量精度为±2 mm。为防止前挡风玻璃反射对激光测距仪造成测量误差,将激光测距仪固定在与双目相机垂直、激光发射头与双目相机水平的车顶位置。车辆以 30 km/h 的速度在真实道路上行驶,双目相机采集的视频为 30 帧/s,本文选取 60 帧作为车辆行驶序列,以视频中灰色前车作为参考对象,每隔 6 帧测量一次前车距离,部分测距结果如图 15 所示,测距各项数据如表 5 所示。



图 15 部分动态车辆测距结果
Fig. 15 Partial dynamic vehicle ranging results

从表 5 可以看出,车辆在动态行驶过程中与前车距离保持在 8 500~10 500 mm 之间,随着帧数的增加,两车距离逐渐减小。在 8 500~10 500 mm 区间两车测距的相对误差在 2%~4%,测距的平误差绝对值为 307.4 mm,平均相对误差为 3.29%,与之前静态测距推断的误差相符。整个测距过程相对稳定,未出现较大波动。动态测距平均每帧耗时 30.12 ms,相当于可处理

33 帧/s 图像,满足本次采集视频帧数的处理速度。综上所述,本测距系统在一定程度上可较好地实现车辆测距的实时性和准确性。

表 5 前方车辆动态测距数据

Tab. 5 Dynamic ranging data of the vehicle in front

帧数	实际 距离/mm	测量 距离/mm	绝对 误差/mm	相对 误差/%	耗时/ ms
6	10 386	9 987	399	3.84	29.52
12	10 201	9 851	350	3.43	30.28
18	10 019	9 652	367	3.66	28.95
24	9 324	9 644	−320	3.43	29.30
30	8 921	9 235	−314	3.52	31.28
36	8 902	8 638	264	2.97	30.65
42	8 946	8 660	286	3.16	29.88
48	8 866	8 590	276	3.18	29.12
54	8 746	8 989	−243	2.78	30.54
60	8 694	8 439	255	2.95	31.63

5 基于嵌入式设备系统移植部署实验

为了验证嵌入式设备部署后的运行效果,本文进行实车搭载实验,实验系统平台组成如图 16 所示,由 NVIDIA Jetson Xavier NX 开发板、显示器、双目相机、电源等构成。实验中,移植的算法为改进的 YOLOv8n 车辆目标检测算法和基于双目视觉 SGBM 立体校正匹配的车辆测距算法。实车搭载测试的 1 号道路为学校周边区域,2 号道路为城市高峰期行驶路段,两个路段均包含本次测量的车辆样本种类及不同道路场景,最终运行效果如图 17 所示。



图 16 实验系统平台
Fig. 16 Experimental system platform

在实车搭载实验中,双目相机作为数据输入源,输入图像分辨率为 2 560×960,经过缩放处理将图像转为 480×480,在该分辨率下成功运行

了改进的车辆目标检测和测距算法。由图17的实验结果可知,经实车搭载采集的视频流统计,本实验测量的4个车辆样本均可被准确检测出,检测精度达85%以上,采集路段涉及阴雨天和高峰拥堵遮挡等复杂场景,在复杂场景下的实车检测精度也达到80%以上;在车辆测距方面,40 m以内的车辆测距误差在6%以内,超过40 m的测距误差在6%以上。最终的运行效果能有效检测出车辆的类别和距离,视频检测帧率为30 FPS,证明整套目标检测测距系统在嵌入式设备上运行良好,在实际应用中基本满足目标检测与测距的实时性和准确性。



图17 不同道路场景下的实验结果。(a) 1号场景道路;
(b) 2号场景道路。

Fig. 17 Experimental results under different road scenarios.
(a) Road in Scene 1; (b) Road in Scene 2.

6 结 论

本文针对传统车辆目标检测算法在更高精度和复杂场景下的稳定性和准确性较差,及YOLOv8在颈端网络的卷积层略显冗余,且计算量较大、推理速度过慢等问题,提出一种基于GGW-YOLOv8的车辆目标检测算法。通过引入全局注意力、GSConv轻量化卷积与WIoU_v3损失函数,在有效降低模型复杂度的同时减少模型参数量和计算量,增强特征提取能力,提升边界框回归精度。搭建基于SGBM半全局立体匹配算法的双目视觉测距系统,开发融合车辆目标检测与测距的上位机系统,并移植部署到嵌入式边缘设备进行实车搭载检测。实验结果表明,改进后的模型在保持轻量化优势的同时,显著提升了车辆检测精度,测距误差满足实际行车需求,可在嵌入式边缘设备上实现30 FPS级实时运行。整套系统可应用于实车环境并能满足基本需求。

本文方法重点解决了车载场景光照多变、遮挡干扰、嵌入式实时性差等工程难题,可为低成本双目视觉辅助驾驶系统提供可行技术方案,但在极端天气、强逆光及密集遮挡场景下的鲁棒性仍需进一步提升。后续将结合多模态信息融合、动态化视差优化与模型蒸馏技术,进一步提高系统在复杂环境下的稳定性与适应性,为高级辅助驾驶系统提供更可靠的感知支撑。

参考文献:

- [1] 程瑶,赵雷,成珊,等. 基于机器视觉的车距检测系统设计[J]. 计量学报, 2020, 41(1): 11-15.
CHENG Y, ZHAO L, CHENG S, *et al.* Design of vehicle distance detection system based on machine vision [J]. *Acta Metrologica Sinica*, 2020, 41(1): 11-15. (in Chinese)
- [2] 龚坤,徐鑫,陈小庆,等. 弱纹理环境下融合点线特征的双目视觉同步定位与建图[J]. 光学精密工程, 2024, 32(5): 752-763.
GONG K, XU X, CHEN X Q, *et al.* Binocular vision SLAM with fused point and line features in weak texture environment [J]. *Optics and Precision Engineering*, 2024, 32(5): 752-763. (in Chinese)
- [3] 杨玉,许四祥,张梦权,等. 基于多尺度注意力机制TransUNet的双目视觉定位与测量方法[J]. 光学精密工程, 2025, 33(16): 2502-2515.
YANG Y, XU S X, ZHANG M Q, *et al.* Binocular vision localization and measurement method based on TransUNet with multi-scale attention mechanism [J]. *Optics and Precision Engineering*, 2025, 33(16): 2502-2515. (in Chinese)
- [4] 刘宇豪,吴福培,吴树壮,等. 基于插值超分辨率的双目三维重建方法[J]. 中国光学(中英文), 2024, 17(4): 862-874.
LIU Y H, WU F P, WU S Z, *et al.* Binocular 3D reconstruction method based on interpolation super-resolution [J].

- Chinese Optics*, 2024, 17(4): 862-874. (in Chinese)
- [5] 李鹏飞,熊召新,王桂宝. YOLOv8-NTS:一种针对交通标志检测的目标识别方法[J]. 液晶与显示,2025,40(12): 1868-1880.
LI P F, XIONG Z X, WANG G B. YOLOv8-NTS: a target detection approach for traffic sign recognition [J]. *Chinese Journal of Liquid Crystals and Displays*, 2025, 40(12): 1868-1880. (in Chinese)
- [6] GIRSHICK R, DONAHUE J, DARRELL T, *et al.* Rich feature hierarchies for accurate object detection and semantic segmentation [C]//*Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition*. Columbus: IEEE, 2014: 580-587.
- [7] CHEN Y H, WANG H R, LI W, *et al.* Scale-aware domain adaptive faster R-CNN [J]. *International Journal of Computer Vision*, 2021, 129(7): 2223-2243.
- [8] JIN L, LIU G D. An approach on image processing of deep learning based on improved SSD [J]. *Symmetry*, 2021, 13(3): 495.
- [9] REDMON J, DIVVALA S, GIRSHICK R, *et al.* You only look once: unified, real-time object detection [C]. *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition*. Las Vegas: IEEE, 2016: 779-788.
- [10] LIU W, ANGUELOV D, ERHAN D, *et al.* SSD: single shot MultiBox detector [C]. 14th European Conference on Computer Vision-ECCV 2016. Amsterdam: Springer, 2016: 21-37.
- [11] REDMON J, FARHADI A. YOLO9000: better, faster, stronger [C]// *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition*. Honolulu: IEEE, 2017: 6517-6525.
- [12] LAWAL M O. Tomato detection based on modified YOLOv3 framework [J]. *Scientific Reports*, 2021, 11(1): 1447.
- [13] HAN L Y, LI F G, YU H Y, *et al.* BiRPN-YOLOvX: a weighted bidirectional recursive feature pyramid algorithm for lung nodule detection [J]. *Journal of X-Ray Science and Technology*, 2023, 31(2): 301-317.
- [14] ARAKI R, HIRAKAWA T, YAMASHITA T, *et al.* MT-DSSD: multi-task deconvolutional single shot detector for object detection, segmentation, and grasping detection [J]. *Advanced Robotics*, 2022, 36(8): 373-387.
- [15] 叶赵兵,段先华,赵楚. 改进YOLOv3-SPP水下目标检测研究[J]. 计算机工程与应用,2023,59(6):231-240.
YE Z B, DUAN X H, ZHAO C. Research on underwater target detection by improved YOLOv3-SPP [J]. *Computer Engineering and Applications*, 2023, 59(6): 231-240. (in Chinese)
- [16] 王浩,雷印杰,陈浩楠. 改进YOLOV3实时交通标志检测算法[J]. 计算机工程与应用,2022,58(8):243-248.
WANG H, LEI Y J, CHEN H N. Real time traffic sign detection algorithm based on improved YOLOV3 [J]. *Computer Engineering and Applications*, 2022, 58(8): 243-248. (in Chinese)
- [17] 盖荣丽,孔祥宙,泰山,等. 基于YOLOv7改进的夜间樱桃检测方法:YOLOv7-Cherry[J]. 计算机工程与应用, 2024,60(21):315-323.
GAI R L, KONG X Z, QIN S, *et al.* Improved cherry detection method at night based on YOLOv7: YOLOv7-Cherry [J]. *Computer Engineering and Applications*, 2024, 60(21): 315-323. (in Chinese)
- [18] 朱世松,高虹,芦碧波,等. 基于改进YOLOv8n的城市道路病害检测算法[J]. 液晶与显示,2025,40(9):1356-1368.
ZHU S S, GAO H, LU B B, *et al.* Urban road defect detection algorithm based on improved YOLOv8n [J]. *Chinese Journal of Liquid Crystals and Displays*, 2025, 40(9): 1356-1368. (in Chinese)
- [19] ZHOU K, TONG Y F, LI X T, *et al.* Exploring global attention mechanism on fault detection and diagnosis for complex engineering processes [J]. *Process Safety and Environmental Protection*, 2023, 170: 660-669.
- [20] WOO S, PARK J, LEE J Y, *et al.* CBAM: convolutional block attention module [C]//*Proceedings of the 15th European Conference on Computer Vision*. Munich: Springer, 2018: 3-19.
- [21] HU J, SHEN L, SUN G. Squeeze-and-excitation networks [C]//*Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Salt Lake City: IEEE, 2018: 7132-7141.
- [22] CAO W H, FENG Z Y, ZHANG D Y, *et al.* Facial expression recognition via a CBAM embedded network [J]. *Procedia Computer Science*, 2020, 174: 463-477
- [23] LI H, LI J, WEI H, *et al.* Slim-neck by GSConv: a better design paradigm of detector architectures for

- autonomous vehicles [J]. *Journal of Real-Time Image Processing*, 2024, 21(3): 62.
- [24] 胡峻峰,李柏聪,朱昊,等. 改进YOLOv8的轻量化无人机目标检测算法[J]. 计算机工程与应用, 2024, 60(8): 182-191.
- HU J F, LI B C, ZHU H, *et al.* Improved YOLOv8 lightweight UAV target detection algorithm [J]. *Computer Engineering and Applications*, 2024, 60(8): 182-191. (in Chinese)
- [25] TONG Z J, CHEN Y H, XU Z W, *et al.* Wise-IoU: bounding box regression loss with dynamic focusing mechanism [J/OL]. *arXiv*, 2023: 2301.10051.
- [26] 刘阳阳. 无人机双目视觉目标检测与测距方法研究[D]. 重庆:重庆大学, 2019:27-29.
- LIU Y Y. Research on binocular vision object detection and ranging for UAV [D]. Chongqing: Chongqing University, 2019: 27-29. (in Chinese)
- [27] YU F, CHEN H F, WANG X, *et al.* BDD100K: a diverse driving dataset for heterogeneous multitask learning [C]// *Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Seattle: IEEE, 2020: 2633-2642.

作者简介:



程 瑶,女,硕士,副教授,2006年于重庆大学获得硕士学位,主要从事红外探测及传感、图像成像系统、图像检测及处理算法等方面的研究。E-mail: chy_cqit@cqut.edu.cn



唐青玲,女,硕士研究生,2024年于绵阳师范学院获得学士学位,主要研究方向为光电检测及机器视觉。E-mail: 631239340@qq.com