

改进的 YOLOv11 智能车辆动态环境目标检测算法^{*}于承峰^{1,2} 高 松^{1,2} 王鹏伟^{1,2} 孙宾宾^{1,2} 张 榕^{1,2}

(1. 山东理工大学交通与车辆工程学院 淄博 255000; 2. 山东省新能源汽车集成设计与智能化重点实验室 淄博 255000)

摘 要: 针对智能车辆复杂动态场景下车辆小目标和遮挡目标存在的检测困难及定位不准问题,提出了一种改进的 YOLOv11 的实时目标检测算法。首先,针对主干网络中因池化层特征丢失导致的小目标识别困难的问题,在 AIFI 基础上提出了 DSEAIIFI 替换了主干网络中的池化层。其次,为了改善颈部网络对特征的利用和融合能力,同时提高对遮挡目标的检测能力提出了 MFFNeck 网络,提高了模型对上下文特征的融合能力和适应性。最后,为进一步提高网络对复杂动态环境的适应能力,并且突出高级特征在特征图中的重要级别,在头部网络中融合了针对检测头设计 LAAFPN 网络。为验证所提出的算法性能,进行了仿真与实车实验,仿真结果表明在 KITTI 数据集上改进的算法 mAP@0.5 和 mAP@0.5:0.95 分别为 91.1% 和 70.1%,与基础模型相比分别提高了 2.1% 和 3.8%。实车实验结果表明所提算法的平均检测精度为 92.7%,相较于基础模型提升了 4.3%,并且具有较好的实时性。

关键词: YOLOv11; 车辆小目标检测; 遮挡目标检测; 特征融合

中图分类号: TP391.41; TN914 **文献标识码:** A **国家标准学科分类代码:** 460.99

Improved YOLOv11 intelligent vehicle dynamic environment
object detection algorithmYu Chengyi^{1,2} Gao Song^{1,2} Wang Pengwei^{1,2} Sun Binbin^{1,2} Zhang Rong^{1,2}

(1. School of Transportation and Vehicle Engineering, Shandong University of Technology, Zibo 255000, China;

2. Key Laboratory of Integrated Design and Intelligence of New Energy Vehicles in Shandong Province, Zibo 255000, China)

Abstract: An improved YOLOv11 real-time object detection algorithm is proposed to address the detection difficulties and inaccurate positioning of small and occluded targets in complex dynamic scenarios of intelligent vehicles. Firstly, in response to the difficulty of small object recognition caused by the loss of pooling layer features in the backbone network, DSEAIIFI was proposed on the basis of AIFI to replace the pooling layer in the backbone network. Secondly, in order to improve the neck network's ability to utilize and fuse features, as well as enhance its ability to detect occluded targets, the MFFNeck network was proposed, which improved the model's ability and adaptability to fuse contextual features. Finally, in order to further improve the adaptability of the network to complex dynamic environments and highlight the importance level of advanced features in the feature map, a LAAFPN network designed for the detection head was integrated into the head network. To verify the performance of the proposed algorithm, simulations and real vehicle experiments were conducted, and the simulation results showed that the improved algorithm was effective on the KITTI dataset mAP@0.5 and mAP@0.5:0.95 is 91.1% and 70.1% respectively, which is an improvement of 2.1% and 3.8% compared to the basic model. The actual vehicle experiment results show that the average detection accuracy of the proposed algorithm is 92.7%, which is 4.3% higher than the basic model and has good real-time performance.

Keywords: YOLOv11; vehicle small target detection; occlusion target detection; feature fusion

0 引 言

当今科技飞速发展,智能网联技术成为备受瞩目的焦

点^[1]。目标检测技术作为智能网联汽车的一项核心技术,逐步显示出其重要性^[2]。目前,目标检测方法主要分为基于手工设计的特征方法、基于传统机器学习的方法和基于

收稿日期:2025-02-20

^{*} 基金项目:国家自然科学基金(52102465)、国家自然科学基金面上项目(52475269)、山东省自然科学基金面上项目(ZR2024ME179, ZR2022MF230)、山东重大科技创新工程项目(2023CXGC010111)、山东省青年基金(ZR2021QF039)、山东省中小企业创新能力提升工程项目(2022TSGC2277)资助

深度学习的方法^[3]。基于深度学习的方法目前虽然取得显著进步^[4]。但是对于复杂动态环境下车辆小目标和遮挡目标的检测,在道路交通拥堵^[5]、交通事故频发的复杂路段^[6],精确度和实时性上还有待提升,因此本文对此展开了研究。

目前,常用的目标检测网络主要有 R-CNN 系列、SSD 和 YOLO 系列。其中 R-CNN 检测精度较高,在小物体检测和精确定位方面表现出色。但训练和推理时间较长,实时性有待提高。SSD 虽然在网络结构方面较为简单,实时性较好,但是对目标形状和姿态适应能力有限。而 YOLO 算法可以通过利用多个不同尺度的特征图来捕捉不同大小和形状的物体,并且使用了基于 Transformer 的骨干网络,能够捕获长程依赖关系,更好地理解图像的全局上下文信息,而且兼顾了算法的实时性。目前 YOLO 系列分为许多版本,其中 YOLOv11 不仅是 YOLO 系列最新的版本而且是最具代表性目标检测方法^[7],相比于 R-CNN 系列和 SSD 具有更强的适应能力。但是 YOLOv11 模型在远距离车辆小目标、遮挡目标、复杂背景干扰检测方面还面临着许多问题。首先,远距离车辆小目标在图像中所占的像素极少,图像提供的特征信息匮乏,模型难以精准的勾勒出目标轮廓与关键特征,增加了检测的不确定性^[8]。其次,遮挡情况频发,尤其是在拥堵的路段,仅露出局部车身,模型难以依据残缺信息还原目标的完整状态。最后,真实环境中复杂背景干扰强烈^[9],背景中相似的纹理、颜色都可能导致漏检、误检。

首先,针对小目标检测问题,Xuan 等^[10]提出了一种轻量级自适应的无人机巡检模型(lightweight and adaptive model of YOLOv11,LAM-YOLOv11)。首先,设计了轻量级自适应模块(lightweight and adaptive model of C2K3,LA-C2K3),从输入图像中提取多个不同尺度的特征图,其次设计了一种自适应模块避免了环境噪声对小目标检测的干扰,但模型的鲁棒性还有待提升。Zou 等^[11]提出了一种基于 YOLOv11 的目标检测网络 YOLOv11-SDC,并且提出了一种新的稀疏特征(sparse feature,SF)模块替换了快速空间金字塔池化(spatial pyramid pooling fast,SPPF)模块,减少了池化过程中的特征丢失的问题,提升了目标检测的精度但没有解决复杂环境干扰的问题。强浩南等^[12]在 YOLOv11 的基础上提出了一种基于 CPM-YOLO 算法的高性能头盔识别模型,提出新的跨尺度特征融合方法(cross scale feature pyramid networks,CS-FPN)。并且提出了 PCT(partially CNN and Transformer block,PCT)模块优化了模型的特征提取能力,但检测目标较为单一。李学威等^[13]提出了一种基于 YOLOv11n 模型的改进算法 YOLOv11n-DPD。在 C3k2 模块中引入可变形卷积(deformable convolution,DCN)并且提出了一种结合 Focal Loss 与 DIoU Loss 的 Focal-DIoU 损失函数,解决了小目标检测中的各目标类别不平衡问题,但检测精度仍需进一步提升。其次,针对目标遮挡问题。林哲等^[14]提出了

YOLO-CA-SD 和语义分割的遮挡目标检测模型及抓取方法,提高了对遮挡目标的检测精度。Zhou 等^[15]提出了一种基于线性注意力机制和自适应下采样的轻量级无人机检测方法(linear attention mechanism based on improved YOLOv11,LAMS-YOLO)。使用了深度可分离卷积并且在注意力机制中结合了 LSTM 的门控系统,增强了语义提取效率,提高了目标遮挡场景下的检测性能。但是对小目标的检测精度还有待提升。涂育智等^[16]提出了一种基于 YOLOv11 改进的轻量化模型(multi-attention YOLOv11s,MA-YOLOv11s)选择性地引入小目标检测层并且设计了融合多种注意力机制的轻量级特征提取模块 C2SCSA 和 C2MCA 提升了模型在密集重叠目标场景中的检测精度与鲁棒性。但还需进一步优化模型的计算量。最后,针对复杂动态环境干扰问题,Tian 等^[17]提出一种结合灰度投影和 YOLOv11 模型的改进车辆识别算法。首先,利用灰度投影技术对采集到的图像进行预处理,增强对比度,减少了背景干扰,提高了目标特征的清晰度,但模型对复杂背景的泛化能力有待提升。Wang 等^[18]提出一种基于 YOLOv11 的目标检测算法 SDS-YOLO。引入了注意力模块增强了对复杂背景的免疫能力,但对小目标的检测并不敏感。另外,Gao 等^[19]提出了一种基于改进 YOLOv11 的轻量级反无人机检测方法。不仅引入 HWD 作为骨干的下采样模块。而且使用 CCFM 结构代替原来的颈部网络。提高模型对小目标的检测能力和复杂环境适应能力。王晓峰等^[20]基于 YOLOv11n,提出了深度特征强化与路径聚合优化的目标检测方法,设计了全局-局部特征增强模块(global local feature enhancement module,GLFEM)结合了局部特征和全局特征,虽然提高了模型对复杂环境的适应能力但参数量和计算开销增大。

综上所述,研究人员提出了许多改进的 YOLOv11 算法,尽管相比于基础模型都有明显的提升,但大都只是针对特定环境下的单一小目标或遮挡目标,而面对复杂多变的实车环境仍存在一定局限性^[21]。针对这一问题,为了兼顾模型对车辆小目标、遮挡目标和动态复杂环境干扰下的检测能力,同时能够满足车辆行驶时的实时性的要求,本文提出了一种改进的 YOLOv11 模型。首先,针对主干网络中因池化层特征丢失导致的小目标识别困难的问题,创新性的在尺度内特征融合模块(attention-based intrascale feature interaction,AIFI)的基础上提出了膨胀注意力尺度内特征融合(dilated squeeze and excitation attention-based intrascale feature interaction,DSEAIFI)模块,增强了主干网络对特征的提取能力,解决了因特征丢失而造成的小目标检测能力下降的问题。其次,为了改善颈部网络对特征的利用和融合能力,同时提高对遮挡目标的检测能力,对颈部网络进行了重新设计,并且提出了多重特征融合颈部网络(multi feature fusion neck,MFFNeck)替换了基础模型中的颈部网络,提高了各层特征的质量,改善了因特征质量

不佳而造成的漏检现象。最后,为进一步提高网络对复杂动态环境的适应能力,并且突出高级特征在特征图中的重要级别,在头部网络中设计了轻量化注意力渐进特征金字塔(lightweight attention asymptotic feature pyramid network, LAAFPN)网络,进一步增强了网络特征渐进融合的能力。通过仿真实验和在实车试验,证明改进的算法对复杂动态场景的车辆小目标和遮挡小目标具有更好的检测效果。

1 改进的 YOLOv11 算法

YOLOv11 是一种高效的目标检测算法,在主干网络中通过引入 C3K2 模块,使得更多的分支实现跨层连接,丰

富了模型的梯度流^[22]。其颈部网络采用 PAN-FPN 架构^[23],使特征更好的融合和交互,提高了对不同尺度目标的检测能力。YOLOv11 在头部网络中使用了解耦的结构^[24],提高了分类的准确性和效率。

总体而言,YOLOv11 是一种性能卓越、环境适应能力强、易于使用的算法模型,但在复杂动态环境下的车辆小目标、遮挡目标检测时仍存在一些局限性。针对此局限性,本文基于 YOLOv11 模型做了以下改进,首先在骨干网络中提出了 DSEAIPI 替换了原有的 SPPF 模块,其次在原有的颈部网络的基础上提出了 MFFNeck 网络,最后在头部网络中融合了基于检测头设计的 LAAFPN。改进的 YOLOv11 结构如图 1 所示。

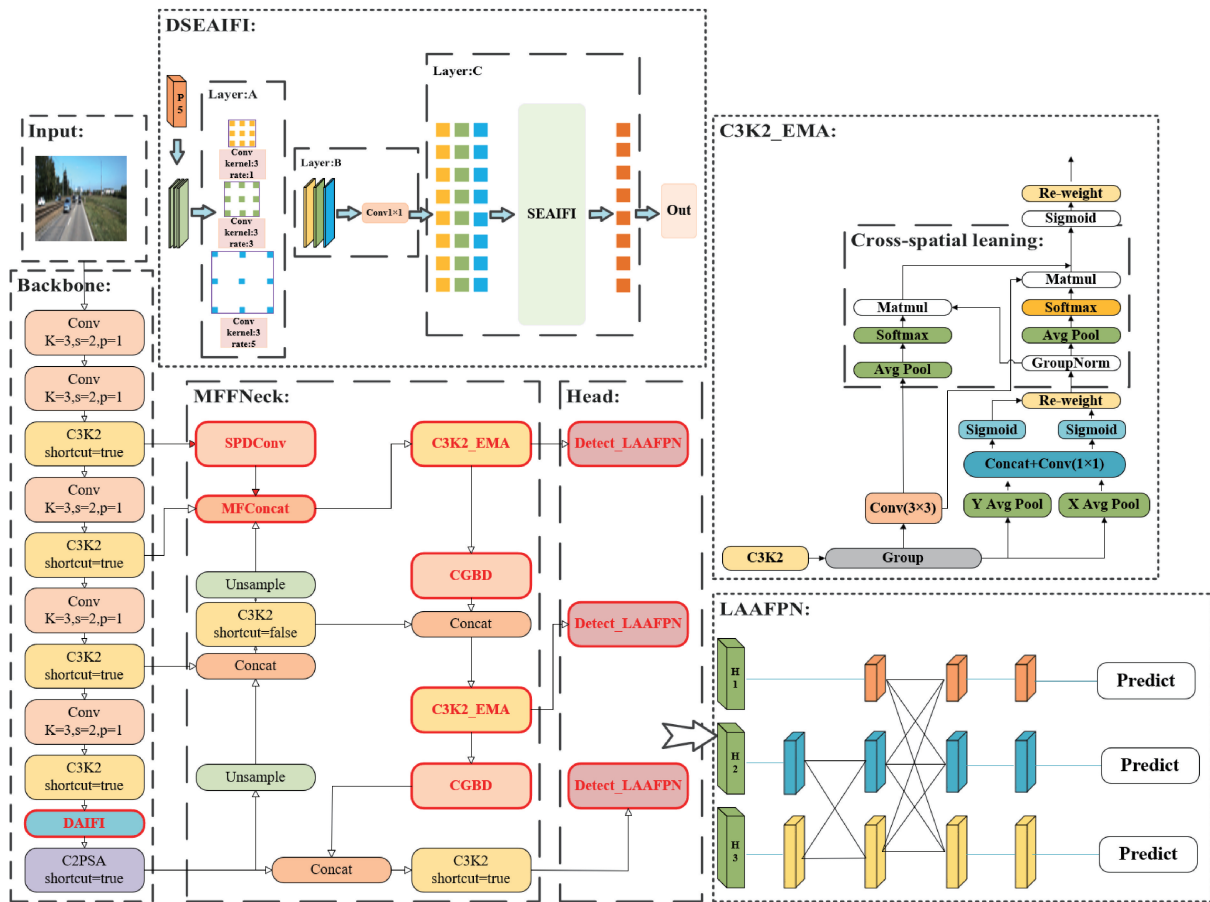


图 1 改进的 YOLOv11 目标检测算法

Fig. 1 Improved YOLOv11 object detection algorithm

1.1 DSEAIPI 模块

基础模型主干网络中的 SPPF 层会造成细节特征信息丢失、计算资源消耗大和对超参数过度敏感等问题,并不能满足对远距离车辆小目标的检测需求。针对此问题本文提出了 DSEAIPI 模块来代替传统 YOLOv11 算法中的 SPPF 层。DSEAIPI 模块是在 AIFI 的基础上嵌入了 SENet(squeeze and excitation network, SENet)网络,从而设计出 SEAIPI (squeeze and excitation attention-based

intrscale feature interaction, SEAIPI) 模块。同时增加了不同膨胀率(1, 3, 5)的膨胀卷积,将提取到的特征在进入 SEAIPI 模块前进行融合,扩大了模型的感受野。SEAIPI 模块一方面通过 SENet 增强了模型在通道上的表达能力,另一方面通过 AIFI 模块进行特征融合。保证了在微小特征不丢失的同时,使得网络可以更深更稳定的训练。DSEAIPI 的具体结构如图 2 所示。

P5 层的特征图在进入 SEAIPI 模块进行混合编码

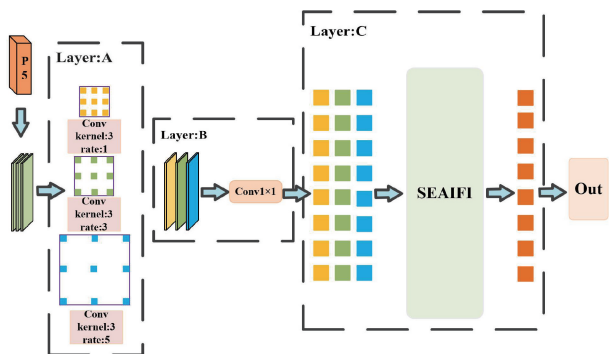


图 2 DSEAIFI 具体结构

Fig. 2 DSEAIFI specific structure

之前首先进入 A 层进行膨胀率为 1, 3, 5 的膨胀卷积, 提取不同尺度的特征信息。接着在 B 层进行特征融合

拼接, 并通过卷积核为 1×1 的卷积将提取到的特征通过压缩后传递到 SEAIFI 模块。确保在尽量不增加计算量的同时, 提高小目标识别的精度。DSEAIFI 过程如式(1)所示。

$$\begin{cases} \psi = \text{Concat}(\text{DConv}(P5)_{r=1}, \text{DConv}(P5)_{r=3}, \text{Conv}(P5)_{r=5}) \\ \mu = \text{Conv}(\psi) \\ \nu = \text{SEAIFI}(\mu) \end{cases} \quad (1)$$

其中, ψ 表示通过三次膨胀卷积融合之后的特征图。P5 表示从主干网络 P5 中输出的特征图。 $\text{DConv}(X)_{r=n}$ 表示对 X 进行膨胀率为 n 膨胀卷积。 μ 表示经过 1×1 卷积后输出的特征图。 ν 表示经过 AIFI 融合之后输出的特征图。 $\text{SEAIFI}(X)$ 表示对 X 进行注意力尺度内特征融合。SEAIFI 过程如图 3 所示。

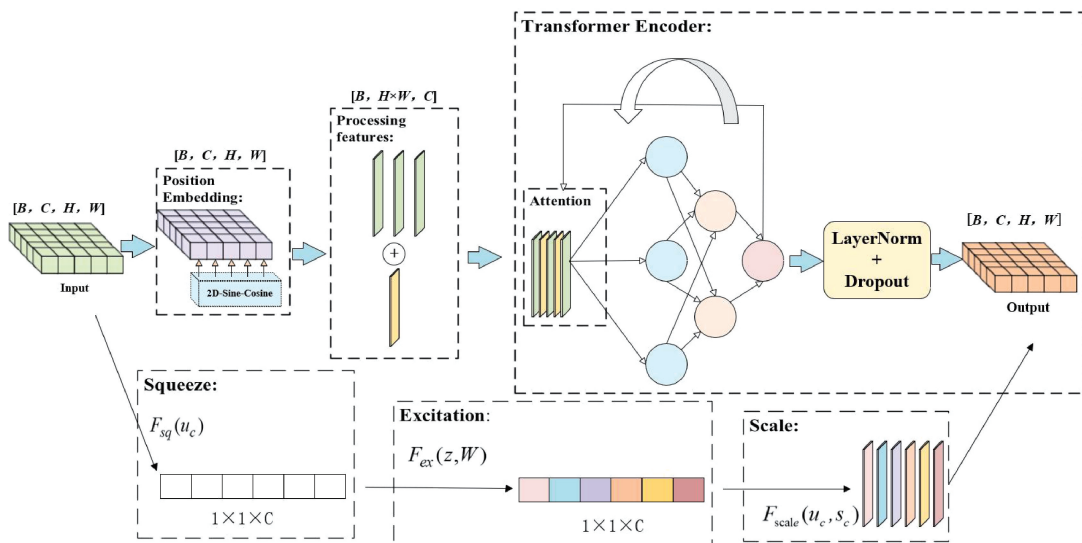


图 3 SEAIFI 具体结构

Fig. 3 Specific structure of SEAIFI

Squeeze 模块如式(2)所示。

$$Z_c = F_{sq}(u_c) = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W u_c(i, j) \quad (2)$$

式中: 对 u_c 进行全局平均池化, 在空间 $H \times W$ 上压缩后得到 Z_c 。

Excitation 模块如式(3)所示。

$$s = F_{ex}(z, W) = \sigma(g(z, W)) = \sigma(W_2 \delta(W_1 z)) \quad (3)$$

式中: W_1 和 W_2 分别表示两个全连接, δ 表示通道权重。

Scale 模块如式(4)所示。

$$\tilde{x}_c = F_{scale}(u_c, s_c) = s_c u_c \quad (4)$$

式中: s_c 表示生成的特征向量 ($1 \times 1 \times C$), $u_c (H \times W \times C)$ 表示特征图。

注意力机制如式(5)所示。

$$\text{Attention}(Q, K, V) = \text{Softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (5)$$

式中: Q 是查询矩阵, K 是关键字, V 是值。其中 d_k 等于向量的长度。

经过对 SPPF 层的替换, 不仅减少了模型的复杂度, 并且保留了图像中的细节特征, 避免了因池化操作而造成的计算资源浪费、超参数敏感等问题。相比于基础模型中的 SPPF 层, DSEAIFI 模块更满足复杂动态场景下车辆小目标的检测需求。

1.2 改进颈部网络

基础模型的颈部网络中难以将 P3、P4、P5 层提取到的小目标特征完美融合^[25], 而且在面对遮挡目标时, 细颗粒特征提取困难。为了提高其特征提取能力, 以满足对大量遮挡目标的检测需求。本文提出了 MFFNeck 网络。首先, 在颈部网络中对 P2 层提取到的特征图经过空间深度转换卷积模块 (space-to-depth convolution, SPD-Conv) 处理, 与 P3 层和上采样层 (Upsample) 通过设计的多重特征

连接模块(multi feature concat,MFConcat)进行融合,不仅保留了浅层网络的细颗粒度信息,而且提高了对低分辨率和小物体的识别性能。其次引入了轻量级下采样模块(context guided block down,CGBD)模块替换了 CBS 模块。在一定程度上降低了模型的复杂度。最后将 C3K2 模块和 EMA 注意力机制进行融合设计了 C3K2_EMA 模块。MFFNeck 具体结构如图 4 所示。

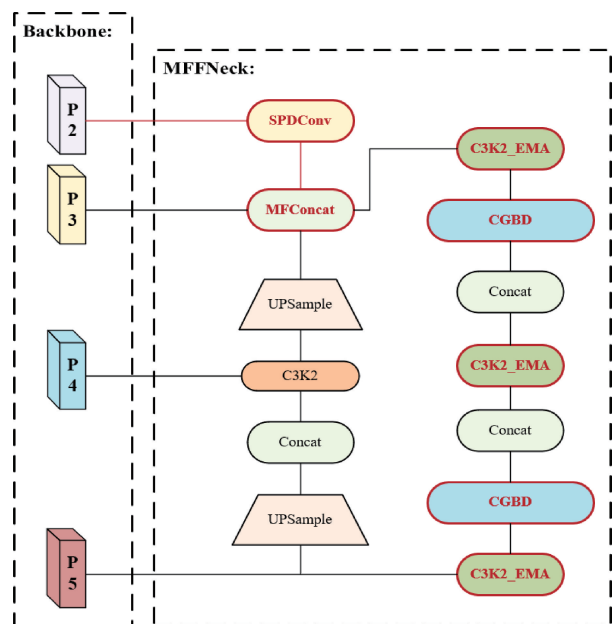


图 4 MFFNeck 网络结构

Fig. 4 MFFNeck network structure

MFConcat 具体结构如图 5 所示。

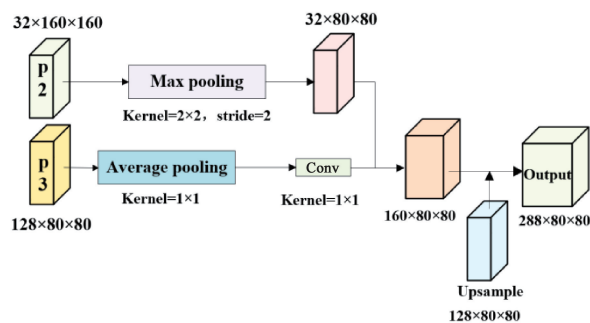


图 5 MFConcat 结构

Fig. 5 MFConcat structure

C3K2 是 YOLOv11 模型中的一个重要的特征提取组件,是基于传统 C3 模块改进而成。通过结合可变卷积核和通道分离的策略,提供了比传统 C3 更加强大的特征提取能力,更加适合复杂场景和深层次的特征提取任务。为了进一步加强其高级特征提取能力。在 C3K2 模块的基础上加入了 EMA 注意力机制,从而设计出了 C3K2_EMA 模块。C3K2_EMA 结构如图 6 所示。

通过对颈部网络的改进,增加了特征在颈部网络中的融合机制,为遮挡目标的检测提供了更多有效的特征信息。提高了模型上下文信息联系的能力,使其能够在结合全局特征信息的同时对遮挡目标进行识别,而且增强了模型对像素级别特征的关注度,进而提高了模型对遮挡目标的检测精度。

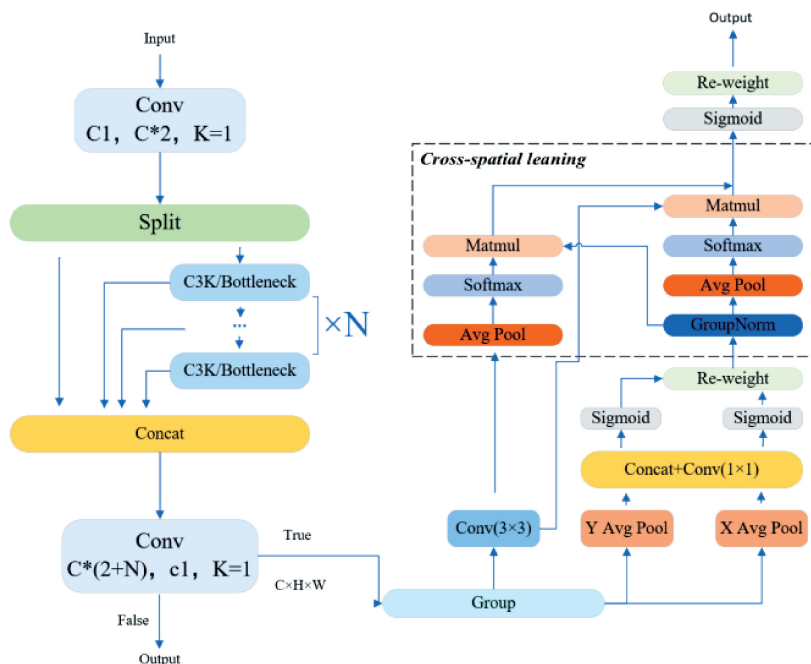


图 6 C3K2_EMA 结构

Fig. 6 C3K2_EMA structure

1.3 改进头部网络

在基础模型的头部网络中虽然引入了多层卷积模块,但是模型对复杂动态环境的适应能力还有待提升。为了满足模型对复杂动态环境的适应能力的需求,本文基于 AFPN 设计了 LAAFPN。LAAFPN 具体结构如图 7 所示。

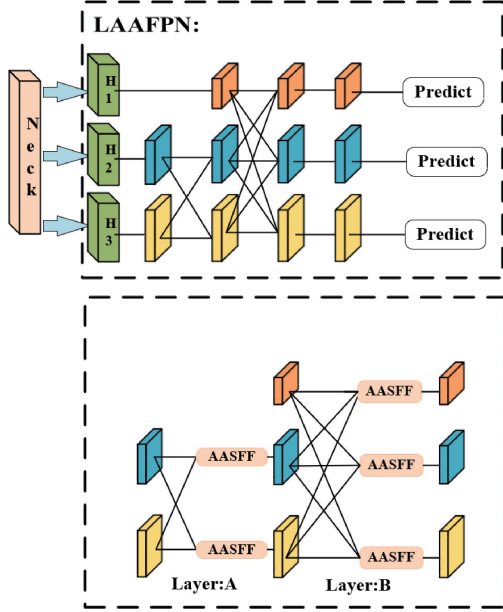


图 7 LAAFPN 结构

Fig. 7 LAAFPN structure

首先对传统的 AFPN 进行剪枝,降低其深度和广度。并且在 ASFF 的基础上设计了注意力自适应空间特征融合模块(attention adaptive spatial feature fusion, AASFF),在多级特征融合过程中各层特征先通过设计的层间交互空间注意力机制(inter layer interaction space attention mechanism, ILISAM),然后为不同级别的特征分配不同的空间权重,增强了关键特征的重要级别,减轻来自不同单元的矛盾信息的影响。AASFF 权重计算公式如式(6)所示。

$$\begin{aligned} \text{LayerA: } y_{i,j}^l &= \delta_{i,j}^l \cdot x_{i,j}^{2 \rightarrow l} + \eta_{i,j}^l \cdot x_{i,j}^{3 \rightarrow l} \\ \text{LayerB: } y_{i,j}^l &= \alpha_{i,j}^l \cdot x_{i,j}^{1 \rightarrow l} + \beta_{i,j}^l \cdot x_{i,j}^{2 \rightarrow l} + \gamma_{i,j}^l \cdot x_{i,j}^{3 \rightarrow l} \end{aligned} \quad (6)$$

式中: $y_{i,j}^l$ 表示第 n 层在 (i, j) 处的加权后获得的向量, $x_{i,j}^{1 \rightarrow l}, x_{i,j}^{2 \rightarrow l}, x_{i,j}^{3 \rightarrow l}$ 分别表示从第 n 层到第 l 层自适应特征融合获得的向量, $\alpha_{i,j}^l, \beta_{i,j}^l, \gamma_{i,j}^l, \delta_{i,j}^l, \eta_{i,j}^l$ 分别表示水平 l 上的五个水平特征的空间权重,且 $\alpha_{i,j}^l + \beta_{i,j}^l + \gamma_{i,j}^l = 1, \delta_{i,j}^l + \eta_{i,j}^l = 1$ 。ILISAM 具体结构如图 8 所示。

ILISAM 模块在保留了每层权重向量的基础上加入各层融合后的权重向量,在保持每层权重特点的同时兼顾了其他各层的权重特点,在复杂动态环境干扰情况下的目标识别上更具有优势。权重融合如式(7)所示。

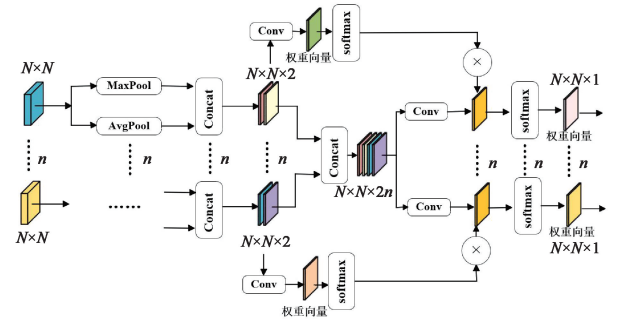


图 8 ILISAM 结构

Fig. 8 ILISAM structure

$$\begin{cases} v_i(f_1, \dots, f_n) = \tau(f_i) \cdot \xi(f_1, \dots, f_n) \\ \tau(f_i) = \text{Softmax}(\text{Conv}(\kappa(f_i))) \\ \xi(f_1, \dots, f_n) = \text{Softmax}(\text{Concat}(\kappa(f_1), \dots, \kappa(f_n))) \\ \kappa(f_i) = \text{Concat}(\text{Conv}(\text{MaxPool}(f_i)), \text{AvgPool}(f_i)) \end{cases} \quad (7)$$

式中: v_i 表示第 i 层融合后的权重系数, f_n 表示第 n 层的特征图, $\tau(f_i)$ 表示对第 i 层的特征图处理得到的权重系数, $\xi(f_1, \dots, f_n)$ 表示融合全部特征图得到的权重系数。ASFF 和 AASFF 结构对比如图 9 所示。

通过对头部网络进行改进,引入了基于 AFPN 设计的 LAAFPN,通过对特征的加权融合,突出了关键特征,同时降低了无效特征对模型检测能力的干扰。通过特征的相互融合,实现了网络之间特征的共享,提高了模型在复杂场景中的检测精度,有效的降低了因复杂动态场景噪声多而造成漏检、误检的可能性。

2 实验验证和分析

为了验证改进的 YOLOv11 在复杂动态环境下的车辆小目标、遮挡目标检测的有效性和实用性,本文分别开展了仿真实验和实车试验。

2.1 仿真试验

1) 数据集

本文主要选用了 KITTI 数据集^[26]。对于目标检测,训练集、验证集和测试集分别有 7 434、2 042 和 2 048 张图片。其中包含了大量复杂场景下车辆小目标、遮挡目标的情况,可以更好的检测模型的性能。为了进一步验证算法的泛化性能分别选用了 COCO 数据集^[27]和 BDD100K 数据集^[28]。其中包含了多种目标的特征信息,可以有效的验证模型的泛化性能。从 COCO 数据集和 BDD100K 数据集上各筛选 20 000 张图片作为训练集、验证集和测试集对模型进行泛化实验。

2) 实验环境

本文采用 Windows11 操作系统,搭载 Intel(R) Core (TM) i9-14900HX,主频为 2.2 GHz。显卡为 NVIDIA GeForce RTX 4070 Laptop GPU,内存为 32 G。实验环境

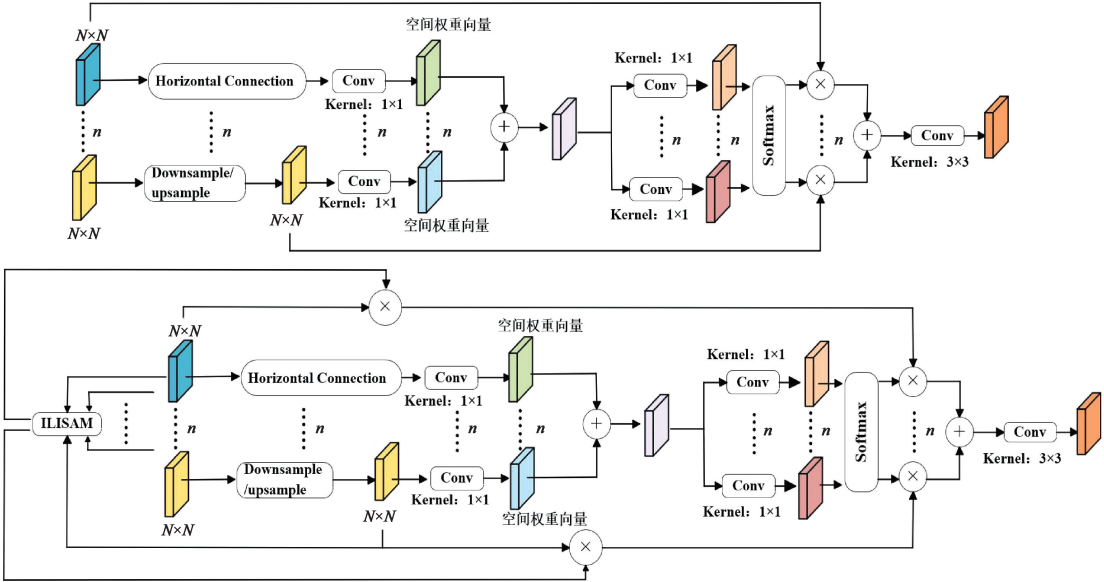


图 9 ASFF 模块和 AASFF 模块对比

Fig. 9 Comparison of ASFF module and AASFF module

为 Pycharm 和 Anaconda,Pytorch 框架版本为 2.4.1, Anaconda 中 Python 版本为 3.8.20,Batch size 设置为 32, 图像大小为 640×640。epoch 设置为 300。具体实验环境信息如表 1 所示。

表 1 实验环境	
Table 1 Experimental environment	
硬件/软件	配置
CPU	Intel(R) Core(TM)i9-14900HX
GPU	NVIDIA GeForce RTX 4070 GPU
内存	32 G
Pycharm	pycharm-community-2023.3.4
Anaconda	Anaconda3-2024.02-1-Windows-x86_64
Pytorch	Pytorch 2.4.1
Python	Python 3.8.20
CUDA	CUDA 11.8

3)评价指标

本文采用 mAP(mean of average precision,mAP)作为检测精度的评价指标,mAP 是目标检测中一个重要的评估指标,用于衡量模型在不同召回率水平下的平均精度,因此是一个综合性评价指标,其值越大模型越好。

精确度(Precision)的计算公式如式(8)所示。

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

(8)

召回率(Recall)的计算公式如式(9)所示。

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

(9)

AP 的计算公式如式(10)所示。

$$AP = \int_0^1 P d(R)$$

(10)

mAP 的计算公式如式(11)所示。

$$mAP = \frac{\sum_{i=1}^k AP_i}{k}$$

(11)

4)对比试验

通过对不同模型在 KITTI 数据集上的对比实验,证明了改进 YOLOv11 算法性能有了较大的提升。对比实验结果如表 2 所示。

表 2 对比试验		
Table 2 Comparative experiments		
算法	mAP@0.5	mAP@0.5:0.95
SSD	0.751	0.316
YOLOv7-tiny	0.876	0.613
YOLOv6s	0.867	0.596
YOLOv5s-OEAG	0.904	0.621
YOLOv11n	0.890	0.663
YOLOv10	0.891	0.687
YOLOv8	0.898	0.673
Faster-RCNN	0.549	0.447
改进的 YOLOv8m 算法 ^[29]	0.743	0.554
SEP-YOLO ^[30]	0.667	0.454
改进的 YOLOv11 ^[25]	0.903	0.716
SES-YOLOV8n ^[31]	0.927	0.692
本文算法	0.911	0.701

分析可得本文改进 YOLOv11 算法的 mAP@0.5 和 mAP@0.5:0.95 分别为 0.911 和 0.701,均处于前列。比

基础模型 YOLOv11 分别提高了 2.1% 和 3.8%，相较于 YOLOv5s-OEAG 提高了 0.7% 和 3.8%。与最新改进的 YOLOv8m 算法^[29]相比提高了 16.8% 和 14.7%。相比于 SES-YOLO8n^[31]，mAP@0.5 稍有下降，但 mAP@0.5:0.95 提高了 0.9%。与同类改进的 YOLOv11 的算法^[25]相比 mAP@0.5 和 mAP@0.5:0.95 基本持平。

综上所述，改进的 YOLOv11 模型的检测效果优

于基础模型，并且相较于最新同类改进算法也具有一定的优势，进一步验证了模型在复杂动态场景下车辆小目标和遮挡目标检测的可行性和有效性。

5)消融实验

为了进一步验证改进的各个模块对模型的提升能力，本文在 KITTI 数据集上模型进行消融实验，√表示添加了该模块。消融实验结果如表 3 所示。

表 3 消融实验
Table 3 Ablation experiment

Baseline	DSEAIPI	MFFNeck	LAAFPN	mAP@0.5	mAP@0.5:0.95	P	R	Param/M
√				0.890	0.663	0.894	0.771	2.58
√	√			0.897	0.687	0.901	0.786	2.74
√	√	√		0.905	0.698	0.926	0.825	2.88
√	√	√	√	0.911	0.701	0.939	0.833	3.15

由表 3 可知本文提出的 3 种改进都对基础模型起到了提升的作用。首先，加入了 DSEAIPI 模块，通过将特征图通过不同膨胀率的膨胀卷积处理并融合拼接，提高了网络提取特征的能力。结果显示 mAP@0.5 和 mAP@0.5:0.95 分别提高了 0.7% 和 2.4%，精确度和召回率提高了 0.7% 和 1.5%，参数量增加了 1.6 M。其次，添加了 MFFNeck 网络，提高了模型全局上下文联系能力。mAP@0.5 和 mAP@0.5:0.95 分别提高了 0.8% 和 1.1%，精确度和召回率提高了 2.5% 和 3.9%，参数量增加了 0.14 M。最后，加入了 LAAFPN 网络，精确度和召回率提高了 1.3% 和 0.8%，参数量增加了 0.27 M。为了研究参数量的增加对实时性和 mAP 值的影响，更好的平衡其之间的关系，并且进一步说明算法的优越性。本文与该领域同类最新算法在 KITTI 数据集上进行对比分析。综合性能分析结果如表 4 所示。

表 4 模型综合性能分析

Table 4 Comprehensive performance analysis of the model			
模型	参数量/	mAP@	推理
	M	0.5:0.95	速度/fps
YOLOv11n	2.58	0.663	186
YOLOv8	3.10	0.673	177
MFF-YOLOv8s ^[32]	2.56	0.691	159
改进的 YOLOv11 ^[25]	3.87	0.716	128
本文算法	3.15	0.701	181

经过分析可知，本文算法相较于基础算法参数量有所增加，但是 mAP@0.5:0.95 有了较大提升，提升了 3.8%，推理速度基本持平。与 YOLOv8 算法相比参数量基本保持不变，但 mAP@0.5:0.95 和推理速度都有较大的提升，分别提升了 2.8% 和 4 fps。与最新提出的 MFF-YOLOv8^[32]相比，参数量和 mAP@0.5:0.95 基本相同，但

推理速度提升了 22 fps。与同类改进的 YOLOv11 算法相比，mAP@0.5:0.95 基本相同，但参数量下降了 0.72 M，推理速度提升了 53 fps。进一步证明了本文改进的 YOLOv11 算法在兼顾车辆小目标和遮挡目标的复杂动态环境下的同时具有更好的实时性和优越性。

6)泛化实验

为了验证改进的算法在其他数据集仍然具有良好的适应性，同时为了防止模型在训练时过拟合。本文在 BDD100K 数据集和 COCO 数据集上对模型进行了泛化实验。泛化实验结果如表 5 所示。

表 5 泛化实验
Table 5 Generalization experiment

数据集/模型	mAP@0.5	mAP@0.5:0.95
BDD100K/YOLOv11n	0.761	0.658
BBD100k/本文算法	0.795	0.699
COCO/YOLOv11n	0.641	0.535
COCO/本文算法	0.651	0.559

经过分析可知，通过在 BDD100K 数据集和 COCO 数据集进行的泛化性实验，在 BDD100K 数据集上与基础模型相比 mAP@0.5 和 mAP@0.5:0.95 分别提高了 3.4% 和 4.1%，在 COCO 数据集上 mAP@0.5 和 mAP@0.5:0.95 分别提高了 1% 和 2.4%。经验证改进的模型具有良好的泛化性能。

7)可视化对比分析

为了验证改进算法的性能，将本文模型与最新改进的 YOLOv11 模型^[25]和基础模型在 KITTI 数据集作可视化分析。可视化对比结果如图 10 所示。

首先，对于骨干网络，文献[25]从增强特征提取能力的角度出发，使用膨胀特征金字塔卷积(dilated feature pyramid convolution, DFPC)模块替代了原有的 SPPF 层。



图 10 改进模型与其他模型效果对比

Fig. 10 Compare the effectiveness of the improved model with other models

本文则是从增强特征融合和特征注意力的角度使用 DSEAIIFI 模块代替 SPPF 层,都避免了特征丢失的问题。从第 1 行图片中可以看出,增强特征融合具有更好的小目标检测效果。其次,针对颈部网络,文献[25]提出了目标上下文多尺度特征金字塔增强了全局和局部特征的联系。本文则提出了 MFFNeck 结构,通过设计的多重特征融合模块进一步对高级特征进行融合,并且使用了轻量化卷积以减少模型复杂度。从第 3、4 行图片中可以看出多重特征融合在遮挡目标识别上具有较大的优势。最后,在头部网络中,文献[25]引入了 DyHead 增强特征表示能力。本文设计了 LAAFPN 网络,渐进的将各层特征进行融合,并且通过权重控制的方式减轻来自不同单元的矛盾信息影响。从第 2 行组图片中可以看出,即使面对更加复杂的环

境背景,本文算法也有着较高的识别精度。综上所述,本文改进的模型在对复杂动态环境车辆小目标检测和遮挡目标的检测能力上具有较大的优势。

2.2 实车实验

为了进一步验证改进的 YOLOv11 算法的实用性,本文选取了较为拥堵路段作为实车实验的场景,并且采用哈弗 H7 自动驾驶汽车作为实验平台。其上搭载了树莓派 5 (树莓派 Ubuntu20.04 操作系统) 开发板,工控机 (Ubuntu20.04)通过 ROS 系统将视频数据传输到树莓派 ROS 系统中,其中树莓派中分别配置基础模型文件、文献[25]算法和本文算法。为了减小检测误差,对哈弗 H7 的摄像头进行标定。实验平台和部分相机标定图如图 11 所示。

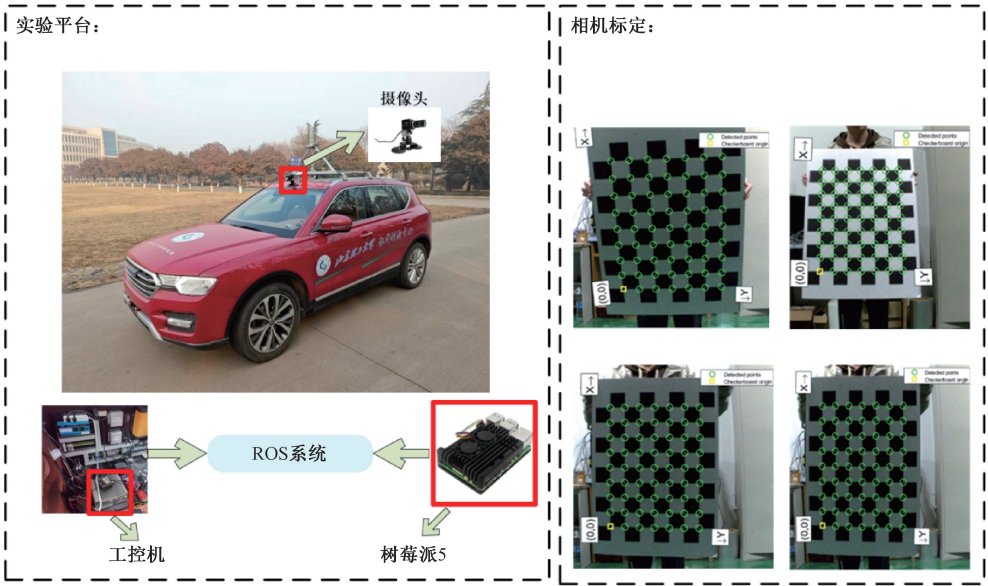


图 11 实验平台

Fig. 11 Experimental platform

相机误差分布如图 12 所示。

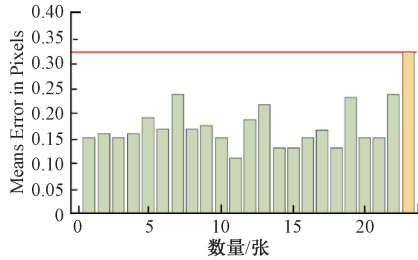


图 12 相机误差分布图
Fig. 12 Camera error distribution diagram

由图可知相机的总体平均误差为 0.18,最大平均误差

不超过 0.33,标定结果可靠。
相机的径向畸变系数如式(12)所示。
 $[k_x, k_y, k_z] = [-0.384\ 2\ \ -0.076\ 25\ \ 5.827]$ (12)
相机的切向畸变系数如式(13)所示。
 $[P_1, P_2, P_3] = [0.004\ 1\ \ 0.003\ 2]$ (13)
相机的内参矩阵如式(14)所示。

$$K = \begin{pmatrix} 958.1 & 0.75 & 329.69 \\ 0 & 955.4 & 197.4 \\ 0 & 0 & 1 \end{pmatrix} \quad (14)$$

依据以上标定的结果对树莓派中的视频数据进行矫正,提高数据的质量,减少模型检测的外在干扰因素。树莓派检测结果如图 13 所示。

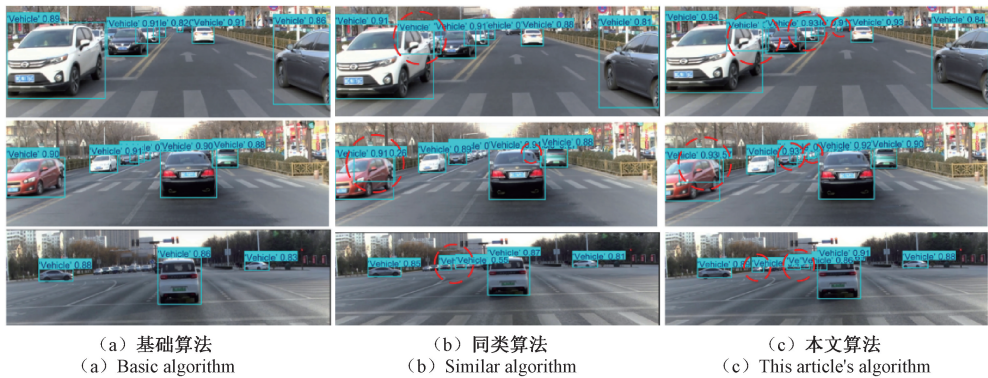


图 13 实车实验对比图
Fig. 13 Comparison of real vehicle experiments

首先,从实车对比图片中可以看出,本文改进的 YOLOv11 通过微小细节可以检测到被遮挡后的车辆,基础模型没有检测出,同类改进算法只检测到部分被遮挡车辆。证明了改进的模型对遮挡目标的识别能力有了较大的提升。进而说明添加了多重融合机制的 MFFNeck 网络在特征融合上具有一定的优势。其次,基础模型和同类算法对复杂环境下小目标的识别能力有限,而本文改进的 YOLOv11 模型则可以通过 DSEAIIFI 层保留更多的细颗粒度特征,从而提高了模型对小目标的识别度。最后,在有大量复杂背景干扰的情况下,尤其是在第 3 行图片中有大量树木影响的情况下,本文改进模型的识别度较高。说明通过 LAAFPN 网络对特征重新整合,提高了模型复杂场景适应能力。总体而言,本文改进的 YOLOv11 模型在复杂动态环境下车辆小目标、遮挡目标检测上具有更好的性能。

3 结 论

针对目前复杂动态环境下车辆小目标、遮挡目标特征提取困难,环境适应能力有限的问题,本文提出了一种改进的 YOLOv11 模型。首先,为了解决基础模型中特征遗漏的问题,提出了 DSEAIIFI 模块替换了基础模型主干网络中的池化层(SPPF),改进的模型在车辆小目标识别方面能

力有所提高,充分说明了 DSEAIIFI 模块的可行性。其次,提出了 MFFNeck 网络,即将 P2 层、P3 层和上采样层的特征进行融合,并且引入了 CGBD 模块,而且在 C3K2 模型的基础上提出了 C3K2_EMA 模块,改进的模型在遮挡目标识别能力方面有了较大的提升。最后,设计了 LAAFPN 网络提高了模型复杂动态环境下的检测能力和适应能力。由仿真实验结果可知,改进模型在 KITTI 数据集上的 mAP@0.5 和 mAP@0.5:0.95 分别为 91.1%和 70.1%,与 YOLOv11 模型相比分别提高了 2.1%和 3.8%,并且对于小目标、遮挡目标和复杂动态场景下的检测能力优于同类算法。经过实车实验的验证,基础模型的平均检测精度为 88.4%,改进的 YOLOv11 模型的平均检测精度为 92.7%,相较于基础模型提高了 4.3%,说明改进的模型具有更好的实用性。

在未来的研究中考考虑不影响模型精确度和性能的情况下对模型进行轻量化,减少模型的参数量,进一步提高模型的综合性能,使其更好的应对真实复杂环境。

参考文献

[1] NGOC T H, NGUYEN H K, HUA K H, et al. Optimizing YOLO performance for traffic light detection and end-to-end steering control for

- autonomous vehicles in Gazebo-ROS2[J]. International Journal of Advanced Computer Science and Applications(IJACSA),2023,14(7): 475-484.
- [2] WEI L. Vehicle detection in foggy weather based on an enhanced YOLO method[J]. Journal of Physics: Conference Series,2022,2284(1):012015.
- [3] ALBUQUERQUE C, HENRIQUES R, CASTELLI M. Deep learning-based object detection algorithms in medical imaging: Systematic review [J]. Heliyon, 2025, 11(1):e41137.
- [4] KIRAN K V, DASH S, PARIDA P. Vehicle detection in varied weather conditions using enhanced deep YOLO with complex wavelet[J]. Engineering Research Express,2024,6(2):025224.
- [5] YADAV K D, MISHRA S. Vehicle detection in high density traffic surveillance data using YOLO. v5[J]. Recent Advances in Electrical & Electronic Engineering,2023, 17(2):216-227.
- [6] ALI L M, ZHANG Z. The YOLO framework: A comprehensive review of evolution, applications, and benchmarks in object detection[J]. Computers, 2024, 13(12): 336-336.
- [7] 游昊,吕文涛,叶丹,等.基于深度学习的车辆检测方法研究进展[J]. 无线电工程,2025,55(2):230-245.
- YOU H, LYU W T, YE D, et al. Research progress on vehicle detection methods based on deep learning[J]. Journal of Radio Engineering,2025,55(2):230-245.
- [8] 李琼,考月英,张莹,等.面向无人机航拍图像的目标检测研究综述[J]. 图学报,2024,45(6):1145-1164.
- LI Q, KAO Y Y, ZHANG Y, et al. A review of research on object detection for unmanned aerial vehicleaerial images[J]. Journal of Graphic Science, 2024,45(6): 1145-1164.
- [9] NIU C, SONG Y, ZHAO X. SE-Lightweight YOLO: Higher accuracy in YOLO detection for vehicle inspection [J]. Applied Sciences, 2023, 13(24):13052.
- [10] XUAN Y, ZHANG Y X, LI C, et al. LAM-YOLOv11 for UAV transmission line inspection: Overcoming environmental challenges with enhanced detection efficiency [J]. Multimedia Systems, 2025, 31(2):110.
- [11] ZOU C, YU S, YU Y, et al. Side-scan sonar small objects detection based on improved YOLOv11[J]. Journal of Marine Science and Engineering, 2025, 13(1):162.
- [12] 强浩南,邹涌波,马立东,等.基于 CPM-YOLO 的道路头盔检测方法[J]. 电子测量技术, 2025, 48(7): 152-162.
- QIANG H N, ZOU Y B, MA L D, et al. Road helmet detection method based on CPM-YOLO[J]. Electronic Measurement Technology,2025,48(7):152-162.
- [13] 李学威,毛小坤,孙滨,等.基于 YOLOv11n 的无人机航拍小目标检测算法[J/OL]. 光通信研究,1-8[2025-03-11]. <http://kns.cnki.net/kcms/detail/42.1266.TN.20250210.1015.002.html>.
- LI X W, MAO X K, SUN B, et al. Small target detection algorithm for drone aerial photography based on YOLOv11n [J/OL]. Journal of Optical Communication Research, 1-8 [2025-03-11]. <http://kns.cnki.net/kcms/detail/42.1266.TN.20250210.1015.002.html>.
- [14] 林哲,潘慧琳,陈丹.融合改进 YOLO 和语义分割的遮挡目标抓取方法[J]. 电子测量与仪器学报, 2025, 38(12):190-201.
- LIN ZH, PAN H L, CHEN D. A method for capturing occluded targets by integrating improved YOLO and semantic segmentation [J]. Journal of Electronic Measurement and Instrumentation, 2025, 38(12):190-201.
- [15] ZHOU S, YANG L, LIU H, et al. A lightweight drone detection method integrated into a linear attention mechanism based on improved YOLOv11 [J]. Remote Sensing,2025,17(4):705.
- [16] 涂育智,王法翔,吴春霖.融合多注意力机制的轻量级无人机航拍小目标检测模型[J]. 计算机工程与应用, 2025,61(11):93-104.
- TU Y ZH, WANG F X, WU CH L. A lightweight unmanned aerial vehicle aerial small target detection model integrating multiple attention mechanisms[J]. Journal of Computer Engineering and Applications, 2025, 61(11):93-104.
- [17] TIAN S, LU Y, JIANG F, et al. Improved campus vehicle detection method based on YOLOv11 and grayscale projection-based electronic image stabilization algorithm[J]. Traitement du Signal,2024, 41(6):3335.
- [18] WANG D, TAN J, WANG H, et al. SDS-YOLO: An improved vibratory position detection algorithm based on YOLOv11[J]. Measurement,2025,244:116518.
- [19] GAO Y, XIN Y, YANG H, et al. A lightweight anti-unmanned aerial vehicle detection method based on improved YOLOv11[J]. Drones,2024,9(1):11.
- [20] 王晓峰,黄俊俊,谭文雅,等.基于深度特征强化与路径聚合优化的目标检测[J/OL]. 计算机科学,1-18[2025-

- 03-11]. <http://kns.cnki.net/kcms/detail/50.1075.tp.20250227.0933.005.html>.
- WANG X F, HUANG J J, TAN W Y, et al. Object detection based on deep feature enhancement and path aggregation optimization[J/OL]. Journal of Computer Science1-18[2025-03-11]. <http://kns.cnki.net/kcms/detail/50.1075.tp.20250227.0933.005.html>.
- [21] 陈金林, 吴一全, 苑玉彬. 无人机视角下目标检测的 YOLO 系列算法研究进展[J/OL]. 北京航空航天大学学报, 1-33[2025-01-23]. <https://doi.org/10.13700/j.bh.1001-5965.2024.0420>.
- CHEN J L, WU Y Q, YUAN Y B. Research progress on YOLO series algorithms for object detection from the perspective of unmanned aerial vehicles [J/OL]. Journal of Beijing University of Aeronautics and Astronautics, 1-33 [2025-01-23]. <https://doi.org/10.13700/j.bh.1001-5965.2024.0420>.
- [22] 茅智慧, 朱佳利, 吴鑫, 等. 基于 YOLO 的自动驾驶目标检测研究综述[J]. 计算机工程与应用, 2022, 58(15):68-77.
- MAO ZH H, ZHU J L, WU X, et al. A review of research on object detection in autonomous driving based on YOLO[J]. Journal of Computer Engineering and Applications, 2022, 58(15):68-77.
- [23] 米增, 连哲. 面向通用目标检测的 YOLO 方法研究综述[J]. 计算机工程与用, 2024, 60(21):38-54.
- MI Z, LIAN ZH. A review of YOLO method research for general object detection[J]. Journal of Computer Engineering and Applications, 2024, 60(21):38-54.
- [24] 兰红, 王惠钊. 结合轻量化与多尺度融合的交通标志检测算法[J]. 计算机工程, 2024, 50(10):381-392.
- LAN H, WANG H ZH. Traffic sign detection algorithm combining lightweight and multi-scale fusion[J]. Journal of Computer Engineering, 2024, 50(10):381-392.
- [25] 李彬, 李生林. 改进 YOLOv11n 的无人机小目标检测算法[J]. 计算机工程与应用, 2025, 61(7):96-104.
- LI B, LI SH L. Improved YOLOv11n UAV small target detection algorithm[J]. Journal of Computer Engineering and Applications, 2025, 61(7):96-104.
- [26] GEIGER A, LENZ P, STILLER C, et al. Vision meets robotics: The KITTI dataset [J]. The International Journal of Robotics Research, 2013, 32(11): 1231.
- [27] MICIKEVICIUS P, NARANH S, ALBEN J, et al. Mixed precision training [C]. 6th International Conference on Learning Representations. Vancouver; ICLR, 2018:1086-1097.
- [28] ABADI M, BARHAM P, CHEN J, et al. Tensorflow: A system for large-scale machine learning[J]. USENIX Association, 2016, DOI:10.48550/arXiv.1605.08695.
- [29] 孙虎成, 臧可, 吕晓晓. 基于改进 YOLOv8m 的道路目标检测算法[J/OL]. 电子设计工程, 1-9[2025-03-11]. <http://kns.cnki.net/kcms/detail/61.1477.TN.20241223.1724.002.html>.
- SUN H CH, ZANG K, LYU X X. Road object detection algorithm based on improved YOLOv8m[J/OL]. Journal of Electronic Design Engineering, 1-9 [2025-03-11]. <http://kns.cnki.net/kcms/detail/61.1477.TN.20241223.1724.002.html>.
- [30] 朱彦, 张月霞. SEP-YOLO: 基于 YOLOv8 改进的道路目标检测算法[J/OL]. 计算机应用与软件, 1-8[2025-03-11]. <http://kns.cnki.net/kcms/detail/31.1260.tp.20240829.1140.002.html>.
- ZHU Y, ZHANG Y X. SEP-YOLO: Improved road object detection algorithm based on YOLOv8[J/OL]. Journal of Computer Applications and Software, 1-8 [2025-03-11]. <http://kns.cnki.net/kcms/detail/31.1260.tp.20240829.1140.002.html>.
- [31] 刘罡, 闫曙光, 刘钰, 等. 基于感受野增强的复杂道路场景目标检测研究[J]. 电子测量技术, 2024, 47(13): 157-166.
- LIU G, YAN SH G, LIU Y, et al. Research on complex road scene object detection based on sensory field enhancement [J]. Electronic Measurement Technology, 2024, 47(13):157-166.
- [32] 许景科, 索祥龙, 周磊. 改进 YOLOv8 的无人机航拍小目标检测算法[J]. 计算机工程与应用, 2025, 61(11): 119-131.
- XU J K, SUO X L, ZHOU L. Improved YOLOv8 drone aerial small target detection algorithm [J]. Journal of Computer Engineering and Applications, 2025, 61(11): 119-131.

作者简介

于承峰, 硕士研究生, 主要研究方向为智能网联汽车感知技术。

E-mail: 2577871042@qq.com

高松, 博士, 教授, 博士生导师, 主要研究方向为新能源汽车和智能网联汽车。

E-mail: 2577871042@qq.com

王鹏伟(通信作者), 博士, 副教授, 硕士生导师, 主要研究方向为自动驾驶汽车运动与决策方法。

E-mail: wpwk16@163.com

孙宾宾, 博士, 副教授, 硕士生导师, 主要研究方向为新能源汽车能量管理与工作模式切换协调控制。

张榕, 硕士研究生, 主要研究方向为智能网联汽车感知技术。