文章编号　1004-924X(2025)19-3135-15

复杂天气条件下基于YOLO-CGT的自动驾驶车辆检测

伍锡如[1,2]，郝家琦[1,2*]，赵一波[1,2]，何佳融[1,2]，葛舒雅[1,2]，梁诗意[1,2]，吴思明[1,2]

(1. 桂林电子科技大学 电子工程与自动化学院，广西 桂林 541004；
2. 桂林电子科技大学 智能综合自动化广西高校重点实验室，广西 桂林 541004)

摘要：针对复杂天气条件下检测车辆目标时，因存在目标模糊及遮挡造成车辆检测精度明显下降的现象，提出一种改进YOLOv8的车辆检测算法YOLO-CGT。该算法面向车载摄像头图像输入场景，通过在YOLOv8结构中引入多项改进，显著提升了在复杂环境下的检测稳定性。其中，设计多尺度残差聚合模块替换原有主干网络结构中的C2f结构，用于增强原始信息的利用并减少网络深度带来的梯度消失问题；引入空间聚合模块，融合全局信息提取和局部信息感知；设计轻量级动态检测头，保证检测精度和效率的平衡；引入内最小点距离交并比(Inner-Minimum Points Distance Intersection over Union，Inner-MPDIoU)度量替换传统IoU，以减少目标框重叠问题。在复杂天气条件下的车辆数据集上进行训练和验证后，实验结果显示，该方法的平均检测精度达到81.4%，提升了6.3%，模型参数量为3.259×10^6，计算量为9.7GFLOPs，在精度显著提升的同时保证了模型的轻量部署能力。该研究方法为自动驾驶系统的安全稳定运行提供了有力保障。

关　键　词：自动驾驶；车辆检测；YOLOv8；复杂天气；多尺度特征

中图分类号：TP394.1；TP181　　**文献标识码**：A

doi：10.37188/OPE.20253319.3135　　**CSTR**：32169.14.OPE.20253319.3135

Autonomous vehicle detection in complex weather based on YOLO-CGT

WU Xiru[1,2], HAO Jiaqi[1,2*], ZHAO Yibo[1,2], HE Jiarong[1,2], GE Shuya[1,2],
LIANG Shiyi[1,2], WU Siming[1,2]

(1. *School of Electronic Engineering and Automation, Guilin University of Electronic Technology, Guilin* 541004, *China*;
2. *Key Laboratory of Intelligence Integrated Automation in Guangxi Universities, Guilin University of Electronic Technology, Guilin* 541004, *China*)
* *Corresponding author, E-mail: haojq@mails.guet.edu.cn*

Abstract: To mitigate the pronounced decline in vehicle detection performance caused by object blur and occlusion under adverse weather conditions, an enhanced YOLOv8-based vehicle detection algorithm, designated YOLO-CGT, is proposed. Tailored for vehicle-mounted camera imagery, the algorithm incor-

收稿日期：2025-06-18；**修订日期**：2025-06-27.

基金项目：国家自然科学基金地区科学基金资助项目(No. 62263005)；广西高校人工智能与信息处理重点实验室开放基金重点项目(No. 2024GXZDSY013，No. 2022GXZDSY004)；桂林电子科技大学研究生教育创新计划资助项目(No. 2025YCXS138)

porates multiple enhancements to the YOLOv8 architecture to substantially improve detection robustness in challenging environments. Specifically, a multi-scale residual aggregation module replaces the original C2f module in the backbone network to increase exploitation of raw feature information and to alleviate gradient vanishing associated with greater network depth. A spatial aggregation module is incorporated to integrate global information extraction with local feature perception. Moreover, a lightweight dynamic detection head is developed to balance detection accuracy and computational efficiency. The conventional IoU metric is supplanted by the Inner-Minimum Points Distance Intersection over Union (Inner-MPDIoU) to reduce bounding-box overlap issues. Trained and validated on a vehicle dataset captured under complex weather conditions, the proposed method attains an average detection accuracy of 81.4%-an improvement of 6.3%-with 3.259×10^6 model parameters and a computational cost of 9.7 GFLOPs, demonstrating suitability for lightweight deployment while delivering substantial accuracy gains. These results provide a robust foundation for the safe and reliable operation of autonomous driving systems.

Key words: intelligent driving; vehide detection; YOLOv8; complex weather; multi-scale features

1 引　言

复杂天气是引发交通事故的重要原因之一。雾、雨、雪等极端天气条件会显著降低驾驶员对周围环境的可见性,从而影响驾驶行为的安全性。在此类天气下,驾驶员往往难以及时识别和判断前方的障碍物与动态目标,极易引发碰撞、刮擦等交通事故,导致道路运行效率下降和交通秩序混乱。近年来,基于深度学习的目标检测方法在标准环境下取得了进展,显著提升了检测的准确率与实时性。然而,这些方法通常在理想光照和清晰图像基础上训练,对于解决复杂天气场景下图像退化、模糊、遮挡、多尺度和背景干扰等问题仍存在明显瓶颈。在雾、雨、雪天气中,图像中的水滴、雾霾和光斑干扰会破坏目标轮廓,使模型难以提取稳定特征;同时,部分小目标或远距离车辆在极端天气中更容易被遮挡或背景淹没,现有检测算法在这类情况下稳定性不足。解决复杂天气条件下的交通安全问题与及时检测车辆和障碍物是维护交通安全的重要一环[1]。

随着世界经济的飞速发展,目标检测已经成为社会安全、公共交通和国防军事等各个领域的基础性研究课题之一,广泛应用在机器人导航、航空航天、工业检测、行人追踪和军事应用等领域[2-5]。相较于传统的目标检测方法,基于深度学习的目标检测方法的性能更优越。这种目标检测方法可以分为两大类:两阶段检测算法和单阶段检测算法。两阶段检测算法将目标物候选框的生成和分类分为两个阶段完成,其中典型的算法有:Faster R-CNN[6],Mask R-CNN[7]和 Cascade R-CNN[8]。Ji 等[9]提出了以 Faster R-CNN 为基准的车辆检测模型,采用更好的特征提取、多尺度特征融合和同源性数据增强。针对 Mask R-CNN 网络掩码分支和过多的全连接层会占用大量网络检测时间的问题,石杰等[10]通过去掉掩码分支和多余的全链接层,将头部轻量化区域卷积神经网络引入到 Mask RCNN 中,调整区域建议网络(Region Proposal Network,RPN)中锚点(Anchor)的比例,提升了检测精度和速度。李松江等[11]使用改进的特征金字塔将浅层信息逐层融入深层网络,引入多支路空洞卷积,减少车辆目标检测过程中小目标及遮挡目标的错检、漏检问题。和两阶段检测方法相比,单阶段检测方法将目标检测任务直接视为一个回归问题。在单个网络中同时完成目标的分类和边界框的回归,这种方法不需要额外的候选区域生成过程,具有较高的实时性和精度,适用于复杂场景下的城市车辆检测任务,典型的算法有 SSD 系列[12]和 YOLO 系列[13]算法。陈家栋等[14]提出一种基于改进 SSD 轻量化的车辆检测方法,引入新型轻量化网络 MobileNetV3、重新设计特征层交叉融合以及更改损失函数的方法。王佳琪等[15]提出了一种红外和可见光融合的多模态检测方法 YOLO-MMF,减少了车辆在受光照变化所导致的性能下降问题。Liu 等[16]考虑到交通场景实时移动,提出一种新的同步检测和跟踪网络 YOLO-

3DMM，以端到端的方式同步进行车辆检测和跟踪。胡丹丹等[17]提出一种基于改进 YOLOv5s 的面向自动驾驶场景的道路目标检测算法，利用深度可分离卷积替换部分普通卷积，减少模型的参数以提升检测速度。郭翠霞等[18]提出了一种双模态轻量化目标检测模型，利用差分模态特征融合模块在不增加网络参数的同时合理互补各模态的特征信息，最终提高了检测精度。赵海丽等[19]提出一种适合部署在边缘终端设备上的基于 YOLOv7-tiny 改进的车辆目标检测算法，通过构造深度强力残差卷积块对主干网络进行改进，降低网络的参数量和计算量。寇发荣等[20]提出了一种高效精准的煤矿井下无人运输车辆目标检测模型—YOLOv8-CBD，该模型显著增强在复杂煤矿井下环境中的定位和检测能力。江旺玉等[21]提出了一种结合多尺度特征聚合扩散和边缘信息增强的小目标检测算法 ADE-YOLO，设计了自适应上下文融合模块，利用通道注意力机制自适应地调整不同特征图的贡献，进一步强化多尺度特征的融合效果，使得重要特征在信息融合过程中更加突出。杜宏等[22]设计了一种新的注意力机制模块 WT-PSA，在复杂天气下提高模型对于车辆目标的关注度，改进的 C2f-OD 模块，提升了主干网络提取图片特征信息的能力，展现了在复杂天气、恶劣天气下车辆检测方面的优越性。卜祥涛等[23]针对现有算法存在处理图像模式单一和通用性受限的问题，提出了一种基于暗通道实验的多模式图像去雾增强算法，在图像清晰度、边缘和细节恢复方面取得显著提升。

然而，基于深度学习的车辆检测研究仍存在以下不足：首先，现有模型如 RT-DETR[24]、Faster R-CNN 和 RetinaNet[25]等的参数量较大，模型复杂度高，在实际交通场景中应用受限；其次，复杂天气条件下，图像存在清晰度下降，遮挡严重的特点，模型难以从复杂的背景中准确地识别车辆目标；另外，实际交通场景数据集中存在远处小目标车辆和近处的大目标车辆，导致特征尺度变化较大，产生一些低质量的标注边框，从而加重了低质量样本对锚框回归的影响。针对上述问题，本文选择 YOLOv8 模型作为基准模型，提出一种复杂天气下车辆检测算法 YOLO-CGT。YOLOv8 作为 YOLO 系列的代表，在结构上融合了多种主干优化设计，具备较强的基线性能和可扩展性。特别是其轻量化版本 YOLOv8n，参数量小、计算量低，适用于边缘计算与车载部署场景，为实现复杂天气下的高效车辆检测提供了良好的基础平台。

2　复杂天气下车辆检测网络

YOLOv8 在目标检测任务中的网络结构由 3 个主要部分组成：输入端、主干网络、颈部结构和检测头。通过这些部分的有机结合，YOLOv8 实现了高效的目标检测和图像处理能力。模型结构如图 1 所示。主干网络通过 C2f 模块、可变形

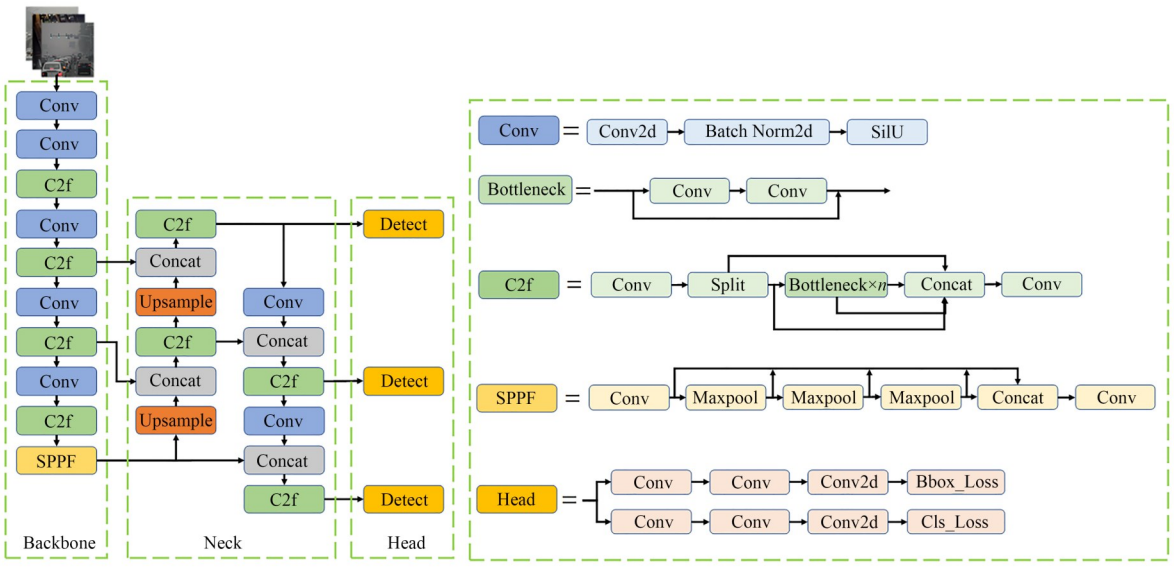

图 1　YOLOv8 网络结构

Fig. 1　Structure of YOLOv8 network

卷积和动态卷积等技术，提升了特征提取和表达能力；颈部结构结合 FPN[26]和 PANet[27]的优点，增强了多尺度特征的融合和信息传递；检测头通过优化的检测头和损失函数，提高了目标定位和分类精度。整体结构的改进，使 YOLOv8在目标检测、图像分割和分类等任务中都表现出了高效性和高精度，特别是在复杂场景和车辆检测上有更好的表现。针对复杂天气下车辆检测任务，提出一种基于 YOLOv8模型的复杂天气下车辆检测模型（YOLO-CGT），结构如图 2所示。主干网络设计 CSP-MSRAM 模块替换原始 YOLOv8的 C2f模块，实现多尺度特征提取；在 Neck 中通过引入引入全局到局部空间聚合（Global-to-Local Spatial Aggregation，GLSA）[28]模块，融合全局信息提取和局部信息感知的结果；在 Head 部分，设计一种轻量级动态检测头，有效减少了模型的参数量；引入 Inner-MPDIoU 损失函数来预测边界框和真实边界框的整体重叠区域。

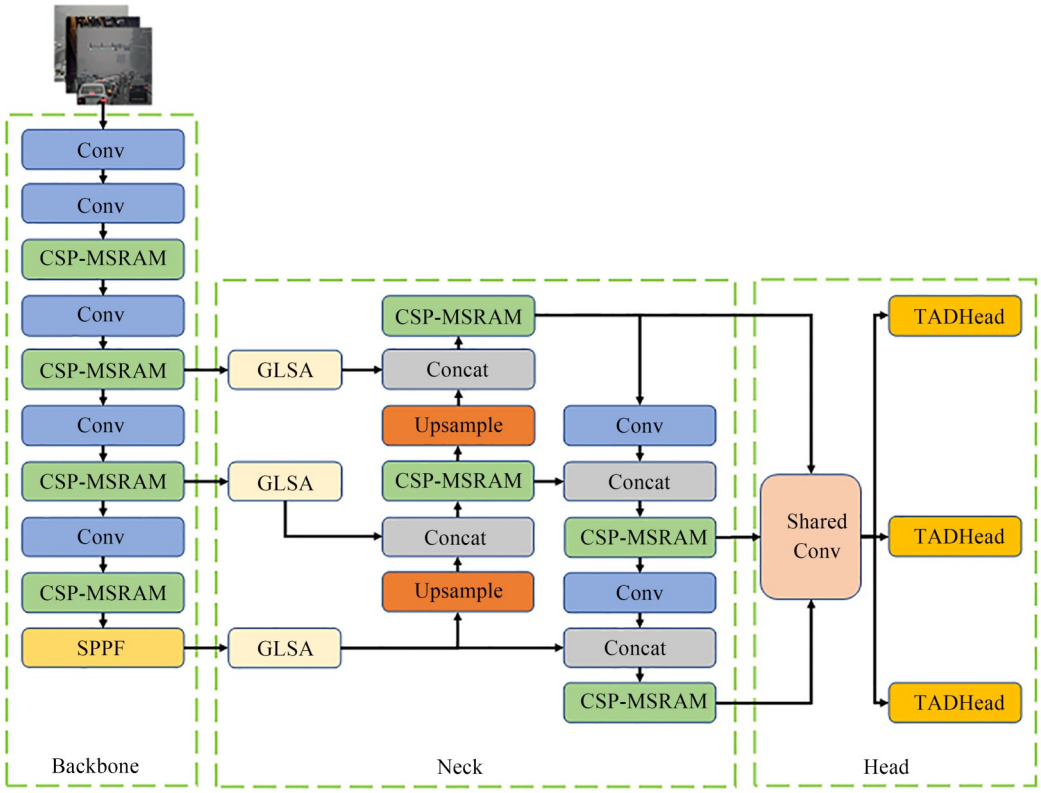

图 2　YOLO-CGT 网络结构

Fig. 2　Structure of YOLO-CGT network

2.1　CSP-MSRAM 模块

在复杂环境下的交通场景中，恶劣天气可能导致图像模糊、低对比度、噪声增加等情况使图像质量下降。为使检测模型可以提升检测精度和可靠性，设计了 CSP-MSRAM 模块代替原网络的 C2f，其结构如图 3所示。C2f模块是一种轻量化的特征提取模块，主要依赖标准卷积进行特征生成，缺少对多尺度信息的显式建模。在处理多尺度目标（如远处的小车与近处的大车）时，可能存在特征表达不足的问题。CSP-MSRAM 模块通过使用多种卷积核大小（3×3，5×5，7×7）进行特征提取，能够有效显示多尺度上下文信息；通过特征拆分和分组卷积，可以有效捕获局部细节和全局特征。对于复杂天气下的目标检测，多尺度特征提取可以更好地适应不同目标大小和形状的变化。为了应对图像噪声和质量下降问题，CSP-MSRAM 模块引入部分残差连接，能够保留输入特征，增强原始信息的利用；在多尺度特征提取的同时减少网络深度带来的梯度消失问题。

CSP-MSRAM 模块通过不同大小的卷积核来捕获多尺度的特征。给定输入特征

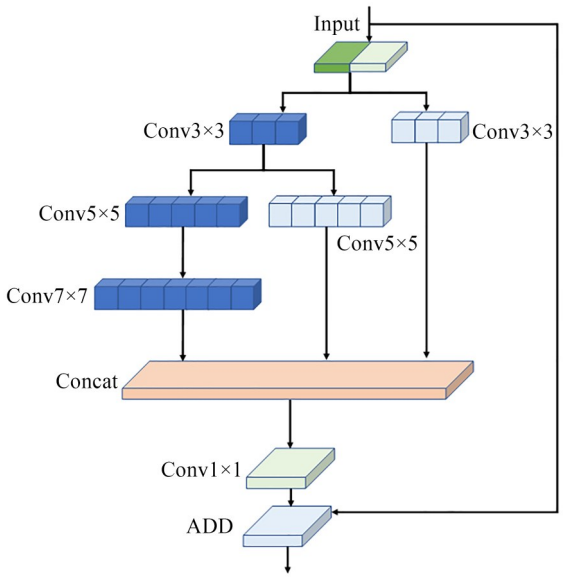

图 3　CSP-MSRAM 结构

Fig. 3　CSP-MSRAM structure

$x \in \mathbf{R}^{C\times H\times W}$，$C$，$H$，$W$ 分别表示通道数、高度和宽度，卷积操作可以表示为：

$$\begin{aligned} f_1 &= \mathrm{Conv}_{3\times 3}(x) \\ f_2 &= \mathrm{Conv}_{5\times 5}(x), \\ f_3 &= \mathrm{Conv}_{7\times 7}(x) \end{aligned} \tag{1}$$

其中：$\mathrm{Conv}_{k\times k}$ 表示核大小为 $k\times k$ 的卷积操作。通过三种卷积操作分别提取不同尺度的特征，以适应复杂天气中不同目标的大小和形状。特征在通道维度上被拆分为不同的部分，进行分组卷积，从而降低计算成本，同时聚焦于局部区域特征。初始特征设为 f_1，拆分操作可以表示为：

$$f_{1,1}, f_{1,2} = \mathrm{Chunk}(f_1, \mathrm{dim}=1, \mathrm{groups}=2). \tag{2}$$

然后在子特征上 $f_{1,1}$ 继续进行分组卷积：

$$f_{2,1} = \mathrm{Conv}_{5\times 5}(f_{1,1}, \mathrm{groups}=1). \tag{3}$$

在多维特征提取完成后，模块通过特征拼接和融合操作将不同路径上的特征聚合。最终的特征来自 3 个路径：

$$f_{\mathrm{out}} = \mathrm{Cat}\left(\left[f_3, f_{2,2}, f_{1,2}\right], \mathrm{dim}=1\right), \tag{4}$$

其中：$\mathrm{Cat}([\cdot])$ 表示在通道维度上的拼接操作；f_3，$f_{2,2}$，$f_{1,2}$ 分别使不同尺度的特征表示。通过引入残差连接，进一步增强梯度流动并保留输入的全局特征。最终输出表示为：

$$y = \mathrm{Conv}_{1\times 1}(f_{\mathrm{out}}) + x. \tag{5}$$

$\mathrm{Conv}_{1\times 1}$ 用于整合多尺度的聚合特征，并匹配输入特征的通道数。残差连接（$+x$）保留了输入的特征信息，使特征学习更加稳定。

2.2　GLSA 模块

在实际交通场景中，往往会出现较多的车辆分散在检测视野中，无法较好地提取全局信息。为了同时捕捉全局和局部空间特征，引入了 GLSA 模块，该模块融合了全局信息提取和局部信息感知的结果。这种双通道设计有效地保留了局部和非局部建模能力有利于分别定位大目标和小目标，减少交通场景图像处理的计算复杂度。GLSA 的网络结构如图 4 所示，F_1 为输入的特征信息，经过 GLSA 操作之后得到新的特征信息，F_1' 的计算过程如下：

全局信息提取（Global Spatial Attention，GSA）强调空间中每个像素的全局关系，作为局部空间注意的补充，有助于模型对整体特征信息

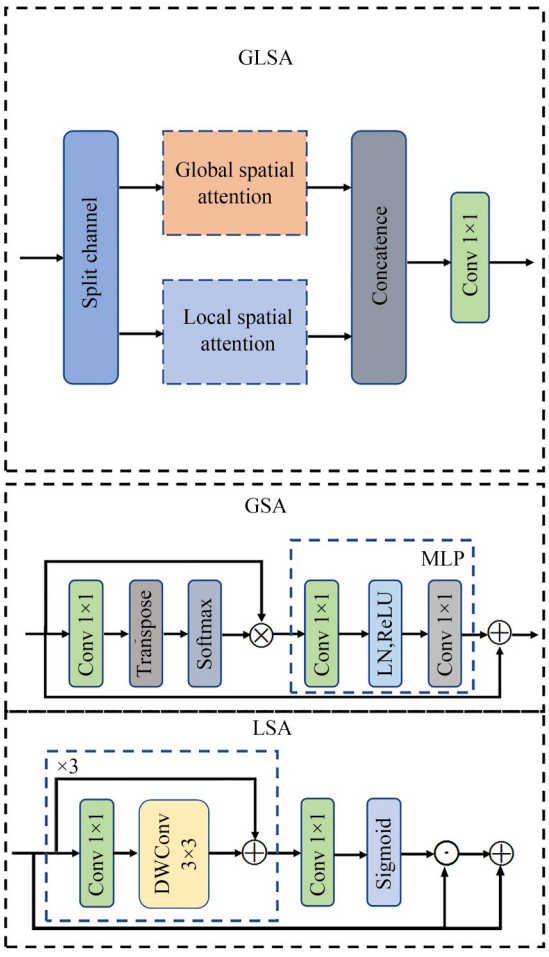

图 4　GLSA 模块

Fig. 4　GLSA structure

的感知。如图4所示，具体公式如下：

$$Att_{\mathrm{G}}(F_i^1)=\mathrm{Soft}_{\max}(\mathrm{Transpose}(C_{1\times1}(F_i^1))),\quad(6)$$

$$G_{\mathrm{sa}}(F_i^1)=\mathrm{MLP}(Att_{\mathrm{G}}(F_i^1)\otimes F_i^1)+F_i^1,\quad(7)$$

其中：F_i^1 是 GSA 的输入，$G_{sa}(F_i^1)$ 是 GSA 的输出。

局部信息感知（Local Spatial Attenion，LSA）模块在给定特征图的空间维度（例如小对象）中有效地提取了感兴趣区域的局部特征。如图4所示，具体公式如下：

$$Att_{\mathrm{G}}(F_i^2)=\partial(C_{1\times1}(F_c(F_i^2)))+F_i^2,\quad(8)$$

$$L_{\mathrm{sa}}(F_i^2)=Att_{\mathrm{L}}(F_i^2)\odot F_i^2+F_i^2,\quad(9)$$

其中：F_i^2 是LSA的输入，$L_{\mathrm{sa}}(F_i^2)$ 是LSA的输出。

GLSA通过关注全局特征捕获全图范围内的上下文信息，并通过捕获局部特征关注小范围内的精细信息。在全局和局部信息感知之后，将全局信息与局部信息进行聚合，并通过1×1卷积和深度可分离卷积降低参数量。通过该双通道设计有效保留了全局与局部的建模功能，模型能够平衡局部信息和局部特征，更好地进行图像检测。

$$F_i^1,F_i^2=\mathrm{Split}(F_i),\quad(10)$$

$$F_i'=C_{1\times1}(\mathrm{Concat}(G_{\mathrm{sa}}(F_i^1),L_{\mathrm{sa}}(F_i^2))).\quad(11)$$

2.3　轻量级检测头

YOLOv8采用解耦头结构，其检测速度、精度和稳定性达难以到自动驾驶的高要求。YOLOv8检测头结构如图5所示。检测头分别通过3×3的卷积和1×1的卷积，进一步由边界框损失函数和分类损失函数输出物体的类别和位置信息。解耦头的设计避免了特征之间的相互干扰，但由于卷积模块产生了堆叠，模型计算量大大增加，不利于实时检测。

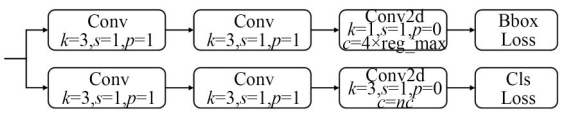

图5　YOLOv8检测头

Fig. 5　YOLOv8 detection head

因此，本文提出了使用共享卷积和动态任务对齐的结构，并引入可变形卷积，以减少模型的参数量，可变形卷积如图6所示。检测头从特征提取器中提取多尺度特征，进而得到联合特征，并生成DCNv2所需要的offset和mask，进一步输出多尺度检测框。为了解决使用共享卷积结构所导致所检测的目标尺度不一致的问题，使用缩放层对特征进行缩放。其结构如图7所示。

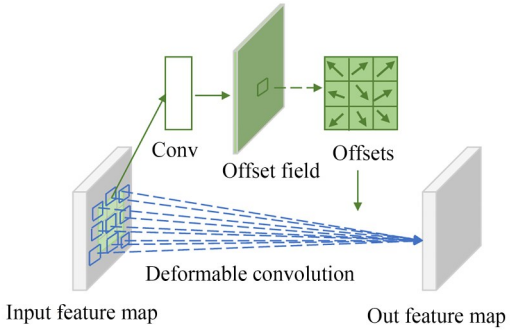

图6　可变形卷积结构

Fig. 6　Deformable convolution structure

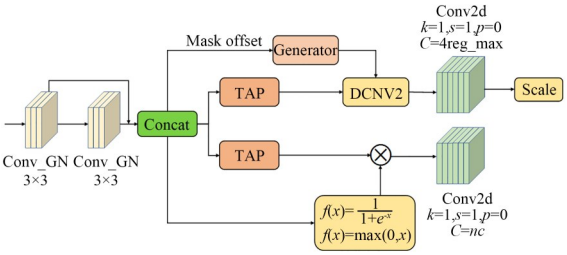

图7　TADHead结构

Fig. 7　Structure of TADHead

在TADHead结构中，共享卷积层由两层卷积操作组成，其特征提取公式为：

$$F_{\mathrm{share}}(x)=\mathrm{Conv2}D(GN(\mathrm{Re}LU(\mathrm{Conv2}D(GN(x))))).\quad(12)$$

输出的特征是堆叠后的共享特征，通过动态任务对齐结构分解为分类特征和回归特征，其公式分别为：

$$F_{\mathrm{cls}}=\mathrm{cls\_decomp}(F_{\mathrm{share}},\mathrm{avg}(F_{\mathrm{share}})),\quad(13)$$

$$F_{\mathrm{reg}}=\mathrm{reg\_decomp}(F_{\mathrm{share}},\mathrm{avg}(F_{\mathrm{share}})),\quad(14)$$

其中：$\mathrm{avg}(F_{\mathrm{share}})$ 是通过全局平均池化得到的共享特征的统计信息，用于特征分解。回归分支进一步利用可变形卷积调整特征图的感受野，公式如下：

$$F_{\mathrm{reg\_aligned}}=DCNv2(F_{\mathrm{reg}},offset,mask),\quad(15)$$

其中：*offset* 为共享卷积模块生成的偏移量；*mask* 为同一模块生成，并通过Sigmoid激活函数调整权重。进一步，分类特征通过两层卷积计算分类

概率权重和生成最终预测，其公式如下：

$$P_{\text{cls}}=\sigma(\text{Conv2}D(\text{Re}LU(\text{Conv2}D(F_{\text{share}})))), \tag{16}$$

$$y_{\text{cls}}=cv3(F_{\text{cls}},P_{\text{cls}}), \tag{17}$$

$$y_{\text{reg}}=\text{Scale}(cv2(F_{\text{reg\_aligned}})), \tag{18}$$

其中：σ表示Sigmoid激活函数，$cv2$，$cv3$为边界框回归和类别预测卷积，Scale适用于尺度调整的模块。

TADHead在设计中通过共享卷积实现了轻量化设计，同时引入可变形卷积解决了共享参数可能导致模型表达能力不足的问题。可变形卷积在训练阶段引入更多的可学习参数，进而弥补轻量化模型后可能带来的精度丢失问题，并且变形卷积可以大大提升参数利用率，并且在推理阶段与普通卷积无差，为模型带来无损的优化方案。通过任务分解增强了分类和回归任务的独立性，结合动态偏移卷积实现特征对齐和感受野优化，最终保证了检测精度和效率的平衡。该设计适用于复杂目标检测任务，尤其是在对模型轻量化和高性能有双重要求的场景中。

2.4　损失函数改进

交并比（Intersection-over-Unions，IoU）损失函数常用于目标检测任务中的回归损失函数，通常预测边界框和真实边界框之间的重叠程度，能够有效衡量模型的匹配程度。其公式如下：

$$L_{\text{IoU}}=1-\frac{|A\cap B|}{|A\cup B|}, \tag{19}$$

其中：A和B分别表示真实框和预测框的面积集合。

在YOLOv8算法中，损失函数由分类损失与回归损失两部分构成。CIoU作为边界框回归损失函数，其结构如图8所示，其表达式为：

$$L_{\text{CIoU}}=1-\text{IoU}+\frac{\rho^2(B_{\text{gt}},B_{\text{pr}})}{(W_{\text{g}}+H_{\text{g}})^2}+\alpha v, \tag{20}$$

$$v=\frac{4}{\pi^2}\left(\arctan\frac{w_{\text{gt}}}{h_{\text{gt}}}-\arctan\frac{w_{\text{pr}}}{h_{\text{pr}}}\right)^2, \tag{21}$$

$$\alpha=\frac{v}{(1-\text{IoU})+v}, \tag{22}$$

式中：$\rho^2(B_{\text{gt}},B_{\text{pr}})$为预测框与真实框中心点的欧氏距离，$B_{\text{gt}}=(x_{\text{gt}},y_{\text{gt}},w_{\text{gt}},h_{\text{gt}})$为真实框的长、宽和中心坐标，$B=(x,y,w,h)$为预测框的长、宽和中心坐标，$v$为纵横比度量函数。

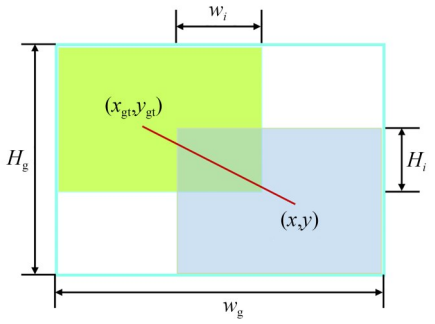

图8　CIoU损失函数

Fig. 8　CIoU loss function

当目标物体的尺度差异较大时，特别是在长宽比差异较大或物体形状不清晰时，基于边界框回归的CIoU损失函数无法充分反映目标检测的真实效果。因此，为解决复杂天气条件下目标不均匀分布、遮挡严重或边界模糊的情况，使用MPDIoU损失函数来替代CIoU损失函数，进一步加强处理复杂背景和不均匀分布的目标时的质量。MPDIoU损失函数可表示为：

$$L_{\text{MPDIoU}}=\frac{A\cap B}{A\cup B}-\frac{d_1^2}{w^2+h^2}-\frac{d_2^2}{w^2+h^2}, \tag{23}$$

$$d_1^2=\left(x_1^B-x_1^A\right)^2+\left(y_1^B-y_1^A\right)^2, \tag{24}$$

$$d_2^2=\left(x_2^B-x_2^A\right)^2+\left(y_2^B-y_1^A\right)^2, \tag{25}$$

式中：(x_1^A,y_1^A)，(x_2^A,y_2^A)分别为真实框左上和右下坐标，(x_1^B,y_1^B)，(x_2^B,y_2^B)分别为预测框左上和右下坐标，d_1^2，d_2^2分别真实框和预测框左上角和右下角间欧氏距离的定位目标。

为了进一步提升模型对目标框内部像素的匹配质量确保预测框的准确性，引入Inner-IoU损失函数与MPDIoU损失函数相结合。Inner-IoU及相关参数如图9所示。通过引入缩放因子比率来控制辅助边界框的比例。其公式为：

$$b_l^{\text{gt}}=x_c^{\text{gt}}-\frac{w^{\text{gt}}s_{\text{f}}}{2},\ b_r^{\text{gt}}=x_c^{\text{gt}}+\frac{w^{\text{gt}}s_{\text{f}}}{2}, \tag{26}$$

$$b_t^{\text{gt}}=y_c^{\text{gt}}-\frac{h^{\text{gt}}s_{\text{f}}}{2},\ b_b^{\text{gt}}=y_c^{\text{gt}}+\frac{h^{\text{gt}}s_{\text{f}}}{2}, \tag{27}$$

$$b_l=x_c-\frac{w\cdot s_{\text{f}}}{2},\ b_r=x_c+\frac{w\cdot s_{\text{f}}}{2}, \tag{28}$$

$$b_t=y_c-\frac{h\cdot s_{\text{f}}}{2},\ b_b=y_c+\frac{h\cdot s_{\text{f}}}{2}, \tag{29}$$

$$inter=(\min(b_r^{\text{gt}},b_r)-\max(b_l^{\text{gt}},b_l))\cdot(\min(b_b^{\text{gt}},b_b)-\max(b_t^{\text{gt}},b_t)), \tag{30}$$

$$union=(w^{gt}h^{gt})(s_f)^2+(w\cdot h)(s_f)^2-inter, \quad (31)$$

$$IoU_{inner}=\frac{inter}{union}, \quad (32)$$

式中：s_f为尺度因子；inter为真实框与预测框的交集；union为真实框与预测框的并集；下标inner表示inner框的相应数据，上标gt表示真实框的相应数据，无上标表示预测框的相应数据。

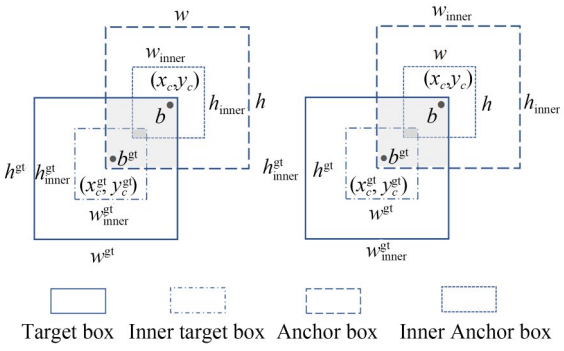

图9　Inner-IoU示意图

Fig. 9　Schematic diagram of Inner-IoU

将Inner-IoU和MPDIoU相结合，得到inner-MPDIoU，如式(33)，既可以通过辅助边框计算IoU，提升泛化能力，也可以解决复杂多目标场景下的边界框重叠问题。

$$L_{inner\text{-}MPDIoU}=L_{MPDIoU}+IoU-IoU_{inner}. \quad (33)$$

3　实验结果和分析

3.1　实验环境和参数

本研究实验环境配置如表1所示，实验参数如表2所示。

3.2　数据集

本文使用的实验数据集为公共数据集RTTS与自建数据集融合构成。RTTS数据集主要包括自然雾天或霾天气候条件下拍摄的真实道路交通场景图像。自建数据集采集主要通过使用高清车载摄像头完成，增加了雾天、雪天和雨天的道路交通图像，一共1 182张图片。合并后数据集包含了不同场景、不同天气情况下道路交通状况的图片，该数据集包含人、自行车、车辆、公交车和摩托车5种类别，一共5 500张图片。图像标注采用LabelImg工具手动完成，标注内容包括人、自行车、车辆、公交车和摩托车的二维边界框。为保证模型训练充分，更好的对模型进行评估，将所有图片按照8∶1∶1的比例划分为训练集、验证集和测试集。其中，训练集4 397张，验证集552张，测试集551张。输入图像的尺寸设定为640×640。

表1　实验环境配置

Tab. 1　Experimental environment configuration

名称	配置名称
操作系统	Windows11
GPU	NVIDIAGeForceRTX4060Ti
CPU	Interi5-12400F
编程语言	Python3.8
深度学习框架	PyTorch2.3.0
加速模块	CUDA12.1

表2　实验参数

Tab. 2　Experimental parameters

参数	参数值
优化器	SGD
学习率	0.01
动量	0.937
权重衰减	0.0005
训练轮数	300
Batch size	16

3.3　评价指标

为验证本文算法的改进效果，实验采用如下评价指标：精确率(Precision，P)、召回率(Recall，R)、(mean Average Precision，mAP)、参数量(Params)、模型计算量(GFLOPs)。

精确率P是指所有被预测为正类的样本中，实际为正类样本的比例，用于评估模型结果的准确性。召回率R指所有实际为正类样本中，被正确预测为正类的比例，反映模型对目标的检测能力。计算公式如下：

$$P=\frac{TP}{TP+FN}, \quad (34)$$

$$R=\frac{TP}{TP+FN}. \quad (35)$$

均值平均精度mAP：计算模型在不同类别上的精确度并进行平均操作，通常使用不同的IOU阈值来表示每个类别的精确度，计算公式

如下：

$$AP=\int_0^1 P(R)\mathrm{d}R, \tag{36}$$

$$mAP=\frac{1}{C}\sum_{i=1}^{C}AP_i. \tag{37}$$

参数量(Params)表示模型中所有可训练参数的总数，用于评估模型的复杂度和存储需求。模型计算量(GFLOPs)是指模型在推理过程中所需的浮点运算次数，用于评估模型的计算效率和推理速度。

3.4　主干感受野可视化

感受野的大小决定了该层网络对输入图像的理解程度，即能够捕捉到多大范围内的特征信息。较大的感受野可以捕捉到更全局的信息，其在模型前向传播时通过中间层的函数提取特征图，然后对该位置的激活值进行反向传播，获取输入图像的梯度，并通过聚合和归一化处理得到贡献度热力图。为了更清晰地展示感受野的范围，通常会采用阈值分析方法，统计达到不同贡献比例(如20%，30%，50%，99%)所对应的最小区域面积，从而量化模型在空间上的感受能力。

通过在主干网络中进行感受野可视化(如图10所示)，可以直观地对比原网络与改进网络在特征提取能力上的差异。可视化结果显示，绿色部分表示感受野范围，相比原网络，改进后网络感受野采样有明显的增强，保持了对关键区域的有效关注，展现出更强的全局特征提取能力。

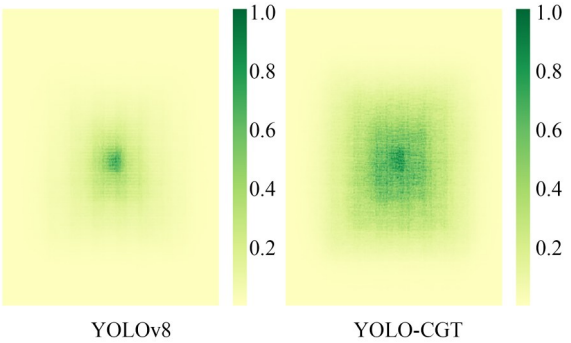

图10　感受野可视化

Fig. 10　Visualization of feel the wild

表3中不同覆盖阈值t下的统计结果表明，改进网络在各梯度阈值条件下感受野面积的占比均高于原网络，表明改进后的模型在提取和集中高贡献特征能力更强，并集中处理关键区域信息，避免对无关背景的过多关注，从而提高检测的精度和效率。

表3　不同阈值下的模型分析

Tab. 3　Model analysis under different thresholds(%)

t	YOLOv8	YOLOv8-CGT
20	2.2	3.9
30	3.9	6.2
50	9.1	12.3
99	70.9	73.6

3.5　损失函数对比实验

为验证不同损失函数对YOLOv8模型的影响，将Inner-IoU与主流的CIoU，DIoU，EIoU，GIoU，SIoU，WIoU和MPDIoU损失函数结合进行实验，结果如表4所示。可以看出，使用Inner-MPDIoU损失函数计算边界损失后模型有更好的精度表现，对复杂天气条件下车辆检测任务有较大的提升。

表4　损失函数对比实验

Tab. 4　Loss function comparison test

损失函数	P	R	mAP@0.5
Inner-CIoU	96.6	87.0	75.5
Inner-DIoU	95.9	86.0	75.6
Inner-EIoU	94.8	**87.6**	75.9
Inner-GIoU	95.7	86.2	74.5
Inner-SIoU	95.8	87.0	75.4
Inner-WIoU	94.7	86.0	75.8
Inner-MPDIoU	**97.4**	86.0	**76.2**

3.6　消融实验

本文算法在YOLOv8算法的基础上采用多个改进策略，为了验证不同模块改动和不同模块组合的改进策略的有效性，设计了8组消融实验进行对比研究。其中，模型1-4分别表示使用CSP-MSRAM模块替换C2f结构、采用GLSA结构、替换TADHead检测头、采用Inner-MPDIoU替换原有边界框损失函数；模型5：同时使用CSP-MSRAM模块和GLSA结构；模型6：在模型5的基础上将检测头替换为TADHead；模型7：本文提出的算法。消融实验结果如表5所示。可以看出，在模型分别对CSP-MSRAM，GLSA，

表 5　消融实验结果

Tab. 5　Ablation experiment results

模　型	CSP-MSRAM	GLSA	TADHead	inner-MPDIoU	mAP@0.5/%	Params/10^6	GFLOPS	FPS
YOLOv8n	×	×	×	×	75.1	3.006	8.1	165.33
Model-1	√	×	×	×	76.4	2.602	7.1	158.65
Model-2	×	√	×	×	76.7	4.017	9.2	152.55
Model-3	×	×	√	×	77.6	2.240	8.6	136.52
Model-4	×	×	×	√	76.2	3.012	8.1	147.62
Model-5	√	√	×	×	78.5	3.906	9.0	134.73
Model-6	√	√	√	×	79.5	3.252	9.7	128.54
Model-7	√	√	√	√	81.4	3.259	9.7	117.44

TADHead，Inner-IoU 设计改进后模型 mAP@0.5分别增加了1.3%，1.6%，2.5%，1.1%，在模型参数量和计算量除引入GLSA模块外，均有不同程度的降低。在同时引入 CSP-MSRAM 和 GLSA 后，模型 mAP@0.5增加了 3.4%，在此基础上将检测头替换为 TADHead 后，模型 mAP@0.5相较于基准模型增加了4.4%，证明了模块之间相互叠加的有效性。进一步替换基准模型的损失函数后，模型 mAP@0.5在参数量和计算量基本不变的情况下增加了 6.3%，模型的检测速度也达到 117 frame/s，证明了模型对于交通场景检测的有效性。

3.7　对比实验

为了进一步验证本文改进算法的有效性，将本文改进的算法与 Faster-RCNN，RT-DETR，RetinaNet 和 YOLO 系列算法在自建数据集上进行对比实验，以模型的准确率、召回率和 mAP 作为评价指标。实验结果如表 6 所示。

从表 6 可以看出，YOLO-CGT 模型的 mAP 指标要优于其他模型，并且在参数量、计算量和检测速度方面都要优于 Faster R-CNN，RT-DETR 和 Retinanet 等模型。与 YOLO 系列算法相比，YOLO-CGT 模型的参数量和计算量相较于同系列小模型仅略有增加，检测速度损失较小，检测精度提高了 7.7%。

表 6　不同算法的对比实验结果

Tab. 6　Comparative test results of different algorithms

损失函数	mAP@0.5/%	mAP@0.5:0.95/%	Params/10^6	GFLOPS	FPS
Faster-RCNN	64.8	39.7	41.14	78.1	16
RT-DETR	75.6	43.6	19.82	57.0	64
RetinaNet	62.3	39.6	36.17	82.2	36
YOLOv5n	74.4	50.9	2.50	7.1	116
YOLOv5s	70.4	48.2	2.72	14.1	132
YOLOv8n	75.1	51.6	3.01	8.1	**165**
YOLOv8s	76.2	51.9	11.1	28.4	165
YOLOv9	73.2	46.5	**2.01**	**7.7**	114
YOLOv10	76.8	49.7	2.69	8.2	121
YOLO-CGT(本文)	**81.4**	**57.9**	3.26	9.7	117

3.8　可视化对比与检测效果分析

3.8.1　热力图可视化

为了进一步清晰地展示改进算法 YOLO-CGT 在复杂天气条件下对车辆检测的效果，本文进行了特征可视化分析，并在图 11 中呈现了改进前后模型的热力图对比结果。如图 11 所示，热力图由 YOLOv8n，YOLOv5n 和改进后的 YOLO-CGT 模型分别基于本文数据集中的复杂天气场景图片进行推理，并使用 Grad-CAM 方法生成。

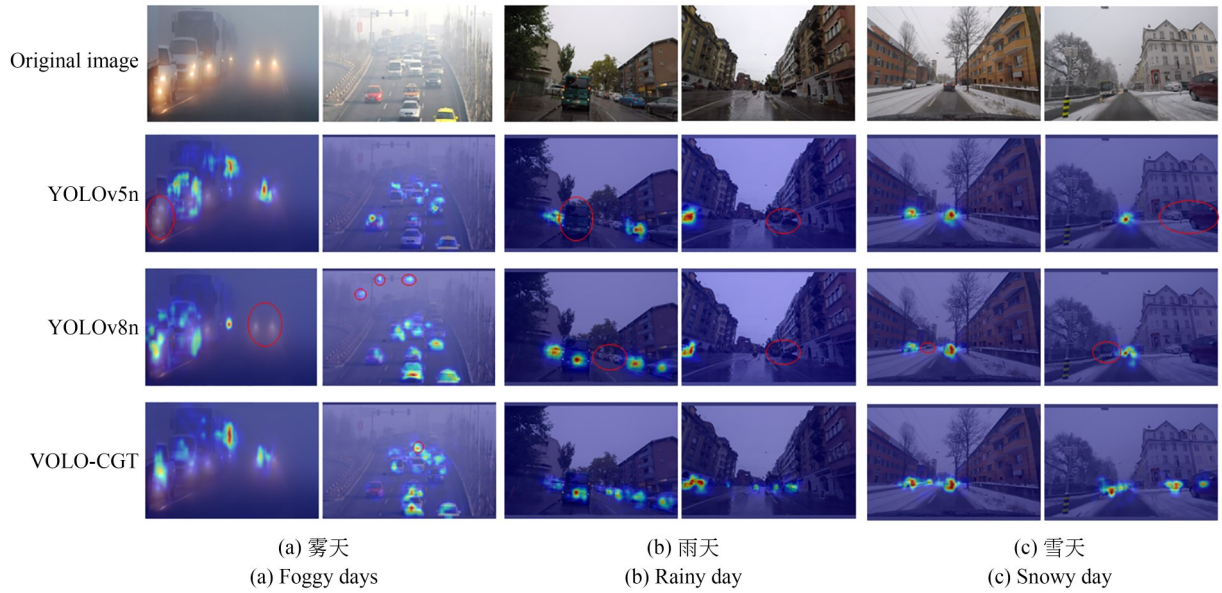

图 11　热力图可视化效果

Fig. 11　Heat map visualization effect

由图 11 可知，YOLOv5n 算法在雾天、雨天和雪天下均出现在距离较近和目标有遮挡情况下聚焦能力较弱的现象，特别在雨天中，第三列图像出现距离较近的目标没有表现出注意力的情况；YOLOv8n 算法在雾天、雨天和雪天下均出现对于远端目标没有体现出较好的注意力的现象，并在第二列图像中对不在检测范围内的目标进行错误的分配注意力；YOLO-CGT 算法更好的关注小目标的提取，减少了在非目标区域的注意力分配，且高亮区域与原始图像中的目标重合度更高，表明本文提出的算法对车辆的聚焦能力更强，对背景噪声有较好的抑制作用。

在图中红色圆圈标注的区域（彩图见期刊电子版）可以观察到，YOLOv5n 和 YOLOv8n 模型的热力响应较弱，存在明显漏检，尤其在图像远端和存在遮挡的情况下表现不佳。而 YOLO-CGT 在相同条件下的热力图更加集中，能够有效从低质量语义区域中挖掘出车辆特征，同时对背景干扰具有良好的抑制能力。整体上，改进后的 YOLO-CGT 算法展现出更强的稳定性与泛化能力，更加适应复杂天气场景下的车辆检测需求。

3.8.2　实验结果分析

为了更直观地对比改进算法的实际效果，将 YOLOv5n 和 YOLOv8n 与改进算法 YOLO-CGT 进行可视化对比，结果如图 12 所示。图中所用图像均来自本文数据集，检测结果基于本文训练模型生成。通过对比可以看出，YOLOv5n 算法在雾天条件下检测效果较好，但在雨天和雪天条件下出现因雨滴、雪花等遮挡物导致检测效果不佳的情况，在雪天条件下对于小目标的检测效果较差，存在因背景噪声导致检测效果下降的问题；YOLOv8n 算法在雾天条件下，对于小目标车辆的检测较差，在雨天和雾天条件下同样存在因遮挡而无法检测的情况，特别在雪天条件下，存在因背景噪声导致的检测效果下降的现象。

为进一步验证改进算法 YOLO-CGT 对多类别目标的检测能力，本文在包含人、自行车、车辆、

图 12　检测效果对比

Fig. 12　Comparison of detection effect

公交车和摩托车五类目标的复杂天气数据集中，分别统计了各类别的检测指标，结果如表 7 所示。图 13 展示了针对上述五类目标的检测可视化结果，不同类别的目标以不同颜色的预测框标注，直观地体现了 YOLO-CGT 在复杂天气条件下对多类别目标的准确定位与识别能力。

表 7　不同类别目标在 YOLO-CGT 算法下的检测效果

Tab. 7　Detection performance of different object categories using YOLO-CGT algorithm

类别	mAP@0.5/%
人	78.6
自行车	68.0
车辆	81.4
公交车	69.1
摩托车	69.2

综上所述，YOLO-CGT 算法通过使用 GLSA 模块，增强了检测网络中全局和局部信息感知的能力，改善了目标因遮挡而无法检测的情况；通过使用多尺度聚合模块，对不同尺度的目标进行特征提取，从而减少了小目标和模糊目标漏检情况的发生；采用可变形卷积[29](Deformable ConvNets，DCNv2)进一步增强多尺度特征提取，显著降低了目标的错检概率。

图 13　不同类别目标在复杂天气条件下的检测可视化结果

Fig. 13　Visualization of detection results for different object categories under complex weather conditions

4　结　论

针对复杂天气条件下，实际道路交通检测存在检测精度不高，实时性差的问题，本文提出了一种改进 YOLOv8 的复杂天气下车辆检测方法 YOLO-CGT。针对图像模糊、低对比度等情况使图像质量下降的问题，设计多尺度残差特征聚合模块替换网络结构中的 C2f 模块。通过使用多种卷积核大小进行特征提取，引入部分残差连接，能够保留输入特征，增强原始信息的利用。

在颈部网络结构中GLSA模块分别定位大目标和小目标,捕捉全局和局部空间特征,减少交通场景图像处理的计算复杂度。为了减少低照度环境下噪声的干扰,提出一种特征共享动态检测头,通过共享卷积实现轻量化设计,同时引入DCNv2解决共享参数可能导致模型表达能力不足的问题。使用Inner-MPDIoU作为边界框损失函数,加强处理复杂背景和不均匀分布目标的质量,并获得更快的收敛速度和更准确的回归结果。实验结果表明,YOLO-CGT算法的mAP@0.5达到了81.4%,模型参数量为3.26×10^6,计算量为9.7GFLOPs,FPS达到了117 frame/s。本文提出的YOLO-CGT算法在保证模型轻量化的前提下具有更高的检测速度和精度,适用于复杂天气条件下的车辆检测。未来的研究将进一步提高模型检测精度,加快车辆检测速度,并将同一模型应用于多类型数据,增强模型的泛化能力,减少模型部署的成本,从而解决实际场景中的难点。

作者贡献声明:

伍锡如:论文构思和项目管理;
郝家琦:实验设计及数据分析,论文撰写;
赵一波:论文审核;
何佳融:实验结果验证;
葛舒雅:测量实验数据分析;
梁诗意:实验结果可视化呈现;
吴思明:研究数据整理和维护。

参考文献:

[1] 宗长富,代昌华,张东. 智能汽车的人机共驾技术研究现状和发展趋势[J]. 中国公路学报,2021,34(6):214-237.
ZONG CH F, DAI CH H, ZHANG D. Human-machine interaction technology of intelligent vehicles: current development trends and future directions [J]. *China Journal of Highway and Transport*, 2021, 34(6): 214-237. (in Chinese)

[2] 范丽丽,赵宏伟,赵浩宇,等. 基于深度卷积神经网络的目标检测研究综述[J]. 光学 精密工程,2020,28(5):1152-1164.
FAN L L, ZHAO H W, ZHAO H Y, *et al*. Survey of target detection based on deep convolutional neural networks[J]. *Optics and Precision Engineering*, 2020, 28(5): 1152-1164. (in Chinese)

[3] 王晶,黄智鑫,徐京邦,等. 面向地下环境机器人的多模态目标检测方法[J]. 机器人,2025,47(4):537-547.
WANG J, HUANG ZH X, XU J B, *et al*. A multimodal object detection method for robots in underground environments [J]. *Robot*, 2025, 47(4): 537-547. (in Chinese)

[4] WANG Z G, ZHAN J, DUAN C G, *et al*. A review of vehicle detection techniques for intelligent vehicles [J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2023, 34(8): 3811-3831.

[5] 王鑫,李鹏程,崔海华,等. 面向机翼线缆支架的装配符合性视觉检测[J]. 光学 精密工程,2025,33(7):1130-1140.
WANG X, LI P CH, CUI H H, *et al*. Visual inspection for assembly conformity of wing cable brackets [J]. *Optics and Precision Engineering*, 2025, 33(7): 1130-1140. (in Chinese)

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

[7] BHARATI P, PRAMANIK A. Deep learning techniques: R-CNN to mask R-CNN: a survey [C]. *Computational Intelligence in Pattern Recognition. Singapore: Springer*, 2020: 657-668.

[8] CAI Z W, VASCONCELOS N. Cascade R-CNN: high quality object detection and instance segmentation [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2021, 43(5): 1483-1498.

[9] JI H, GAO Z, MEI T C, *et al*. Improved faster R-CNN with multiscale feature fusion and homography augmentation for vehicle detection in remote sensing images [J]. *IEEE Geoscience and Remote Sensing Letters*, 2019, 16(11): 1761-1765.

[10] 石杰,周亚丽,张奇志. 基于改进Mask RCNN和Kinect的服务机器人物品识别系统[J]. 仪器仪表学报,2019,40(4):216-228.
SHI J, ZHOU Y L, ZHANG Q ZH. Service robot item recognition system based on improved

Mask RCNN and Kinect[J]. *Chinese Journal of Scientific Instrument*, 2019, 40(4): 216-228. (in Chinese)

[11] 李松江, 吴宁, 王鹏, 等. 基于改进Cascade RCNN的车辆目标检测方法[J]. 计算机工程与应用, 2021, 57(5): 123-130.
LI S J, WU N, WANG P, *et al*. Vehicle target detection method based on improved cascade RCNN[J]. *Computer Engineering and Applications*, 2021, 57(5): 123-130. (in Chinese)

[12] 蒋弘毅, 王永娟, 康锦煜. 目标检测模型及其优化方法综述[J]. 自动化学报, 2021, 47(6): 1232-1255.
JIANG H Y, WANG Y J, KANG J Y. A survey of object detection models and its optimization methods[J]. *Acta Automatica Sinica*, 2021, 47(6): 1232-1255. (in Chinese)

[13] 米增, 连哲. 面向通用目标检测的YOLO方法研究综述[J]. 计算机工程与应用, 2024, 60(21): 38-54.
MI Z, LIAN ZH. Review of YOLO methods for universal object detection[J]. *Computer Engineering and Applications*, 2024, 60(21): 38-54. (in Chinese)

[14] 陈家栋, 雷斌. 基于改进SSD轻量化的交通路口目标检测[J]. 传感器与微系统, 2022, 41(10): 117-121, 126.
CHEN J D, LEI B. Traffic intersection target detection based on improved SSD lightweight[J]. *Transducer and Microsystem Technologies*, 2022, 41(10): 117-121, 126. (in Chinese)

[15] 王佳琪, 张淇, 黄巍. 不同光照下多模态注意力融合的车辆检测[J]. 计算机工程与应用, 2024, 60(16): 116-123.
WANG J Q, ZHANG Q, HUANG W. Vehicle detection of multi-modal attention fusion under different illumination[J]. *Computer Engineering and Applications*, 2024, 60(16): 116-123. (in Chinese)

[16] LIU L C, SONG X Y, SONG H S, *et al*. Yolo-3DMM for simultaneous multiple object detection and tracking in traffic scenarios[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2024, 25(8): 9467-9481.

[17] 胡丹丹, 张忠婷. 基于改进YOLOv5s的面向自动驾驶场景的道路目标检测算法[J]. 智能系统学报, 2024, 19(3): 653-660.
HU D D, ZHANG ZH T. Road target detection algorithm for autonomous driving scenarios based on improved YOLOv5s[J]. *CAAI Transactions on Intelligent Systems*, 2024, 19(3): 653-660. (in Chinese)

[18] 郭翠霞, 徐永涛, 邹章煌, 等. 基于可见和红外双模态融合的轻量化行人车辆检测算法[J]. 光子学报, 2025, 54(6): 153-166.
GUO C X, XU Y T, ZOU ZH H, *et al*. Lightweight pedestrian vehicle detection algorithm based on visible and infrared bimodal fusion[J]. *Acta Photonica Sinica*, 2025, 54(6): 153-166. (in Chinese)

[19] 赵海丽, 许修常, 潘宇航. 基于改进YOLOv7-tiny的车辆目标检测算法[J]. 兵工学报, 2025, 46(4): 103-113.
ZHAO H L, XU X CH, PAN Y H. Vehicle target detection algorithm based on improved YOLOv7-tiny[J]. *Acta Armamentarii*, 2025, 46(4): 103-113. (in Chinese)

[20] 寇发荣, 吕庚毅, 谢伟华, 等. 基于改进YOLOv8的煤矿井下无人运输车辆目标检测研究[J/OL]. 煤炭学报, 2025: 1-13. (2025-05-20). https://link.cnki.net/doi/10.13225/j.cnki.jccs.2025.0194.
KOU F R, LÜ G Y, XIE W H, *et al*. Research on target detection of unmanned vehicles in coal mines based on improved YOLOv8[J/OL]. *Journal of China Coal Society*, 2025: 1-13. (2025-05-20). https://link.cnki.net/doi/10.13225/j.cnki.jccs.2025.0194. (in Chinese)

[21] 江旺玉, 王乐, 姚叶鹏, 等. 多尺度特征聚合扩散和边缘信息增强的小目标检测算法[J]. 计算机工程与应用, 2025, 61(7): 105-116.
JIANG W Y, WANG L, YAO Y P, *et al*. Multi-scale feature aggregation diffusion and edge information enhancement small object detection algorithm[J]. *Computer Engineering and Applications*, 2025, 61(7): 105-116. (in Chinese)

[22] 杜宏, 顾宸瑜, 张孝峥, 等. 基于YOLOv10-vehicle算法的复杂天气情况下车辆目标检测方法[J/OL]. 吉林大学学报(工学版), 2025: 1-10. (2025-01-16). https://link.cnki.net/doi/10.13229/j.cnki.jdxbgxb.20250016.
DU H, GU CH Y, ZHANG X ZH, *et al*. Vehicle target detection method based on the YOLOv10-vehicle algorithm under complex weather conditions[J/OL]. *Journal of Jilin University* (*Engineering and Technology Edition*), 2025: 1-10. (2025-01-

16). https://link.cnki.net/doi/10.13229/j.cnki.jdxbgxb.20250016.(in Chinese)

[23] 卜祥涛, 宋亚芳, 王晓宇, 等. 面向多模式图像的改进暗通道先验去雾增强[J]. 光学 精密工程, 2025, 33(13): 2124-2135.
BU X T, SONG Y F, WANG X Y, *et al*. Improved dark channel prior dehazing enhancement for multi-modal images[J]. *Optics and Precision Engineering*, 2025, 33(13): 2124-2135. (in Chinese)

[24] ZHAO Y A, LV W Y, XU S L, *et al*. DETRs beat YOLOs on real-time object detection[C]. 2024 *IEEE/CVF Conference on Computer Vision and Pattern Recognition* (*CVPR*). *June* 16-22, 2024, *Seattle*, *WA*, *USA*. IEEE, 2024: 16965-16974.

[25] LIN T Y, GOYAL P, GIRSHICK R, *et al*. Focal loss for dense object detection[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2020, 42(2): 318-327.

[26] LIN T Y, DOLLÁR P, GIRSHICK R, *et al*. Feature pyramid networks for object detection[C]. 2017 *IEEE Conference on Computer Vision and Pattern Recognition* (*CVPR*). *July* 21-26, 2017, *Honolulu*, *HI*, *USA*. IEEE, 2017: 936-944.

[27] PIAO Y R, JIANG Y Y, ZHANG M, *et al*. PANet: patch-aware network for light field salient object detection[J]. *IEEE Transactions on Cybernetics*, 2023, 53(1): 379-391.

[28] TANG F L, XU Z X, HUANG Q M, *et al*. *DuAT*: *Dual-aggregation Transformer Network for Medical Image Segmentation*[M]. Pattern Recognition and Computer Vision. Singapore: Springer Nature Singapore, 2023: 343-356.

[29] ZHU X Z, HU H, LIN S, *et al*. Deformable ConvNets V2: more deformable, better results[C]. 2019 *IEEE/CVF Conference on Computer Vision and Pattern Recognition* (*CVPR*). *June* 15-20, 2019, *Long Beach*, *CA*, *USA*. IEEE, 2019: 9300-9308.

作者简介:

伍锡如(1981—),男,湖南新化人,博士,教授,博士生导师,2007年、2012年于湖南大学分别获得硕士和博士学位,主要从事机器人视觉与环境感知、多机器人的协调与编队控制、深度学习与模式识别方面的研究。E-mail: xiru@guet.edu.cn

通讯作者:

郝家琦(2001—),男,山西长治人,硕士研究生,2023年于华北科技学院获得学士学位,主要研究方向为深度学习、目标检测、图像处理和计算机视觉方面的研究。E-mail: haojq@mails.guet.edu.cn