基于改进 YOLOv8n 的雾天行人车辆检测算法

安 旭[1] 胡蓉华[1]
AN Xu HU Ronghua

摘 要 针对雾天场景下，行人车辆检测算法存在检测精度低、参数量大的问题，文章提出了一种基于 YOLOv8n 改进的轻量化检测算法 YOLOv8n-TPDM。首先，在传统 C2f 模块中引入三重注意力机制，强化模型对空间和通道的关系理解和复杂数据的处理能力；同时在检测头位置增加小目标检测层，解决模型对小目标易漏检问题；将原始损失函数 CIoU 替换为 MPDIoU，提升锚框的精度，并加快收敛速度；最后，采用深度可分离卷积降低模型的参数量和计算量，并在雾天数据集 RTTS 上完成训练和验证。实验结果表明，与基准模型 YOLOv8n 相比，改进后的算法平均精度提高了 2.1%，模型参数量减少约 55.5%，能够更精准地完成雾天目标检测任务。

关键词 YOLOv8；轻量化；三重注意力机制；MPDIoU；检测头；深度可分离卷积

doi：10.3969/j.issn.1672-9528.2026.05.008

0 引言

当前行人车辆检测在智能交通以及自动驾驶等方面的应用十分广泛。但在雾天环境下，光照强度低，相机拍摄图片对比度下降、辨识度较低，算法对行人车辆检测的准确性下降，进而容易引发安全事故。因此，针对雾天交通图像的目标检测研究，对精确感知道路交通信息，减少交通事故发生，具有重要的现实意义 [1-3]。

随着技术不断发展，基于深度学习的行人车辆检测算法取得了诸多进展，在大规模数据集上训练达到较好的泛化能力。在雾天环境下，进一步提升目标检测精度仍具挑战，赖镜安等 [4] 提出基于 YOLOv5 轻量级雾天目标检测算法，采用感受野注意力模块，并使用 Slimneck 网络优化颈部结构提升准确率；Li 等 [5] 基于 YOLOv3 设计了一种图像处理模块，虽然在检测精度上有所提升，但计算量也有一定的增长；苏彤等 [6] 在 YOLOv5 的基础上，提出了 YOLOv5-SGE 网络，为提升模型的特征提取能力而在相应模块中引入了 SimAM 三维注意力机制，使用 GSConv 卷积替换标准卷积，减少参数量，实现模型的轻量化，最后替换传统损失函数为 EIoU，

1. 长江大学计算机科学学院 湖北荆州 434100

[5] 毛星亮 , 陈晓红 , 宁肯 , 等 . 全局和局部信息融合的案情关键要素识别 [J]. 软件学报 , 2023, 34(12):5724-5736.
[6]Lu Rui , Li Linying .Named entity recognition method of Chinese legal documents based on parallel instance query network[J].International Journal of Digital Crime and Forensics (IJDCF), 2024, 16(16):1-19.
[7] 杨长春 , 严鑫杰 , 顾晓清 , 等 . 融合字形特征的法律文书命名实体识别方法 [J]. 中文信息学报 , 2025, 39(9):91-99.
[8]Mao Xingliang, Jiang Jie, Zeng Yongzhe, et al. Generative named entity recognition framework for Chinese legal domain[J]. PeerJ Computer Science, 2024, 10: e2428.
[9] 张天宇 , 孙媛媛 , 杜文玉 , 等 . 基于语义边界增强的司法命名实体识别 [J]. 清华大学学报 (自然科学版), 2024, 64(5): 749-759.
[10] 彭晗 , 阮日青 , 胡颖 , 等 .JURIS: 基于理解增强型指令微调的司法命名实体识别方法 [J]. 电子学报 , 2025, 53(9): 3117-3133.

【作者简介】

李林瑛（1975—），男，辽宁大连人，博士，教授，研究方向：自然语言处理。

曲云平（1999—），男，辽宁大连人，硕士，研究方向：自然语言处理。

付子轩（2001—），女，辽宁鞍山人，硕士研究生，研究方向：自然语言处理。

刘美汐（2002—），女，辽宁沈阳人，硕士研究生，研究方向：自然语言处理。

（收稿日期：2026-01-21 修回日期：2026-05-06）

加快模型的收敛速度。

以上算法虽在一定程度上解决了雾天环境下算法检测精度低的问题，但是模型的参数量与计算量均有所增长，难以同时平衡精度和参数量。为了解决此类问题，本文提出了一种基于 YOLOv8n 改进的目标检测算法。本文的主要模型改进点包括以下几点：

（1）在骨干网络（Backbone）中的 C2f 模块融合三重注意力机制，增强关键特征的融合能力；

（2）将传统 YOLOv8 中的传统损失函数替换为 MPDIoU，加快模型收敛速度；

（3）针对小目标漏检的问题，增加小目标检测层，提升模型对小目标检测能力；

（4）深度可分离卷积替换标准卷积，减少模型参数量和计算量，在保持模型检测精度的同时，有效提高模型检测速率。

1 改进的 YOLOv8n 模型框架

YOLOv8 是继 YOLOv5 之后迭代的一个重大版本，秉持 YOLO 系列特点。YOLOv8 分别有五个不同尺寸的版本，从大至小依次为 x、l、m、s、n，相应地，模型参数量递减，精度也不断递减。本研究在 YOLOv8[7-11] 参数最小的版本的基础上进行改进，提出了适应于雾天行人车辆检测算法 YOLOv8n-TPMD，其模型结构如图 1 所示。

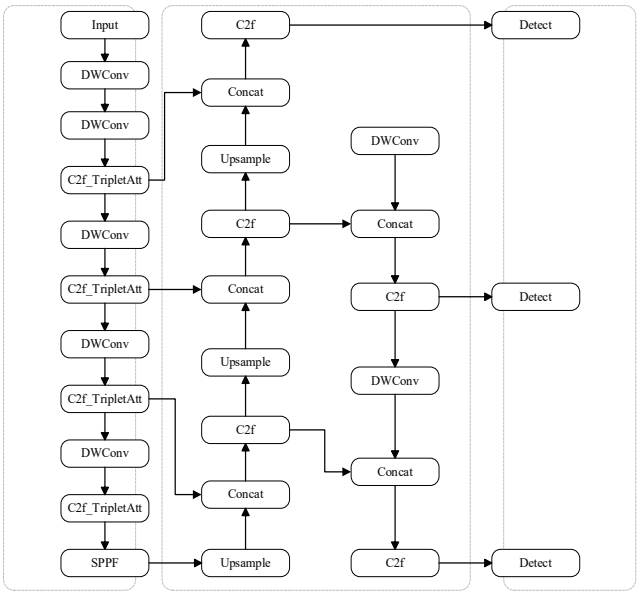

图 1 YOLO-TPDM 模型结构

1.1 三重注意力机制

在雾天环境下，低分辨率使目标信息的特征无法表达充分，致使检测目标准确率下降。故本文引入注意力机制，传统注意力机制通常只能单方面地处理通道维度或空间维度，从而丢失了跨维度信息，而本研究采用的三重注意力机制（Triplet Attention）[12-13] 能够捕获特征图在通道、宽度、高度三个维度间的跨维度交互依赖，从而提升模型对复杂特征的感知能力。同时结合旋转操作和残差变换，将通道与空间维度动态关联，以实现跨维交互，弥补传统注意力的局限性。此外，其参数量较少，计算开销可忽略不计。其结构如图 2 所示。

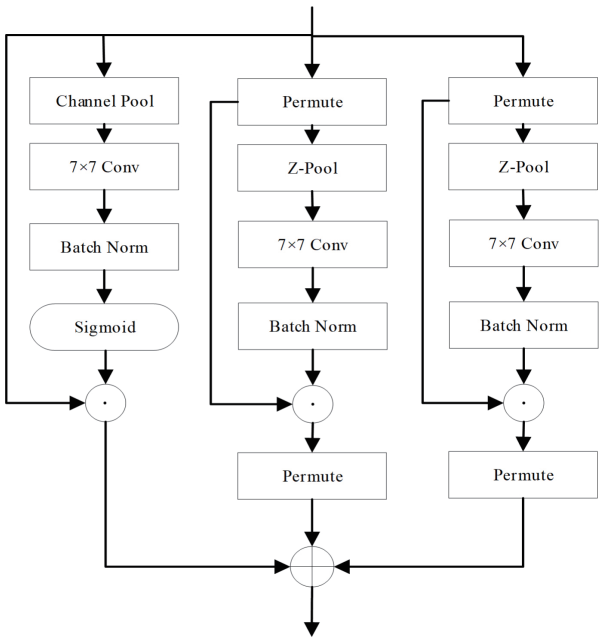

图 2 Triplet Attention 模块结构图

第一个分支为高度维度与通道维度之间建立的交互；第二个分支为宽度维度与通道维度之间建立的交互；第三分支为高度维度与宽度维度之间建立的交互。最后三分支细化张量由平均运算聚合得到一个特征图。用公式表示为：

$$y=\frac{1}{3}(A_1\omega_1+\overline{A_2\omega_2}+\overline{A_3\omega_3}) \tag{1}$$

其中，将该维上的最大汇集特征和平均汇集特征连接到 Z-Pool 层，使通道维度的张量缩减到二维。以便此层增强表征能力和保留真实张量的丰富表示，同时缩小其深度以进一步降低计算量。用公式表示为：

$$Z\text{-Pool}(\chi)=[\text{MaxPool}_{0d}(\chi),\text{AvgPool}_{0d}(\chi)] \tag{2}$$

1.2 增加小目标检测层

在 YOLOv8n 的头部结构，采用三个检测头设计，处理图像中不同尺度目标，但是针对复杂场景中小目标，模型受限，容易漏检。在特征提取过程中，浅层网络能获取丰富的局部细节信息，因其感受野较小且特征图分辨率较高，因而具备较强的小目标细节捕捉能力。而深层网络随层数加深，感受野逐渐扩大，虽能获得更具全局性的语义信息，但其特征图分辨率降低，主要包含中等到大尺度目标的特征。尽管该结构可应对多尺度目标检测，但在小目标识别方面的表现

仍显不足。为提升对小尺寸目标的检测性能，添加一个针对小目标检测层 P2，同时去除最后的大目标检测层 P5，在保持三个检测层基础上，减少模型参数量，虽在一定程度上提升了精度，但在计算成本上不可避免地有所增加。

1.3 损失函数改进

YOLOv8n 中原有的损失函数为 CIoU，对高宽比的定义较为模糊，在某些情况下容易无法发挥作用，导致质量较差的回归样本对损失的影响较大，而对高质量样本回归优化困难。在此将其替换为 MPDIoU，该损失函数利用几何特点仅作两点间距离比较，简化了两个边界框之间的相似性比较。针对相同宽高比不同尺寸的失效场景，非重叠区域的梯度消失，计算复杂度简化。针对模糊目标检测，点间距对模糊边缘更具有鲁棒性，收敛速度因梯度计算更快，如式（3）～（5）所示。

$$\mathrm{MPDIoU_{Loss}} = 1 - \mathrm{IoU} + \frac{d_1^2}{h^2+w^2} + \frac{d_2^2}{h^2+w^2} \tag{3}$$

$$d_1^2 = (x_1^{\mathrm{pred}} + x_1^{\mathrm{gt}})^2 + (y_1^{\mathrm{pred}} + y_1^{\mathrm{gt}})^2 \tag{4}$$

$$d_2^2 = (x_2^{\mathrm{pred}} + x_2^{\mathrm{gt}})^2 + (y_2^{\mathrm{pred}} + y_2^{\mathrm{gt}})^2 \tag{5}$$

式中：$(x_1^{\mathrm{pred}} + y_1^{\mathrm{pred}})$表示预测框左上角，$(x_2^{\mathrm{pred}} + y_2^{\mathrm{pred}})$表示预测框右下角，$(x_1^{\mathrm{gt}} + y_1^{\mathrm{gt}})$和$(x_2^{\mathrm{gt}} + y_2^{\mathrm{gt}})$则分别表示真实框的左上角和右下角。

1.4 深度可分离卷积

标准卷积兼顾通道特征和空间特征两个方面，利用多个通道卷积核对输入的多通道图像进行处理，这样计算量偏大，效率低，容易造成资源浪费。为此，在本研究算法中引入深度可分离卷积（depthwise separable convolution, DW），该卷积是一种高效的卷积操作，被广泛应用于轻量级神经网络中，旨在减少模型的计算量和参数量，同时保持模型的性能与表达能力。相较于标准卷积，深度可分离卷积的具体实现过程可分解为两个步骤：深度卷积和逐点卷积，其中，深度卷积独立地对输入数据的每个通道进行卷积操作，不融合跨通道信息，生成的特征图通道数与输入通道数保持一致；逐点卷积则是融合通道特征，使用 1×1 卷积核跨通道加权组合，输出通道数由卷积核数量决定，参数量为 1×1×M×N，其中 N 为输出通道数。

假定 D_K×D_K 为卷积核大小，M 为输入通道，N 为输出通道，D_F×D_F 为输出特征图的尺寸。对于标准卷积，用公式表示为：

$$F(\mathrm{Conv}) = D_k \times D_k \times M \times N \times D_F \times D_F \tag{6}$$

深度可分离卷积的计算公式为：

$$F(\mathrm{DWC}) = D_k \times D_k \times M \times D_F \times D_F + M \times N \times D_F \times D_F \tag{7}$$

由上述两式相比较可知，深度可分离卷积在计算量和参数量上比标准卷积的更小，故以此来使模型更加轻量化。

2 实验结果与分析

2.1 数据集

本文所采用的数据集为雾天真实数据集 RTTS，共有 4 322 张图片，该数据集标注了五个类别：行人、汽车、公交车、自行车、摩托车，在实验之前，将数据集划分为两个部分：训练集、验证集，比例为 8:2。

2.2 实验环境

本文的实验配置：CPU 型号为 12490F，GPU 型号为 RTX4060Ti，操作系统为 Windows11 专业版。训练环境为 Python3.10.15 版本，深度学习框架为 PyTorch2.4.1，选用 YOLOv8n 版本为 8.2.78，训练过程采用随机梯度下降优化器，初始学习率（lr0）设为 0.01，动量因子 0.937，权重衰退（weight decay）设置为 0.000 5，训练轮次（epochs）设置为 200 epoch，批量大小（batch size）设置为 16，图像分辨率设置为 640 px×640 px，后续训练过程参数均保持一致。

2.3 评价指标

本文实验的评价指标包括平均精确率（mean average precision, mAP）、模型参数量（Params）、计算量（GFLOPs）。其中，mAP 为所有类的检测精度的平均值，反映了模型的综合性能，模型计算量反映了它的计算成本，mAP 相关计算公式如（8）～（11）所示。

$$\mathrm{Precision} = \frac{\mathrm{TP}}{\mathrm{TP+FP}} \tag{8}$$

$$\mathrm{Recall} = \frac{\mathrm{TP}}{\mathrm{TP+FN}} \tag{9}$$

$$\mathrm{AP} = \int_0^1 P(R)\mathrm{d}R \tag{10}$$

$$\mathrm{mAP} = \frac{1}{n}\sum_{i=1}^{n} \mathrm{AP_i} \tag{11}$$

式中：TP 是正确样本预测正确的数量，FP 是错误样本预测为正确的数量，FN 是正类预测为负类的数量，P 是准确率，R 是召回率，AP 是某一类别的识别平均准确率，n 是总类别数。从式中可知，mAP 可以综合评价模型性能，说明最终的检测效果。

2.4 消融实验

为了验证本文的改进方法对模型的有效性，设计了相应的消融实验，分别为改进 1、改进 2、改进 3、YOLOv8n-TPMD 四个模型实验，改进 1 将骨干网络中 C2f 模块替换为 C2f_TripletAtt 模块；改进 2 添加一个小目标检测层，并去除最后一层大目标检测层；改进 3 在此前改进的基础上，替换传统损失函数为 MPDIoU 损失函数；最后一个实验是本文融合四个改进点后的算法，结果如表 1 所示。

表 1 YOLOv8n-TPMD 消融试验结果

改进	方法	mAP/%	参数量 /10^6	计算量 /10^9
基础模型	YOLOv8n	72.2	3.007	8.1
改进 1	YOLOv8n+C2f_TripletAtt	72.4	2.766	7.5
改进 2	YOLOv8n+ C2f_TripletAtt+p2	74.1	1.770	10.9
改进 3 YOLO-TPMD	YOLOv8n+ C2f_ TripletAtt+p2+MPDIoU	73.9	1.770	9.7
	YOLOv8n+ C2f_ TripletAtt+p2+DW+MPDIoU	74.3	1.337	9.7

实验结果表明：基准模型 YOLOv8n 的平均精度为 72.2%，参数量为 3.007×10^6，计算量为 8.1×10^9。而本文的 YOLOv8n-TPMD 相较于基准模型 YOLOv8n，mAP 提升了 2.1%，参数量减少了 55.5%，模型权重减少 50.7%，计算量略有上升。通过以上消融实验的综合评估，验证了本文算法在基准模型的基础上对雾天行人车辆检测的改进方法的有效性与可行性。

3 结语

在雾天环境下，算法检测精度大幅下降，故本文在现有的 YOLOv8n 基础上做出改进，提出改进后 YOLOv8n-TPDM 算法，通过对头部网络的优化、骨干网络中 C2f 模块和标准卷积改进以及损失函数的优化，实现了雾天行人车辆检测的准确性提升，从而将其用于雾天环境检测任务。模型进行了轻量化处理，使其更易于部署在移动端。最终通过验证得出，本文通过改进的 YOLOv8n-TPDM 算法可以有效解决雾天对检测目标精度的影响，满足了雾天行人车辆检测的需求。未来的研究将考虑进一步优化算法，扩充数据集以提升模型的泛化能力。

参考文献：

[1]Lin T Y, Goyal P, Girshick R, et al. Focal loss for dense object detection[C]//2017 IEEE International Conference on Computer Vision (ICCV).Piscataway:IEEE,2017:2980-2988.

[2] 范佳琦 , 李鑫 , 霍天娇 , 等 . 基于单阶段算法的智能汽车跨域检测研究 [J]. 中国公路学报 ,2022,35(3):249-262.

[3] 王娟敏 , 皮建勇 , 黄昆 , 等 . 面向复杂场景的多尺度行人和车辆检测算法 [J]. 现代电子技术 ,2025,48(9):143-153.

[4] 赖镜安 , 陈紫强 , 孙宗威 , 等 . 基于 YOLOv5 的轻量级雾天目标检测方法 [J]. 计算机工程与应用 ,2024,60(6):78-88.

[5]Li Boyi, Peng Xiulian, Wang Zhangyang, et al. AOD-Net:all-in-one dehazing network[C]//2017 IEEE International Conference on Computer Vision (ICCV).Piscataway: IEEE, 2017: 4770-4778.

[6] 苏彤 , 王颖 , 邓启扬 , 等 . 基于 YOLOv5 改进的雾天行人与车辆检测算法 [J]. 系统仿真学报 , 2024,36(10):2413-2422.

[7]Redmon J, Divvala S, Girshick R,et al.You only look once:unified,real-time object detection[C]// 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE,2016:779-788.

[8]Bochkovskiy A, Wang C Y, Liao H Y M. YOLOv4: optimal speed and accuracy of object detection[PP/OL].(2020-04-23)[2026-05-08].https://doi.org/10.48550/arXiv.2004.10934.

[9]Ge Zheng, Liu Songtao, Wang Feng, et al. YOLOx: exceeding YOLO series in 2021[PP/OL].(2021-08-06)[2026-05-08]. https://doi.org/10.48550/arXiv.2107.08430.

[10]Li Chuyi, Li Lulu, Jiang Hongliang, et al. YOLOv6: a single-stage object detection framework for industrial applications[PP/OL].(2022-09-07)[2026-05-08].https://doi.org/10.48550/arXiv.2209.02976.

[11]Xu Xianzhe, Jiang Yiqi, Chen Weihua, et al. DAMO-YOLO: a report on real-time object detection design[PP/OL].(2023-04-24)[2026-05-08].https://doi.org/10.48550/arXiv.2211.15444.

[12]Zhang Pan, Deng Hongwei, Chen Zhong. RT-YOLO:a residual feature fusion triple attention network for aerial image target detection[J].Computers, Materials & Continua, 2023, 75(1): 1411-1430.

[13] Misra D, Nalamada T, Arasanipalai A U, et al. Rotate to attend: convolutional triplet attention module[C/OL]//2021 IEEE Winter Conference on Applications of Computer Vision (WACV). Piscataway: IEEE, 2021[2026-05-08]. https://ieeexplore.ieee.org/document/9423300.DOI: 10.1109/WACV48630.2021.00318.

【作者简介】

安旭（2002—），男，湖北黄冈人，硕士研究生，研究方向：深度学习与目标检测。

胡蓉华（1985—），男，湖北荆州人，博士，讲师、研究生导师，研究方向：人工智能及安全。

（收稿日期：2025-10-30 修回日期：2026-05-14）