

基于YOLOv8的雾天环境下车辆行人检测算法研究

杜雨亭, 何 嘉

(成都信息工程大学 计算机学院, 四川 成都 610000)

摘要: 雾天天气对自动驾驶系统性能影响较大, 导致基于深度学习的目标检测模型在雾天条件下的漏检率和误检率偏高。基于该问题, 提出一种雾天环境下的目标检测算法——PF-YOLOv8-AL。通过借鉴AOD-Net轻量网络思想, 改变网络结构, 融合快速空间金字塔池化结构SPPF, 使用SSIM损失函数构建了PF-Net去雾网络, 再着重对YOLOv8检测网络进行改进, 从而构建了YOLOv8-AL网络。对YOLOv8进行的改进包括融合全局平均池化及LSKA注意力的SPPF-AL模块, 用MPDIoU损失函数替代CIoU损失函数及增加P2检测头, 再将PF-Net与YOLOv8-AL整合成PF-YOLOv8-AL网络。模型在Foggy Cityscapes数据集上进行训练, 在真实雾天RTTS数据集上进行测试。实验结果表明, 该算法在雾天环境场景中具有较高检测精度, 平均精度值达82.2%, 比YOLOv8提升5.4%, 平均精度均值达61.6%, 比YOLOv8提升8.1%。

关键词: 雾天环境; 图像去雾; 目标检测; YOLOv8; 损失函数

DOI: 10.11907/rjtk.241766

中图分类号: TP391.41

文献标识码: A

文章编号: 1672-7800(2025)012-0018-07



扫描二维码阅读全文:

Research on Vehicle and Pedestrian Detection Algorithm Based on YOLOv8 in Foggy Environment

DU Yuting, HE Jia

(School of Computer Science, Chengdu University of Information Technology, Chengdu 610000, China)

Abstract: Due to the significant impact of foggy weather on the performance of autonomous driving systems, the target detection model based on deep learning has a high false positive rate and false negative rate in foggy conditions. Based on this problem, this paper proposes a target detection algorithm for foggy environments – PF-YOLOv8-AL. First, by borrowing the idea of lightweight network from AOD-Net, changing the network structure, integrating the fast spatial pyramid pooling structure SPPF and using the SSIM loss function, a PF-Net dehazing network is constructed. Then, the YOLOv8 detection network is specially improved to construct the YOLOv8-AL network. The improvements made to YOLOv8 include integrating the SPPF-AL module with the global average pooling and LSKA attention, replacing the CIoU loss function with the MPDIoU loss function, and adding the P2 detection head. Finally, the PF-Net and YOLOv8-AL are integrated into the PF-YOLOv8-AL network. The model is trained on the Foggy Cityscapes dataset and tested on the real foggy RTTS dataset. The experimental results show that the proposed algorithm has a high detection accuracy in foggy environment scenes, with an average precision value of 82.2%, an improvement of 5.4% over YOLOv8, and an average precision mean value of 61.6%, an improvement of 8.1% over YOLOv8.

Key Words: foggy environment; image dehazing; object detection; YOLOv8; loss function

0 引言

当前, 计算机视觉技术突飞猛进, 推动着科学研究和工程应用领域不断向前发展, 例如智能监控、医学影像分

析和自动驾驶等。其中, 目标检测作为计算机视觉领域的一个重要分支, 对于提高相关应用的安全性和效率具有重要作用。在雾天, 大气中的水滴和颗粒物会散射光线, 导致图像能见度下降、对比度降低, 甚至使目标的外观和结构变得模糊不清。因此, 在雾天环境下提升目标检测准确

收稿日期: 2024-09-30

基金项目: 四川省科技厅国际合作项目(SYZ202068)

作者简介: 杜雨亭(1998-), 女, 成都信息工程大学计算机学院硕士研究生, 研究方向为图像处理与深度学习; 何嘉(1968-), 女, 博士, 成都信息工程大学计算机学院教授、硕士生导师, 研究方向为人工智能、数据挖掘。本文通讯作者: 何嘉。

性和鲁棒性, 成为该领域亟待解决的问题。

根据去雾原理不同, 常用的去雾算法可分为 3 类: 基于图像增强、基于图像复原以及基于深度学习技术^[1]。Soni 等^[2]结合自适应直方图均衡和引导过滤法以增强图像对比度, 扩大图像灰度范围, 但像素值的重新分配会导致一些图像纹理信息及细节丢失。He 等^[3]提出暗通道先验 (DCP) 去雾算法, 利用非天空区域的 RGB 某通道的像素值很低的特点进行去雾, 虽去雾能力好但对天空部分的去雾效果差, 并且其处理速度较慢。Cai 等^[4]提出的 DehazeNet 算法是通过获得雾霾退化模型中的介质透射率实现去雾, 但是其将全局大气光值选取为固定经验值, 导致得到的去雾图像可能存在偏差。Li 等^[5]设计了轻量级 CNN 去雾模型 AOD-Net, 将其与 Faster R-CNN^[6]结合训练, 速度快但效果一般。考虑去雾网络的实时性, 本文采用 AOD-Net 去雾模型进行改进, 并进行图像去雾。

目前常见的目标检测算法可分为两类: 一是以 R-CNN^[7]系列、Mask-RCNN^[8]系列为代表的两阶段目标检测; 二是以 YOLO^[9-12]系列、SSD^[13] (Single Shot MultiBox Detector) 系列为代表的单阶段目标检测。相较于两阶段目标算法, 单阶段目标算法检测速度快、实时性高, 已广泛运用于雾天场景下的目标检测。Abbasi 等^[14]设计了雾天自适应 YOLO 算法, 对其自建的雾天数据集集中的车辆和行人检测准确率达 72%, 存在检测准确率偏低的问题。Liu 等^[15]在 YOLOv5 中添加了 CBAM 注意力机制和 Transformer 编码层以优化模型结构, 有效增强了模型的特征提取能力, 在 Citylandscapes 雾天数据集上获得了 43.3% 的准确率。Liu 等^[16]提出一种将 GCANet 去雾算法与基于检测算法的卷积特征融合训练相结合的方法, 将晴天图像和相应的边缘特征图像通过 YOLOv5 进行障碍物检测训练, 结果显示检测精度有所提高, 但也存在实时性较低的问题。

上述算法在雾天环境下的目标检测中虽取得了不错的检测性能, 但依然存在检测实时性与准确性难以平衡的问题。为了解决此问题, 本文选择了去雾速度较快的 AOD-Net 和单阶段检测算法 YOLOv8, 提出一种基于 AOD-Net 与 YOLOv8 结合的雾天目标检测算法 PF-YOLOv8-FPM。在此基础上, 本文作出如下贡献: ①调整 AOD-Net 网络结构, 融入快速空间金字塔池化模块 SPPF, 并使用 SSIM^[17] 损失函数, 以形成新的去雾网络 PF-Net; ②提出融合全局平均池化和注意力机制 LSKA^[18] 的 SP-PF-AL 模块; ③用 MPDIoU^[19] 损失函数替代 CIoU 损失函数; ④增加 P2 检测头。本文通过对以上改进进行对比及消融实验, 在 RTTS 数据集上测试发现, 其有效提升了雾天场景下的目标检测精度。

1 AOD-Net 图像去雾算法改进

1.1 大气散射模型与 AOD-Net 去雾模型算法

大气散射模型认为在强散射介质下, 目标反射光受大气中悬浮粒子的吸收和散射作用, 造成目标反射光能量的衰减, 导致探测系统的成像结果亮度降低, 对比度下降。并且太阳光等环境光受大气中散射介质的散射作用形成背景光, 通常这背景光强度大于目标光, 因而造成成像结果模糊不清。其数学模型可简单表示为如式 (1) 所示。

$$I(x) = J(x)t(x) + A[1 - t(x)] \quad (1)$$

其中, $I(x)$ 表示待去雾图像 (原始图像), $J(x)$ 表示去雾后的图像, A 表示无穷远处的大气光成分, $t(x)$ 表示透射率。AOD-Net 核心思想是将式 (1) 中的 $I(x)$ 和 A 两个未知量统一成一个参数 $K(x)$, 其变换公式如式 (2) 所示, 得到简化的大气散射模型如式 (3) 所示。

$$K(x) = \frac{A - b + \frac{I(x) - A}{t(x)}}{I(x) - 1} \quad (2)$$

$$J(x) = K(x)I(x) - K(x) + b \quad (3)$$

式 (2)、式 (3) 中的 b 是默认为 1 的常数, 将透射系数和全局大气光值统一到一个变量中并且估算这个变量, 可以避免由于某些背景光的影响而导致大气光值 A 估算过大。

AOD-Net 结构分为 K-估计模块和干净图像生成模块。K-估计模块是 AOD-Net 的关键, 负责估计深度和相对雾度水平, 其使用了 5 个卷积层, 并通过融合不同尺寸的滤波器形成多尺度特征。干净图像生成模块由一个逐元素乘法层和多个逐元素加法层组成, 以便生成恢复图像。AOD-Net 每个卷积层都只用了 3 个过滤器, 这是该模型轻量化的原因。AOD-Net 结构如图 1 所示。

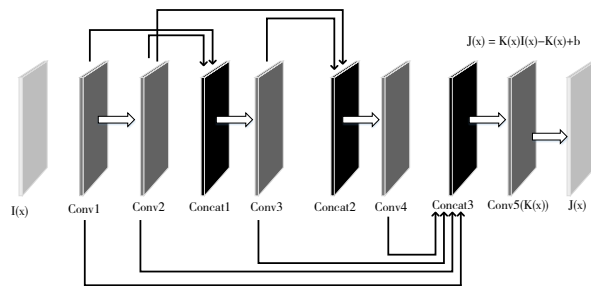


Fig. 1 AOD-Net network architecture

图 1 AOD-Net 网络结构

1.2 PF-Net 去雾算法构建

由于 AOD-Net 的网络深度浅、感受野小, 使其只关注局部特征而忽略整体, 导致去雾后的图像存在纹理不清、模糊等问题。为了解决此问题, 在 AOD-Net 架构基础上实现了 PF-Net, 其结构如图 2 所示。首先, 有雾图像输入后, 将依次经过 3 个卷积层以提取图像的浅层特征; 其次, 在第一维度上拼接以完成特征融合, 并提取图像的深层次特征信息; 再次, 输入改进的 SPPF 空间金字塔结构。该结构

对输入信息进行 1×1 的卷积后连续进行 3×3 、 5×5 、 7×7 的最大池化操作,再将每一次池化后的输出和初始 1×1 卷积后的输出相融合并输入到下一个卷积中,最后进行 1×1 的卷积操作以分离上下文特征。SPPF结构在保证多尺度融合的同时,降低了计算量,并进一步增大感受野。最后输出通过大气散射模型得到去雾图像。

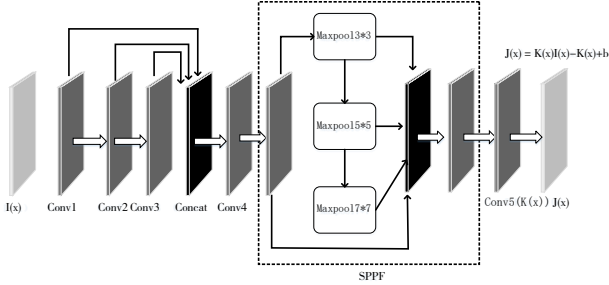


Fig. 2 PF-Net network architecture

图2 PF-Net网络结构

基于卷积神经网络的去雾算法在选择损失函数时,常采用对抗损失、平均绝对误差损失(MAE)、均方误差(MSE)等指标。然而,仅依赖单一指标对网络进行调参,可能会产生偏差并存在某些局限性。PF-Net选择SSIM指标构建损失函数,SSIM从亮度、对比度、结构3个方面对图像进行评价,具有一定的综合性。其公式如式(4)所示。

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \quad (4)$$

其中, μ_x 与 μ_y 是图片 X 和 Y 的平均值, σ_x^2 与 σ_y^2 是图片 X 和 Y 的方差, σ_{xy} 是图片 X 和 Y 的协方差。 $C_1=(k_1L)^2$, $C_2=(k_2L)^2$ 是用来维持稳定的常数。 L 是像素值的动态范围, $k_1=0.01$, $k_2=0.03$ 。本文改进的损失函数如式(5)所示。

$$Loss = 1 - SSIM \quad (5)$$

SSIM值越大表示去雾效果越好。由于SSIM的取值范围是 $[0, 1]$,则SSIM值越大Loss值越小,去雾算法最终效果越好。相较于AOD-Net,PF-Net能捕获并融合更为丰富的图像特征,增强模型学习能力,在保证去雾速率的同时还拥有更优的去雾效果。

2 YOLOv8目标检测算法改进

2.1 YOLOv8算法改进原理

为了有效解决雾天环境下对远处模糊、被遮挡行人与车辆检测准确率低的问题,同时保证检测实时性,本文选用YOLOv8作为雾天行人与车辆的检测网络基础,在此基础上进行改进,从而提出一种新型检测网络YOLOv8-AL。

2.1.1 融合全局平均池化与注意力机制LSKA的SPPF-AL模块

YOLOv8采用了快速空间金字塔SPPF结构,SPPF首先将输入的特征信息经过标准卷积进行特征提取,接着连续3次经过残差结构的最大池化(卷积核为 5×5),最后将

每次池化前和池化后的结果相融合后的特征信息再进行 1×1 卷积。尽管SPPF通过最大池化级联融合不同感受野的特征图,丰富了特征图的表达能力,但当其应用于雾天情境下,对图像中较小、模糊的人或车辆进行检测时发现,SPPF获取图像局部信息过多,忽略了背景信息,影响了模型准确率。因此,在SPPF模块基础上加入全局平均池化以帮助模型获取图像背景信息并且减轻不同尺度大小所带来的影响。全局平均池化直接对整个特征图进行平均池化,将每个通道的特征值平均生成对应于每个通道的汇聚特征值。

由于雾天检测环境复杂多变,为了使模型增强雾天检测行人车辆相关的特征通道,在加入了全局平均池化的SPPF模块基础上,继续引入了LSKA注意力,结构如图3所示。

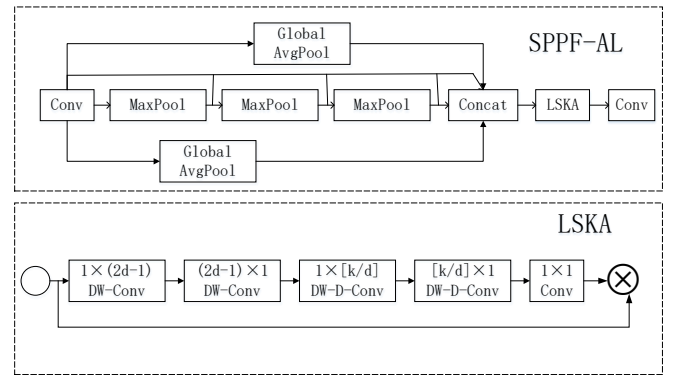


Fig. 3 Improved SPPF structure and LSKA structure

图3 改进的SPPF结构和LSKA结构

LSKA是一种大核可分离注意力模块,通过标准深度卷积(DW-Conv)捕获图像的局部上下文信息,再使用具有大感受野内核的深度膨胀卷积(DW-D-Conv)模拟长程依赖性。LSKA将深度卷积层的2D卷积核分解为级联的水平1D和垂直1D内核,在不增加计算成本的情况下覆盖更大的感受野,以捕获更广泛的上下文信息。LSKA通过最后一个 1×1 卷积层融合得到的特征,生成最终注意力图。通过应用该注意力图,原始特征图中每个元素根据注意力图的值进行加权,抑制不重要的特征而突出重要特征。LSKA输出公式如式(6)一式(9)所示,其中 Y 代表哈达玛乘积, $\times$ 代表卷积。

$$X^C = \sum_{H,W} W_{(2d-1) \times 1}^C * \left(\sum_{H,W} W_{1 \times (2d-1)}^C * F^C \right) \quad (6)$$

$$Z^C = \sum_{H,W} W_{\left[\frac{k}{d}\right] \times 1}^C * \left(\sum_{H,W} W_{1 \times \left[\frac{k}{d}\right]}^C * X^C \right) \quad (7)$$

$$A^C = W_{1 \times 1} * Z^C \quad (8)$$

$$T^C = A^C Y F^C \quad (9)$$

其中, d 代表膨胀率, Z^C 代表通过将大小为 $k \times k$ 的核与输入特征图进行卷积而获得的深度卷积的输出。 Z^C 经过 1×1 卷积后得到注意力图 A^C ,最终LSKA的输出 T^C 是注意力图 A^C 与输入特征图 F^C 的哈达玛乘积。对SPPF融入这两处改进可以增强对待检测目标的识别能力,提高主干特

征的特征提取能力。

2.1.2 损失函数优化

YOLOv8使用了CIoU损失,CIoU考虑了中心点距离和纵横比,但是它定义的纵横比是相对值而不是绝对值,没有考虑宽、高与其置信度的真实差异。相较于CIoU,EIoU可解决上述问题,但是当预测的边界框和真实边框具有相同的宽高比而宽度和高度值不同时,EIoU将失去有效性。因此,本文使用MPDIoU Loss代替CIoU Loss,通过最小化预测边界框和真实边界框之间的左上和右下点距离,更好地进行网络训练。MPDIoU Loss计算公式如式(10)所示,相关距离计算公式如式(11)、式(12)所示。

$$L_{MPDIoU} = 1 - IOU + \frac{d_1^2}{h^2 + w^2} + \frac{d_2^2}{h^2 + w^2} \quad (10)$$

$$d_1^2 = (x_1^{prd} - x_1^{gt})^2 + (y_1^{prd} - y_1^{gt})^2 \quad (11)$$

$$d_2^2 = (x_2^{prd} - x_2^{gt})^2 + (y_2^{prd} - y_2^{gt})^2 \quad (12)$$

其中, (x_1^{prd}, y_1^{prd}) 是预测框的左上角坐标, (x_2^{prd}, y_2^{prd}) 是预测框的右下角坐标, (x_1^{gt}, y_1^{gt}) 是真实框的左上角坐标, (x_2^{gt}, y_2^{gt}) 是真实框的右下角坐标, w 与 h 分别为输入图像的宽和高。

MPDIoU Loss 是一种基于最小点距离的新型损失函数,其考虑了现有损失函数中要考虑的所有相关因素,例如重叠或非重叠面积、中心点距离及宽高的偏差,通过比较边界框回归过程中预测边界框和实际标注边界框之间的相似性,最终获得了更快的网络收敛速度和更准确的回归结果。

2.1.3 P2 目标检测头增加

在雾天场景中,图像存在远处行人车辆遮挡率高、图像模糊及目标尺寸小等问题,这使得YOLOv8在检测此类目标时存在一定局限性。YOLOv8的原始网络结构仅包含3个检测头,分别是P3、P4和P5。在模型中,由于下采样操作会导致小目标信息丢失,进而使得在较深的特征图上难以精确识别小目标。针对这一问题,本文增加了一个P2目标检测头。将模型提取的浅层特征层P2引入进Neck网络,使模型在早阶段获取小目标、高分辨率的特征信息。通过融合不同尺寸目标的高分辨率特征信息,利用高层特征图的语义信息和低层特征图的细节信息,提升模型对于雾天场景下对远处行人车辆的感知能力和检测能力。

通过以上各模块对YOLOv8加以改进,得到YOLOv8-AL网络结构模型如图4所示。

3 实验与结果分析

3.1 实验配置与数据集

本文实验的操作系统是Windows 10 专业版,64位操作系统,运行内存32 G,13th Gen Intel(R) Core(TM) i5-13400F 2.50 GHz 处理器,NVIDIA GeForce RTX 3060 型GPU,计算框架版本为CUDA 11.7,Python版本3.9。图像

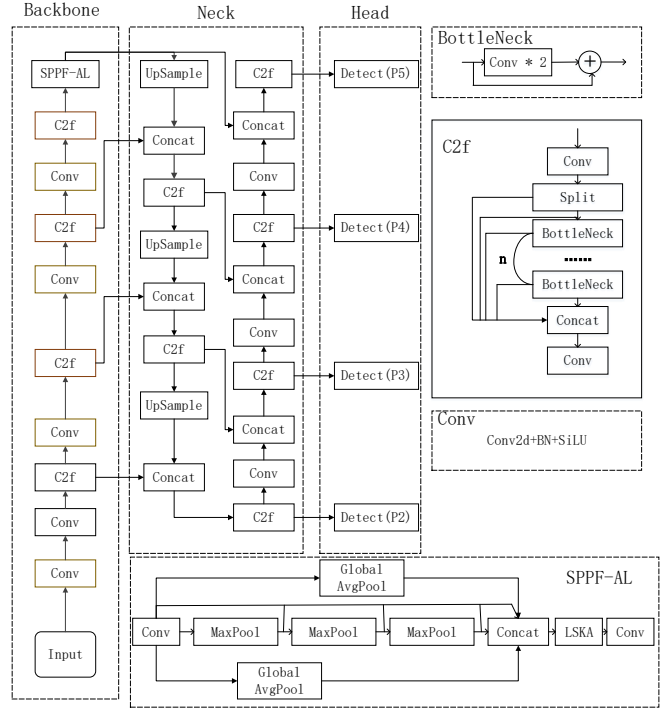


Fig. 4 YOLOv8-AL network architecture

图4 YOLOv8-AL网络结构

去雾实验采用Python编程语言完成模型训练和测试, batch-size为8,初始学习率为0.000 1,权重衰减为0.000 1, epochs为50。去雾后的目标检测实验是在Pytorch深度学习框架下完成模型训练和测试, batch-size为16,初始学习率为0.01,权重衰减为0.000 5, epochs为300。

本文实验以Cityscapes数据集为基础,采用Christos Sakaridis提出的雾模拟管道,人工生成3种浓度(beta=0.005,0.01,0.02)的Foggy Cityscapes数据集,不同雾浓度的示例图像展示如图5所示。该数据集一共10 365张图片,包含了50个城市不同场景、不同背景、不同季节的街景,图像中包含多类目标。本文筛选自行车(bicycle)、公交车(bus)、汽车(car)、行人(person)4类目标作为检测对象,将数据集随机分为训练集和验证集两个部分,比例为9:1,得到9 328张训练集照片和1 037张验证集照片。本文测试集选用数据集RTTS,RTTS包含4 322幅自然雾天的图像,是一个已经标注的真实雾天数据集。

3.2 评价指标

3.2.1 图像去雾算法评价指标

本文采用峰值信噪比(PSRN)和结构相似性(SSIM)两个指标。PSRN是基于MSE(均方误差)定义,用于评价两幅图像之间的相似程度。其公式如式(1)所示。

$$PSRN = 20 \times \log_{10} \left(\frac{MAX_I}{\sqrt{MSE}} \right) \quad (13)$$

其中, MAX_I 为信号峰值,在图像中为最大像素值。PSRN值越大则图像质量越好。与PSRN衡量的绝对误差不同,SSIM通过关注边缘和纹理相似性以模仿人类感知,更加符合人眼的直观感受。SSIM主要考量图片的3个关



Fig. 5 Images of three different fog concentrations

图5 3种雾浓度的图像

键特征:亮度、对比度、结构,其取值范围为 $[-1, 1]$ 。其公式为上文提到的式(4)。

3.2.2 目标检测算法评价指标

本文选用 $mAP@0.5$ 、 $mAP@0.5:0.95$ 作为目标检测模型精度衡量指标,两个指标依赖于精确率(Precision, P)和召回率(Recall, R)的计算。其计算公式分别如式(14)、式(15)所示,其中 TP 为正确检测到物体,FP 为误检为物体, FN 为误检为非物体。

$$P = \frac{TP}{TP + FP} \quad (14)$$

$$R = \frac{TP}{TP + FN} \quad (15)$$

平均精确率(AP)是 PR 曲线与横坐标召回率和纵坐标精确率围起来的面积,其计算公式如式(16)所示。

$$AP = \int_0^1 PR dr \quad (16)$$

mAP 是平均精确率均值, $mAP@0.5$ 和 $mAP@0.5:0.95$ 计算公式分别如式(17)、式(18)所示。

$$mAP@0.5 = \frac{1}{C} \sum_{k=1}^C AP@0.5_k \quad (17)$$

$$mAP@0.5:0.95 = \frac{1}{10} mAP@0.5 + \frac{1}{10} mAP@0.55 + \dots + \frac{1}{10} mAP@0.95 \quad (18)$$

其中, C 为总类别数,此处 $C=4$ 。

参数量(Params)是指网络结构的参数总数,代表了算

法的复杂度。计算量(GFLOPs)是指每秒10亿次的浮点运算数,理论上该数值越高,对设备性能越高。每秒帧数(FPS)是指网络每秒能处理的图片数量,即网络能够实时处理的数据量。

3.3 实验结果对比分析

3.3.1 去雾算法实验对比

为了有效评价本文 PF-Net 去雾算法与 AOD-Net 和其他去雾算法之间的差异,选择了 PSNR 和 SSIM 两个指标进行评价。采用 RESIDE 数据集中混合主观测试集 HSTS 进行实验评价,实验结果如表1所示。

PSNR 值越大,SSIM 值越接近 1,说明图片失真越小,去雾效果越好。从表1可以清楚看出,AOD-Net 算法去雾效果明显优于暗通道先验和直方图均衡化去雾方法,而本文 PF-Net 去雾算法效果优于 AOD-Net,但要逊色于 DehazeNet。不过,在精度上取得优势的 DehazeNet 在运行速度上远远慢于 PF-Net,PF-Net 在精度与处理速度的平衡上更佳,可满足实际应用场景需求。

Table 1 Evaluation results

表1 评价结果

去雾算法	PSNR	SSIM
暗通道先验(DCP)	17.264	0.835
直方图均衡化	18.845	0.685
AOD-Net	19.656	0.856
DehazeNet	23.429	0.868
PF-Net	20.592	0.862

3.3.2 模型算法实验对比

本文通过在 PF-Net 去雾后的 Foggy Cityscapes 数据集上进行实验,评估了 SPPF-AL 模块、P2 检测头模块以及更换 MPDIoU 损失函数改进措施对目标检测模型性能的影响。实验使用 YOLOv8 作为基础网络,实验过程记录了不同实验组合的结果,如表2所示。图6展示了消融实验前后的检测结果对比。

Table 2 Results of ablation experiment

表2 消融实验结果

SPPF-AL	MPDIoU	P2	$mAP@50\%$	$mAP@50:90\%$	GFLOPS	FPS
			78.3	54.2	28.4	77
√			79.0	55.4	30.6	82
√	√		79.5	56.6	30.6	83
√		√	80.8	58.9	36.8	73
	√	√	81.1	60.2	34.6	73
√	√	√	82.2	61.6	36.8	78

根据表2数据可知,通过引入 SPPF-AL 模块,模型在 $mAP@50$ 和 $mAP@50:95$ 指标上相较于基线网络分别提升了 0.7% 和 1.2%,验证了该模块是一种有效的特征提取技术改进方案。当同时应用 SPPF-AL 模块和 P2 模块时,模型在两个主要指标上都实现了显著提升,其中 $mAP@50$ 增长至 80.8%, $mAP@50:95$ 提升至 58.9%,表明了增加 P2 检测头的有效性。此外,当用 MPDIoU 损失函数替代原有的 CIoU 损失函数后,模型精度再次获得提升, $mAP@50$ 和

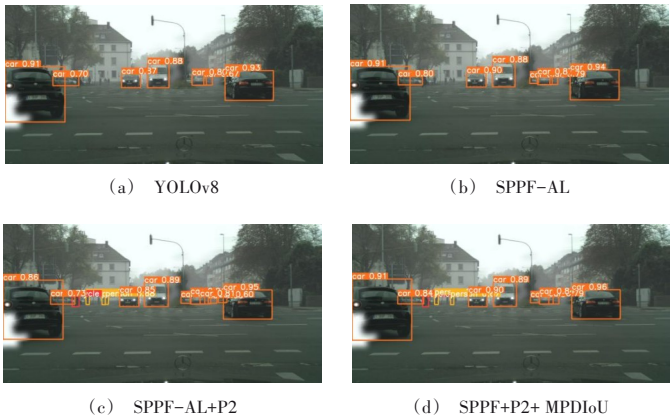


Fig. 6 Comparison of ablation experiment detection results

图 6 消融实验检测结果比较

mAP@50:95 分别提升至 82.2% 和 61.6%, 这是由于 MPDIoU 提高了对行人车辆边框的定位精度且加快了模型检测速度。值得注意的是, 这些改进虽然带来了计算量的增加, 但是在图像处理速度上有些许提升, 在两个准确率指标上都得到了较大提升, 特别是 mAP@50:95 提升了 7.4%。由图 6 可知, 本文模型在准确率、检测出的类别数、检测远处模糊遮挡的行人和车辆这 3 个指标上都明显优于初始的 YOLOv8 网络。综上所述, 本文改进模型在雾天场景下的目标检测任务中, 能够有效地提升模型检测精度。

为了验证本文算法的优越检测性能, 将 PF-YOLOv8-AL 与 YOLOv5s、YOLOv8n、YOLOv8s、YOLOv10s、AOD-YOLOv8s、PF-YOLOv8s 等算法在 Foggy Cityscapes 数据集下进行对比。其中, AOD-YOLOv8s 算法首先利用 AOD-Net 进行图像去雾处理, 随后采用 YOLOv8s 进行目标检测。而 PF-YOLOv8s 算法则首先运用本文提出的 PF-Net 进行图像去雾, 再使用 YOLOv8s 进行目标检测。实验对比结果如表 3 所示。

Table 3 Comparative experimental results

表 3 对比实验结果

方法	mAP@50/%	mAP@50:95/%	Params/M	GFLOPS	FPS
YOLOv5s	75.9	52.8	7.0	15.8	63
YOLOv8n	60.0	40.6	3.0	8.1	83
YOLOv8s	76.8	53.5	11.1	28.4	77
YOLOv10s	76.3	53.3	8.0	24.5	137
AOD-YOLOv8s	77.9	54.0	11.1	28.4	77
PF-YOLOv8s	78.3	54.2	11.1	28.4	77
本文算法	82.2	61.6	12.7	36.8	78

由表 3 可知, 本文提出的 PF-Net 去雾算法相较于 AOD-Net 展现出了更佳效果。在采用相同 YOLOv8s 的目标检测模型时, 使用 PF-Net 去雾算法模型的检测精度高于使用 AOD-Net, 充分展示了 PF-Net 去雾效果的优势。YOLOv5s、YOLOv8n 和 YOLOv8s 模型具有模型参数较小、计算复杂度低的优点, 而本文算法虽参数量较 YOLOv8s 略有增加 (主要是因为 P2 模块增加了网络复杂度), 但是图

像处理速率没有损失, 且准确率得到大幅度提升。值得一提的是, YOLO 系列目标检测算法更迭较快, YOLOv10s 在准确性上低于 YOLOv8s, 但在处理速率上优于 YOLOv8s, 这为后续优化提供了新的思路。由于硬件限制以及考虑实时和准确性的平衡, 对于大多数实时应用场景, 60 帧/s 即可满足需求。本文算法相较以上所有算法在 mAP 值上有明显提升, 展现出了一定的性能优势。

3.3.3 算法检测结果分析

本文采用 RTTS 数据集作为测试集, 对模型进行检测。受本文篇幅限制, 仅展示出 YOLOv8 算法、AOD-YOLOv8 算法与本文 PF-YOLOv8-AL 算法的检测结果比较。并且, 选择了两种不同构图的雾图进行展示, 图 7 直观地呈现了这 3 种模型在选定雾图上的检测结果。



Fig. 7 Comparison of test results

图 7 检测结果比较

图 7(a) 为 YOLOv8 的检测结果, 图 7(b) 为 AOD-YOLOv8 的检测结果, 图 7(c) 为本文模型检测结果。可以看出, PF-Net 去雾后的图片对比度及图片亮度高于 AOD-Net, 使用 PF-Net 去雾对目标检测精度的提升有帮助。在不同场景下的雾天图片中, YOLOv8s 与 AOD-YOLOv8s 模型在面对模糊有遮挡、清晰度不高的目标检测时, 检测能力有所下降且存在误检漏检现象。而本文模型在感知能力上具有优势, 能够检测出更多目标, 有更低的误检漏检概率和更高的检测精度, 同时具有较强的泛化能力, 在不同场景下表现出良好性能。

4 结语

本文针对雾天能见度低的场景下目标检测模型性能下降的问题,提出了新的目标检测模型 PF-YOLOv8-AL。该模型融合了 PF-Net 去雾模块和 YOLOv8-AL 目标检测模块。PF-Net 去雾模块在 AOD-Net 的基础上,作了 2 处改进:①融合改进的空间金字塔池化 SPPF;②使用 SSIM 损失函数。在 HSTS 数据集的验证下,PF-Net 在去雾精度和去雾速度的平衡上达到更佳。YOLOv8-AL 检测算法首先在快速空间金字塔池化 SPPF 结构融合全局平均池化及 LS-KA 注意力机制,增强网络特征提取能力;接着,将 YOLOv8 的损失函数 CIOU Loss 替换为 MPDIoU Loss,在加快网络收敛速度的同时提高检测网络精度;最后,增加 P2 检测头,增强模型检测远处模糊行人和车辆目标的能力。在 RTTS 数据集上对新模型 PF-YOLOv8-AL 进行验证,其检测结果优于其他对比模型,验证了模型的有效性。

但同时,新模型也存在一定局限性,比如真实雾天情况多样,影响因素也是多方面的,而非仅限于本文研究中的情况。针对未来智能驾驶需求,将继续对模型进行联合优化以适应多样情况的雾天场景。

参考文献:

- [1] XIE Y, JIA H Z, WANG T H, et al. Overview of image defogging algorithms [J]. Computer & Digital Engineering, 2022, 50(12): 2765-2774.
谢勇, 贾惠珍, 王同军, 等. 图像去雾算法综述[J]. 计算机与数字工程, 2022, 50(12): 2765-2774.
- [2] SONI B, MATHUR P. An improved image dehazing technique using CLAHE and guided filter [C]//Proceedings of the 7th International Conference on Signal Proceeding and Integrated Networks (SPIN), 2020: 902-907.
- [3] HE K M, SUN J, TANG X O, et al. Single image haze removal using dark channel prior [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2011, 33(12): 2341-2353.
- [4] CAI B L, XU X M, JIA K, et al. DehazeNet: an end-to-end system for single image haze removal [J]. IEEE Transactions on Image Processing, 2016, 25(11): 5187-5198.
- [5] LI B Y, PENG X L, WANG Z Y, et al. AOD-Net: all-in-one dehazing network [C]//Proceedings of The IEEE International Conference on Computer Vision, 2017: 4770-4778.
- [6] REN S Q, HE K M, GIRSHICK R, et al. Faster R-CNN: towards real time object detection with region proposal networks [C]//International Conference on Neural Information Processing Systems, 2015: 91-99.
- [7] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation [C]//Columbus: Proceedings of The IEEE Conference on Computer Vision and Pattern Recognition, 2014.
- [8] ZHU L, LI D B, YAN X Z, et al. Intelligent detection method of tunnel cracks based on improved Mask R-CNN deep learning algorithm [J]. Journal of Graphics, 2023, 44(1): 177-183.
朱磊, 李东彪, 闫星志, 等. 基于改进 Mask R-CNN 深度学习算法的隧道裂缝智能检测方法[J]. 图形学报, 2023, 44(1): 177-183.
- [9] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. YOLOv4: optimal speed and accuracy of object detection [DB/OL]. <https://arxiv.org/abs/2004.10934>, 2020.
- [10] JOCHER G, CHAURASIA A, QIU J. Ultralytics YOLOv8 [EB/OL]. <https://github.com/Ultralytics/Ultralytics>.
- [11] WANG C, YEH I, LIAO H. YOLOv9: learning what you want to learn using programmable gradient information [DB/OL]. <https://arxiv.org/abs/2402.13616>, 2024.
- [12] AO W, HUI C, KAI C, et al. YOLOv10: real-time end-to-end object detection [DB/OL]. <https://arxiv.org/abs/2405.14458>, 2024.
- [13] CHEN Q H, JI J, CHONG Y F, et al. Vehicle and pedestrian detection based on AOD-Net and SSD algorithm in hazy environment [J]. Journal of Chongqing University of Technology (Natural Science), 2021, 35(5): 108-117.
陈琼红, 冀杰, 种一帆, 等. 基于改进 AOD-Net 和 SSD 的雾天车辆和行人检测[J]. 重庆理工大学学报(自然科学), 2021, 35(5): 108-117.
- [14] HASAN A, MARZIEH A, RICAHRD Y. Fog-aware adaptive yolo for object detection in adverse weather [C]//Ottawa: 2023 IEEE Sensors Applications Symposium, 2023.
- [15] LIU X Y, ZHANG B F, LIU N. Cast-yolo: an improved yolo based on a cross-attention strategy transformer for foggy weather adaptive detection [J]. Applied Sciences, 2023, 13(2): 1176.
- [16] LIU Z H, ZHAO S J, WANG X. Research on driving obstacle detection technology in foggy weather based on GCANet and feature fusion training [J]. Sensors, 2023, 23(5): 2822.
- [17] ZHOU W, BOVIK A C, SHEIKH H R, et al. Image quality assessment: from error visibility to structural similarity [J]. IEEE Trans Image Process, 2004, 13(4): 600-612.
- [18] LAU A K, PO L M, REHMAN Y A U. Large separable kernel attention: rethinking the large kernel attention design in CNN [J]. Expert Systems with Applications, 2024, 236: 121352.
- [19] MA S L, XU Y. MPDIoU: a loss for efficient and accurate bounding box regression [DB/OL]. <https://arxiv.org/abs/2307.07662>, 2023.
- [20] CHEN J R, HE H, KAO S H, et al. Run, don't walk: chasing higher FLOPS for faster neural networks [C]//Vancouver: 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2023.
- [21] HUANG S C, LE T H, JAW D W. DSNet: joint semantic learning for object detection in inclement weather conditions [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 43(8): 2623-2633.
- [22] XU S, WANG X, LYU W, et al. PP-YOLOE: an evolved version of YOLO [DB/OL]. <https://arxiv.org/abs/2203.16250>, 2022.
- [23] LI B, REN W, FU D, et al. Benchmarking single-image dehazing and beyond [J]. IEEE Transactions on Image Processing, 2018, 28(1): 492-505.
- [24] LIU W, REN G, YU R, et al. Image-adaptive YOLO for object detection in adverse weather conditions [C]//Proceedings of the AAAI Conference on Artificial Intelligence, 2022: 1792-1800.

(责任编辑:孙娟)