

基于 YOLOv7 的雾天实时目标检测方法^{*}

谢祖华 李海涛 胡建文

(长沙理工大学电气与信息工程学院 长沙 410114)

摘 要: 雾天场景下,拍摄的图像模糊、细节信息缺失,目标与背景难以区分。针对现有深度学习目标检测算法易出现漏检、误检、识别速度慢等问题,提出一种基于 YOLOv7 的雾天实时检测算法。以 YOLOv7 为基线,在其前端设计包含 AOD 去雾子网络和雾天图像生成子网络的循环去雾双子网络。轻量化的 AOD 去雾子网络占用很少的计算资源,有效克服雾天对图像造成的负面影响,增强模型的特征提取能力;雾图像生成子网络在模型训练阶段协助提升 AOD 子网络的去雾性能,并在测试时不参与计算,显著减少推理时间。改进的图像重建损失函数引入模糊图像信息,并将整体网络统一训练有效地结合去雾和检测任务。将 CityScapes 数据集合成得到两个不同雾浓度的雾天图像数据集,在两个数据集上的实验结果表明,该方法的平均精确度分别为 65.2%、64.2%,检测速度 FPS 为 42.4,模型精度在所有对比方法中最优且能实现实时检测;最后将训练好的各模型在 RTTS 数据集上验证,所设计的模型泛化能力超过其他方法。

关键词: 雾天场景;图像去雾;YOLOv7;实时目标检测

中图分类号: TP391;TN911.73 **文献标识码:** A **国家标准学科分类代码:** 520.604

Real-time object detection based on YOLOv7 in foggy weather

Xie Zuhua Li Haitao Hu Jianwen

(School of Electrical and Information Engineering, Changsha University of Science and Technology, Changsha 410114, China)

Abstract: In the foggy scene, the captured image is blurred, the detail information is missing, and the target and the background are difficult to distinguish. Aiming at the problems of missed detection, false detection and slow recognition speed of existing deep learning target detection algorithms, a real-time object detection algorithm based on YOLOv7 in foggy weather is proposed. Taking YOLOv7 as the baseline, a cyclic defogging double sub-network including AOD defogging sub-network and foggy image generation sub-network is designed at the front end. The lightweight AOD defogging sub-network takes up little computing resources, effectively overcomes the negative impact of foggy days on the image, and enhances the feature extraction ability of the model. The fog image generation sub-network improves the dehazing performance of the AOD sub-network in the model training stage, and does not participate in the calculation during the test, which significantly reduces the inference time. The improved image reconstruction loss function introduces blurred image information, and the overall network unified training effectively combines defogging and detection tasks. The CityScapes data set is integrated into two foggy image data sets with different fog concentrations. The experimental results on the two data sets show that the average accuracy of the method is 65.2 % and 64.2 %, and the detection speed FPS is 42.4. The model accuracy is the best among all the comparison methods and can achieve real-time detection. Finally, the trained models are verified on the RTTS dataset, and the generalization ability of the designed model is better than other methods.

Keywords: foggy weather; image defogging; YOLOv7; real-time object detection

0 引 言

近年来,深度学习和大规模数据集的出现,使得目标检

测性能有了显著提升。尤其是基于卷积神经网络的目标检测方法,在基准数据集上取得了振奋人心的性能。然而,现有的大多数目标检测模型是在高质量图像上进行研究的。

收稿日期:2024-12-24

^{*} 基金项目:国家自然科学基金面上项目(62271087)、湖南省自然科学基金(2024JJ5039, 2023JJ60141)、湖南省教育厅科学研究重点项目(24A0243)资助

在实际环境中通常难以获取同等质量的图像,例如雾天城市交通监控、自动驾驶^[1-2]等场景。雾天成像时目标与雾天天气信息相互影响,造成图像模糊、目标与背景难以区分等问题,给雾天场景下目标检测任务带来了挑战。

相比一般的检测任务,目前对雾天场景下目标检测任务的研究较少。雾天目标检测方法大体可以分为两种:传统方法和深度学习方法。传统的雾天图像目标检测方法通常采用直方图均衡化^[3]、暗通道先验^[4]等方法对图像进行去雾,之后将增强后的图像送入检测模型识别。该类方法去雾效果不理想,去雾后的图像易出现块效应、光晕、过度曝光等问题,最终导致检测精度低。基于深度学习的雾天目标检测算法引入了卷积神经网络,包括领域自适应以及将图像去雾与目标检测相结合两种思路。领域自适应方法通常研究源域和目标域的特征对齐,模型更关注目标所在区域的特征即区域建议网络(region proposal network, RPN)生成的目标候选区域,因此大部分领域自适应方法是基于二阶段目标检测,难以实现实时检测;第二种思路有两种实现方式,其一与传统方法相同,先利用图像去雾方法对输入雾天图像进行预处理,然后将处理后的图像送入目标检测模型中进行检测;其二是将图像增强或图像去雾网络与目标检测网络组成一个整体网络,对网络进行统一训练。基于深度学习的方法相较于传统方法检测精度更高、泛化能力更强。因此,学者们对基于深度学习的图像去雾方法以及雾天目标检测方法进行了大量研究。例如,文献[5]提出了一种基于先验学习的图像去雾深度网络,它结合了传统去雾方法和深度学习方法的优点。Li 等^[6]提出基于光学散射模型的 AOD-Net 用来去雾,并与目标检测网络 Faster R-CNN 结合形成端到端的训练,这是第一个端到端可训练的雾天目标检测模型。Huang 等^[7]利用两个子网络共同学习能见度增强和目标检测,并通过共享特征提取层减少图像退化的影响,提出了一个联合学习能见度增强、目标分类和目标定位的雾天目标检测多任务网络 DSNNet,显著提升了在恶劣环境下的目标检测精度。文献[8]提出一种基于改进 YOLOv8s 的恶劣天气车辆行人检测方法,利用可扩张残差结构对主干网络的 C2F 模块优化并引入注意力机制,满足了恶劣天气下的车辆和行人检测需求。文献[9]对于弱光雾天图像,提出通过图像去雾(AOD)和图像增强(SAIP)的联合优化来提高图像质量,提高了微光雾图像质量,使用 YOLOv7 对增强后的弱光图像进行检测,提升了弱光雾天场景下的检测性能。Liu 等^[10]提出了图像自适应目标检测(IA-YOLO)方法,设计了一个可微分图像处理模块 DIP,该模块的超参数由一个基于轻量 CNN 的参数预测器 CNN-PP 自适应学习。CNN-PP 根据输入图像的亮度、颜色、色调和雾天天气信息自适应预测 DIP 的超参数。图像经过 DIP 模块处理后,抑制天气信息对图像的干扰,同时恢复潜在信息,最终得到了一种端到端学习 DIP、CNN-PP 和 YOLOv3^[11]骨干检测网络的联合优

化方案,在雾天和弱光场景下都取得了良好的性能。

上述算法通过对基准算法进行改进,提升了雾天场景目标检测性能,但是仍存在以下问题:首先,去雾模型的性能很大程度上决定了整体模型的准确性,而去雾的优劣主要与去雾网络的参数数量和计算量相关联,导致模型难以权衡检测精度与检测速度;其次,将雾天目标检测拆分成去雾任务和目标检测任务,预处理之后的图像会出现伪影、颜色失真等问题,在预处理后的图像上进行检测并不一定能提升检测精度。这些问题给雾天场景下目标检测算法带来了挑战。

对此,本文提出一种基于 YOLOv7 的雾天实时目标检测方法。本文以 YOLOv7^[12]为基线,在其前端设计双分支的循环去雾网络,推理时仅须轻量化的分支参与,运行时间非常少。对整体模型进行两阶段的统一训练以更好地结合去雾与检测两个任务。

1 图像去雾

大气散射模型是广泛应用于图像去雾的物理模型^[13-15],有雾图像 I 可以表示为:

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

其中, $I(x)$ 表示有雾图像 I 在像素点 x 处的值, J 表示无雾清晰图像, A 为大气光照强度, $t(x)$ 为透射率, $t(x)$ 定义如下:

$$t(x) = e^{-\beta d(x)} \quad (2)$$

其中, β 为大气散射系数, $d(x)$ 为物体与相机之间的距离。通过式 (1) 中的大气散射模型,可以推导出由有雾图像获取去雾后的清晰图像的公式:

$$J(x) = \frac{1}{t(x)}I(x) - A \frac{1}{t(x)} + A \quad (3)$$

因此,基于大气散射模型去雾的核心是求解大气光照强度 A 和透射率 $t(x)$ 两个未知量。在许多基于深度学习的去雾方法中,分别单独估计大气光照强度 A 和透射率 $t(x)$ 。

这种分别估计大气光照强度 A 和透射率 $t(x)$ 的方法虽然直观且有物理基础,但不能直接测量得到准确的大气光照强度 A 或透射率 $t(x)$,导致次优的图像恢复质量。此外,每个单独估计步骤中的误差将会累积导致误差增大。对此,Li 等^[6]对式 (3)进行了转换,将 $J(x)$ 重新表示为式 (4)。在式 (4)中, $t(x)$ 和 A 两个未知量被统一为一个新的未知量,即式 (5)中的 $K(x)$ 。于是,基于大气散射模型的图像去雾核心变成了求解 $K(x)$, b 是默认值为 1 的常数偏置。

$$J(x) = K(x)I(x) - K(x) + b \quad (4)$$

$$K(x) = \frac{I(x) - A + t(x)(A - b)}{t(x)(I(x) - 1)} \quad (5)$$

2 CD-YOLO 雾天目标检测网络

2.1 整体概述

为了准确地估计 $K(x)$,学者们不断提出模型更复杂、

层数更深的神经网络提取图像信息,然而仅仅沿用这种方式难以满足许多应用场景中的实时性要求。对此,受到文献[16]和[17]的启发,我们提出了循环去雾网络 CD-Net (cyclic de-fog network),其网络结构如图1所示,包含两个子网络:AOD去雾网络和雾天图像生成网络,双子网络的

设计策略是通过循环互助的方式优化图像去雾性能,从而增强整体模型的特征提取能力。将 CD-Net 与先进检测框架 YOLOv7 级联,得到一个端到端的雾天检测方法 CD-YOLO。最后,对图像重建损失函数进行了改进,提出两阶段的训练策略对整体网络进行统一训练。

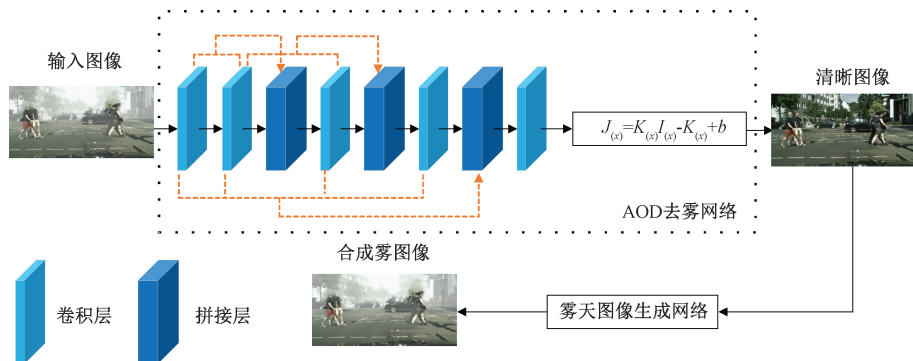


图1 CD-Net 网络结构

Fig. 1 Structure of CD-Net

2.2 循环去雾网络 CD-Net

如图1中 AOD 去雾网络所示,AOD 去雾网络包括 $K(x)$ 估计模块和清晰图像生成模块两个部分。 $K(x)$ 估计模块至关重要,因为它同时估计大气光照强度和透射率。通过结合不同尺寸的滤波器,利用5个卷积层生成多尺度特征。这种多尺度架构有效捕捉了不同尺度的特征,拼接层在卷积过程中减少了信息损失。此外,网络仅有5个卷积层和3个拼接层,因此 AOD 去雾网络十分轻量化。

雾天图像生成网络如图2所示,由浅层特征提取模块、残差密集特征融合模块、雾图像生成模块3个模块组成。网络输入为去雾后的清晰图像,输出为合成后的雾天图像。雾图像生成网络分别使用一个 1×1 卷积和一个 3×3 卷积提取浅层特征。残差密集特征融合模块由3个残差密集块^[18]、 1×1 卷积和 3×3 卷积组成。最后通过雾生成模块输出合成雾天图像。为了充分利用不同层次的特征,利用了残差连接和密集跳跃连接。

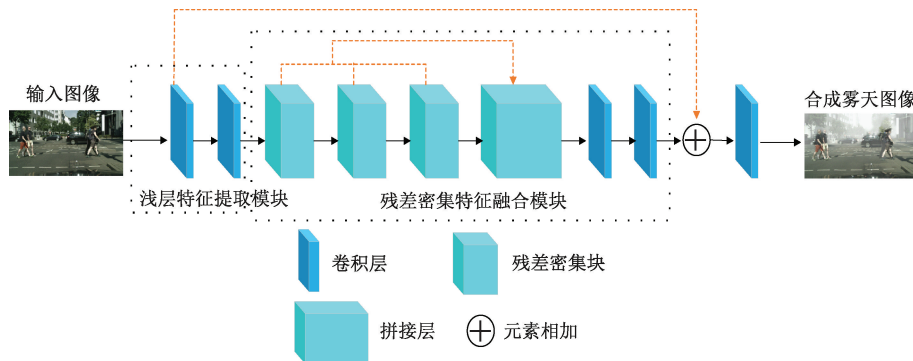


图2 雾天图像生成网络

Fig. 2 Structure of fog image generating network

两个子网络组成了从输入雾天图像到去雾清晰图像,再由去雾清晰图像到生成雾天图像的循环。网络训练时,两个子网络都参与训练,利用雾图像生成网络协助优化 AOD 去雾网络的去雾性能。在测试时,雾图像生成网络无需参与,因此轻量化的 CD-Net 有效地提升图像质量的同时又满足了实时性要求,且可以将其轻松地嵌入到任何检测框架中进行雾天目标检测任务。

2.3 YOLOv7 目标检测模型

YOLO 是一类快速目标检测算法,其主要特点是在一

定检测精度的基础上实现快速化和轻量化^[19]。从 v1 到 v7, YOLO 的检测精度和速度不断提高。YOLOV7 在模型结构(从 CSP 到 ELAN)、部分卷积策略方法(从普通卷积到重参数化卷积)和标签分配方法(从交并比、基于相似度的匹配方法到粗到细深度监督方法)方面与之前的版本有很大不同^[20]。YOLOv7 侧重于训练过程的优化,在不增加计算消耗的情况下提高准确率。与其他先进的目标检测方法相比,YOLOv7 极大地减少了模型的参数量和计算量。

YOLOv7 的网络结构如图 3 所示,在网络架构中提出了基于 ELAN 的扩展 ELAN(E-ELAN)。E-ELAN 在不破坏原有梯度路径的情况下,利用分组卷积扩展通道数提升网络的学习能力。YOLOv7 网络由输入、主干网络和头部网络三部分组成。主干网络用于提取特征,整个主干网络由若干个 CBS、E-ELAN 层和最大池化层交替组成,这

些层分别用于长宽比减半、通道数加倍以及特征提取。头部网络用于特征融合和预测,由若干个空间金字塔池化卷积层(SPPCPC)、卷积层、最大池化层、拼接层和重参数化 RepConv 层组成。在头部网络输出 3 个特征图后,分别通过 3 个重参数化 RepConv 层和 1×1 卷积输出 3 个不同尺寸的预测结果。

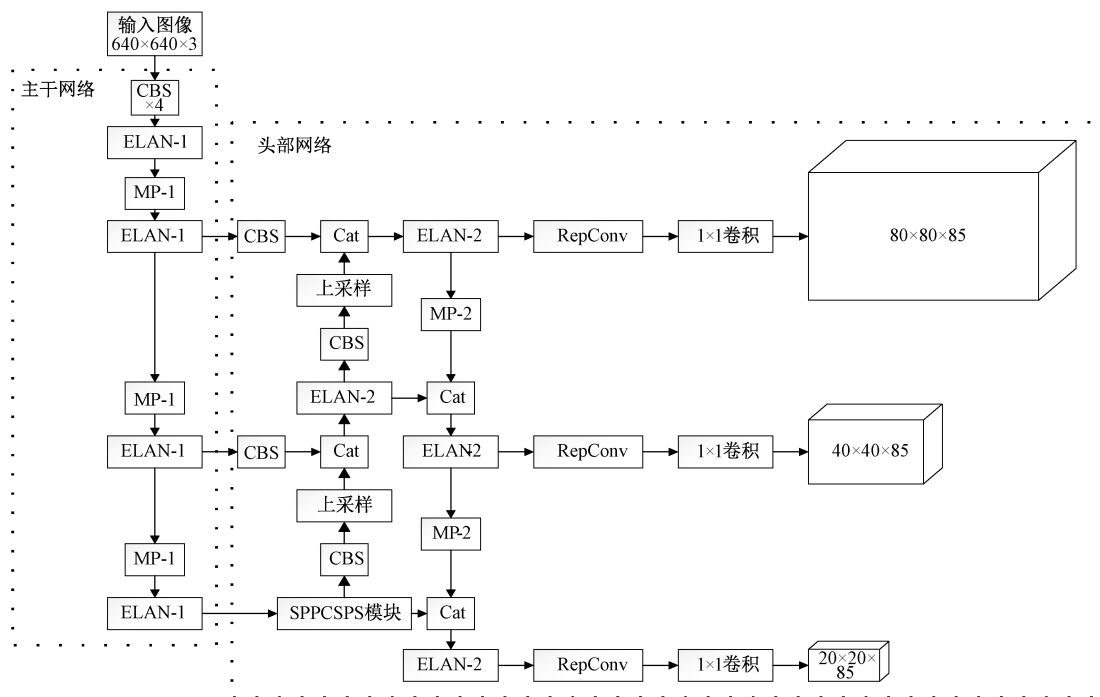


图 3 YOLOv7 网络结构

Fig. 3 Structure of YOLOv7

2.4 CD-YOLO 训练及损失改进

由于整个网络涉及到去雾任务和检测任务的结合,整体训练难度较大、模型难以收敛。对此,本文设计的训练方法分为两个阶段,第 1 个阶段单独训练 CD-Net,使其具备一定的去雾能力,YOLOv7 采用原作者提供的预训练权重;第 2 个阶段统一训练整体网络 CD-YOLO,这种训练方式更有利于结合底层去雾任务和高层检测任务。大多数图像复原或增强网络的训练都需要一组相对应的图像,退化图像作为输入,与之匹配的清晰图像作为标签,图像重建损失如下:

$$Loss1 = \frac{1}{N} \sum_{i=1}^N (R(x) - J(x))^2 \quad (6)$$

其中, J 是去雾模型预测的清晰图像, R 是标签清晰图像。而与之不同的是,本算法提出的 CD-Net 去雾模型包含雾天图像生成子网络,为了利用输入雾天图像中的潜在信息训练 CD-YOLO,改进的损失函数如下:

$$Loss2 = \frac{1}{N} \left[\sum_{i=1}^N (R(x) - J(x))^2 + \lambda \sum_{i=1}^N (F(x) - I(x))^2 \right] \quad (7)$$

其中, F 是模型通过预测清晰图像合成的雾天图像, λ 取 0.5。在第 2 阶段训练 CD-YOLO 时,CD-Net、YOLOv7 分别加载第 1 阶段的训练权重,在这一阶段损失函数仅使用正常的检测损失,整体网络仅朝检测精度更高的方向优化,与模型最终任务一致,不易陷入局部最优解,训练花费的时间更短。损失函数由三部分组成,分为回归损失(regression loss)、目标置信度损失(object confidence loss)和分类损失(classification loss)。其中,目标置信度损失和分类损失使用交叉熵损失,回归损失使用 CIoU 损失。

3 实验结果与分析

3.1 数据集与评价指标

本文在两个不同雾浓度的 CityScapes_foggy 数据集上进行对比实验,在 RTTS 数据集上进一步验证模型的泛化能力。CityScapes_foggy 是由 CityScapes 数据集^[21]合成的雾天图像数据集, CityScapes 数据集是一个专门针对城市街道场景的大规模数据集,它包含了从 50 个不同城市的街道场景中记录的各类图像。CityScapes 数据集分为精细标注和粗略标注两部分,精细标注中包含 5 000 幅图像,其中 2 975 幅训练图像、500 幅验证图像和 1 525 幅测试图

像。本文将 CityScapes 数据集中精细标注的 5 000 幅图像合成^[22]两个不同雾浓度的雾天图像数据集,为了区分分别命名为 CityScapes_foggy_low 和 CityScapes_foggy_high。CityScapes_foggy_low 是大气散射系数 $\beta=0.01$ 的低浓度雾天数据集,包括 Car、Bicycle、Person、Motorcycle、Bus 共 5 类目标,其数据集的划分与 CityScapes 数据集一致。CityScapes_foggy_high 是大气散射系数 $\beta=0.02$ 的中等浓度雾天数据集,除雾浓度不同之外,其余均与 CityScapes_foggy_low 一致。CityScapes、CityScapes_foggy_low 和 CityScapes_foggy_high 中同一幅图像对比效果如图 4 所示。RTTS 数据集是 RESIDE 数据集的子类数据集,包含 4 322 幅真实世界的雾天图像,主要涵盖交通和驾驶场景,目标类别与前文提到的两个数据集相同,本文仅在 RTTS 雾数据集上做验证实验。

由于雾天目标检测任务难度较大,故本文采用常用的 $\text{IoU} = 0.5$ 的平均精确度 AP50(average precision,AP)

作为精度评价指标,各类别 AP50 的平均值为 $\text{mAP}@50$ 。利用准确率(precision, P)和召回率(recall, R)得到 P-R 曲线,AP 即为 P-R 曲线的面积。P、R 和 AP 计算方式如下:

$$P = \frac{T_{TP}}{T_{TP} + F_{FP}} \times 100\% \tag{8}$$

$$R = \frac{T_{TP}}{T_{TP} + F_{FN}} \times 100\% \tag{9}$$

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

其中, T_{TP} 表示正确检测出的目标数量, F_{FP} 表示被错误检测成目标的数量, F_{FN} 表示目标被漏检的数量。将模型参数数量作为比较模型大小的指标。针对 CityScapes_foggy_low 数据集,计算每秒内检测出 640×640 大小的图像数量,即每秒帧数(frames per second,FPS)作为模型检测速度评价指标。



(a) CityScapes数据集样例 (b) CityScapes_foggy_low数据集样例 (c) CityScapes_foggy_high数据集样例
(a) Samples of CityScapes dataset (b) Samples of CityScapes_foggy_low dataset (c) Samples of CityScapes_foggy_high dataset

图 4 数据集样例

Fig. 4 Samples of dataset

3.2 实验配置

本文基于 Pytorch 深度学习框架,所有实验均在 64 位 Windows10 操作系统上进行,CPU 型号为 Intel Core i7-7800X,GPU 型号为 NVIDIA GeForce RTX 1080 Ti。在实验中,批量大小设置为 4,采用随机梯度下降法(stochastic gradient descent,SGD)对模型进行优化,初始学习率设置为 0.001,动量为 0.9,权重衰减为 0.000 5,在非极大值抑制时将阈值设置为 0.2,以去除重复检测框。将本文方法与当前具有代表性的通用目标检测算法

YOLOv5^[23]、YOLOv7、YOLOv9^[24]以及雾天图像检测算法 AOD-Net+Faster R-CNN、AOD-Net+YOLOv7 和 IA-YOLO 等进行比较。本文方法与所有对比方法,均训练 300 epochs。

3.3 对比实验

为了验证本文方法的有效性,首先将本文方法与上述介绍的几种目标检测算法在低雾浓度数据集 CityScapes_foggy_low 上进行对比实验,检测结果如表 1 所示。

表 1 本文算法与其他算法在 CityScapes_foggy_low 上的精度

Table 1 Precision comparison on CityScapes_foggy_low between the proposed algorithm and other algorithms								%
方法	Car	Bicycle	Person	Motorcycle	Bus	P	R	mAP@50
YOLOv5-L	74.0	48.6	53.4	47.0	57.9	76.0	50.7	56.2
YOLOv7	79.4	54.5	63.6	53.1	66.2	76.9	56.6	63.4
YOLOv9-C	76.0	50.8	57.9	51.0	63.2	76.4	52.0	59.8
AOD-Net+Faster R-CNN	64.4	39.5	40.1	28.4	54.4	63.9	53.1	45.4
AOD-Net+YOLOv7	79.8	55.6	64.3	54.7	67.1	78.2	57.2	64.3
IA-YOLO	63.6	44.0	44.1	26.3	33.8	59.4	56.4	42.4
CD-YOLO	80.3	56.1	65.7	55.6	68.4	78.6	58.4	65.2

YOLOv5、YOLOv7 和 YOLOv9 检测精度分别为 56.2%、63.4%和 59.8%。这些算法均为当前先进的通用目标检测算法,在雾天数据集上对通用检测方法微调难以取得令人满意的结果。雾天检测方法中 IA-YOLO 的精度较低,其原因一方面是 IA-YOLO 的检测网络是 YOLOv3,相对于更先进的 YOLO 检测方法精度存在一定差距;另一方面是 IA-YOLO 中的 DIP 模块是基于传统的去雾、锐化等方法增强输入图像,效果有限,且手工设计、固定排序的方式导致 DIP 模块性能受限。对于另一个雾天检测方法,将 AOD-Net 分别与 Faster R-CNN 和 YOLOv7 组成两个模型进行实验,可以看到检测网络对整体精度的影响很大;AOD-Net+YOLOv7 的精度较 YOLOv7 有一定提升但幅度有限,主要原因是轻量化的 AOD-Net 难以对 $K(x)$ 进行准确估计,去雾效果不佳。基准方法 YOLOv7 的检测精度为 63.4%,本文提出的 CD-YOLO 的检测精度为 65.2%,比基准算法高 1.8%。

表 2 统计了 YOLOv5、YOLOv7、YOLOv9、AOD-Net+YOLOv7、IA-YOLO 和本文方法 CD-YOLO 的参数数量和 FPS。上述方法中 IA-YOLO 模型最大,IA-YOLO 的参数数量大部分来自其检测框架 YOLOv3,还有小部分源于参数预测器 CNN-PP 和可微分图像处理模块 DIP,推理速度与其他方法有较大差距。基准算法 YOLOv7 的参数数量为

36.5 M,本文提出的 CD-YOLO 模型仅比基准算法大 0.6 M。在检测速度方面,一般认为人眼视觉的时间特性最低要求为 15 fps,当 FPS >15 时,人眼认为检测流畅即能实时检测。表 2 中,在处理 640×640 大小的图像时除 IA-YOLO 的 FPS 小于 15 之外,其他方法都能实现实时检测。实时检测方法中 CD-YOLO 的精度最高,其每秒能处理 42.4 帧图像。

表 2 本文算法与其他算法的模型大小及检测速度对比

Table 2 Speed and model size comparison between the proposed algorithm and other algorithms

方法	图像大小	参数量/M	FPS
YOLOv5-L	640×640	46.1	29.6
YOLOv7	640×640	36.5	49.5
YOLOv9-C	640×640	53.0	43.5
AOD-Net+YOLOv7	640×640	36.5	42.6
IA-YOLO	640×640	62.7	13.5
CD-YOLO	640×640	37.1	42.4

为了进一步评估本文算法的有效性,将所有方法在雾浓度更大的 CityScapes_foggy_high 数据集上进一步验证,结果如表 3 所示。

表 3 本文算法与其他算法在 CityScapes_foggy_high 上的精度

Table 3 Precision comparison on CityScapes_foggy_high between the proposed algorithm and other algorithms %

方法	Car	Bicycle	Person	Motorcycle	Bus	P	R	mAP@50
YOLOv5-L	71.4	45.1	50.8	44.9	57.1	76.9	47.4	53.9
YOLOv7	78.2	55.3	60.4	50.9	64.5	78.0	54.6	61.9
YOLOv9-C	75.2	52.1	56.3	47.4	61.2	79.7	49.8	58.4
AOD-Net+Faster R-CNN	63.9	38.7	39.9	27.2	50.1	63.1	52.4	44.0
AOD-Net+YOLOv7	79.3	55.9	62.8	52.1	65.4	80.6	54.2	63.1
IA-YOLO	62.5	31.9	42.4	24.4	31.4	57.1	55.4	38.5
CD-YOLO	80.4	56.3	63.5	53.4	67.3	80.7	56.1	64.2

CityScapes_foggy_high 数据集的雾浓度更重,识别难度更大,因此整体上各检测方法在 CityScapes_foggy_high 数据集上的精度有所下降。YOLOv5、YOLOv7 和 YOLOv9 三个通用目标检测方法的检测精度分别为 53.9%、61.9%和 58.4%;雾天目标检测方法中,IA-YOLO 的精度仅有 38.5%,AOD-Net+Faster R-CNN 和 AOD-Net+YOLOv7 的检测精度分别为 44.0%、63.1%。本文提出的 CD-YOLO 方法精度为 64.2%,超过了对比方法中的其他所有算法。

为了进一步验证模型的泛化能力,将 YOLOv5、YOLOv7、YOLOv9、AOD-Net+YOLOv7、IA-YOLO 和 CD-YOLO 在 RTTS 数据集上进行验证实验,结果如表 4 所示。

表 4 本文算法与其他算法在 RTTS 上的验证精度

Table 4 Precision comparison on RTTS between the proposed algorithm and other algorithms %

方法	P	R	mAP@50
YOLOv5-L	71.5	30.2	34.7
YOLOv7	68.8	37.8	41.9
YOLOv9-C	62.2	39.8	41.7
AOD-Net+YOLOv7	69.6	38.5	42.9
IA-YOLO	54.2	25.4	26.8
CD-YOLO	72.7	40.1	43.6

由表 4 可知,本文提出的 CD-YOLO 在 RTTS 数据集上的验证精度超过了 YOLOv5、YOLOv7、YOLOv9、

AOD-Net+YOLOv7 和 IA-YOLO 等通用检测方法和雾天检测方法,CD-YOLO 的泛化能力最佳。

为了更直观地展现出本文算法的优越性,给出 YOLOv5、YOLOv7、YOLOv9、AOD-Net+YOLOv7、IA-YOLO 和本文方法 CD-YOLO 在雾浓度更大的 CityScapes_foggy_high 上的检测结果,如图 5 所示。

图 5(a)是将真实框标注在对应的清晰图像上,为了进一步观察对比,将图 5(a)中待检测目标聚集区域放大得到图 5(b),其他方法检测结果图仅展示放大部分,图 5(c)~(e)分别是通用目标检测方法 YOLOv5、YOLOv7、YOLOv9 的检测结果图,图 5(f)、(g)分别是雾天目标检测方法 AOD-Net+YOLOv7、IA-YOLO 的检测结果图,图 5(h)是本文方法 CD-YOLO 的检测结果图。从图 5 中各算法检测结果和真实标注框对比,可以看出,YOLOv5、YOLOv9、AOD-Net+YOLOv7 和 IA-YOLO 并未识别出待测图中最左侧的 person 类目标,YOLOv7 和 IA-YOLO 存在对 car 类目标漏检的情况;此外,YOLOv9 和 AOD-Net+YOLOv7 出现了对 car 类目标错误识别的现象。本文提出的 CD-YOLO 正确地识别出每一个目标,没有漏检和误检现象,模型检测结果最佳。

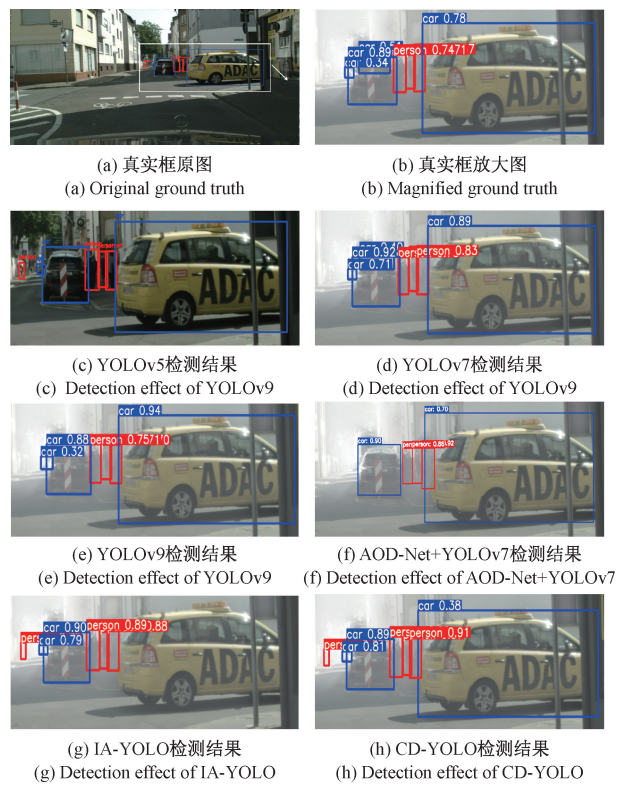


图 5 检测算法结果对比

Fig. 5 Comparison of the effect between each algorithm

3.4 消融实验

为了评估改进的 CD-YOLO 中 CD-Net 循环去雾网络的有效性,本节在 CityScapes_foggy_low 数据集上设计消

融实验,在 YOLOv7 网络模型的基础上,依次加入 AOD 去雾网络和雾天图像生成网络,结果如表 5 所示。

由表 5 可知,基准算法在 CityScapes_foggy_low 数据集上的精度为 63.4%。在基准方法的基础上增加 AOD 去雾网络,精度为 64.3%,相比基准算法精度增加了 0.9%,表明其在重建图像、提取特征上发挥了重要作用,为模型性能的改善做出了积极贡献。最后在加入 AOD 去雾子网的基础上,加入雾图像生成子网组成循环去雾网络,模型精度为 65.2%,较添加 AOD 去雾网络的算法精度提升 0.9%,说明雾图像生成子网能够帮助 AOD 去雾子网更精准地重建图像,进一步提升模型的识别性能。

表 5 在 CityScapes_foggy_low 上的消融实验结果			
Table 5 Ablation experiments on CityScapes_foggy_low			
YOLOv7	AOD 去雾子网络	雾图像生成子网络	mAP@50/%
✓			63.4
✓	✓		64.3
✓	✓	✓	65.2

4 结 论

本文针对雾天场景下图像受天气因素干扰大、细节信息缺失,使用深度学习通用目标检测算法易出现精度低、检测速度慢等问题,提出了一种基于 YOLOv7 的雾天实时目标检测算法。该算法以 YOLOv7 为基线,在其前端设计包含 AOD 去雾子网络和雾天图像生成子网络的循环去雾双子网络。通过雾图像生成网络协助优化 AOD 去雾网络的去雾性能,提升整体网络的特征提取能力;在测试时,雾图像生成网络无需参与,显著减少了模型推理时间;并且引入包含模糊图像信息的损失函数和采用两阶段统一训练方法,有效地结合去雾和检测任务。通过在两个合成的不同雾浓度数据集上的实验表明,改进后的算法提升了雾天图像目标检测整体性能,在准确率和速度上均优于其他对比方法,模型泛化能力更强。

下一步将对 YOLOv7 算法网络结构进行改进工作,提升其对雾天目标检测的适应能力,进一步提升整体模型的检测精度。

参考文献

[1] 宋绍京, 陆婷婷, 孙翔, 等. 面向自动驾驶的多任务环境感知算法[J]. 电子测量技术, 2023, 46(24): 157-163.
SONG SH J, LU T T, SUN X, et al. Multi-task environment awareness algorithm for autonomous driving[J]. Electronic Measurement Technology, 2023, 46(24): 157-163.
[2] 宋建辉, 刘鑫, 庄爽, 等. 面向无人驾驶的多任务环境感知算法研究[J]. 电子测量与仪器学报, 2025, 39

- (1):122-132.
- SONG J H, LIU X, ZHUANG SH, et al. Research on multi-task environment perception algorithm for unmanned dri-ving [J]. Journal of Electronic Measurement and Instrumentation, 2025, 39 (1): 122-132.
- [3] 吴成茂. 直方图均衡化的数学模型研究[J]. 电子学报, 2013, 41(3): 598-602.
- WU CH M. Research on mathematical model of histogram equalization[J]. Acta Electronica Sinica, 2013, 41(3): 598-602.
- [4] HE K M, SUN J, TANG X OU. Single image haze removal using dark channel prior [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2011, 33(12): 2341-2353.
- [5] DONG Y SH, SUN J. Proximal dehaze-net: A prior learning-based deep network for single image dehazing[C]. European Conference on Computer Vision, 2018: 702-717.
- [6] LI B Y, PENG X L, WANG ZH Y, et al. Aod-net: All-in-one dehazing network[C]. IEEE International Conference on Computer Vision, 2017: 4770-4778.
- [7] HUANG S CH, LE T H, JAW D W. DSNET: Joint semantic learning for object detection in inclement weather conditions[J]. IEEE Transaction on Pattern Analysis and Machine Intelligence, 2020, 43 (8): 2623-2633.
- [8] 梁天添, 杨淞淇, 钱振明. 基于改进 YOLOv8s 的恶劣天气车辆行人检测方法[J]. 电子测量技术, 2024, 47(9): 112-119.
- LIANG T T, YANG S Q, QIAN ZH M. Vehicle pedestrian detection method in bad weather based on improved YOLOv8s [J]. Electronic Measurement Technology, 2024, 47(9): 112-119.
- [9] QIU Y SH, LU Y Y, WANG Y T, et al. IDOD-YOLOV7: Image-dehazing YOLOV7 for object detection in low-light foggy traffic environments[J]. Sensors, 2023, 23(3): 1347.
- [10] LIU W Y, REN G F, YU R SH, et al. Image-adaptive YOLO for object detection in adverse weather conditions [C]. AAAI Conference on Artificial Intelligence, 2022, 36(2): 1779-1800.
- [11] FARHADI A, REDMON J. YOLOv3: An incremental improvement[C]. Computer Vision and Pattern Recognition, Berlin/Heidelberg, Germany: Springer, 2018, 1804: 1-6.
- [12] WANG CH Y, BOCHKOVSKIY A, LIAO H Y. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 7464-7475.
- [13] MCCARTNEY E J. Optics of the atmosphere: Scattering by molecules and particles[C]. New York: John Wiley and Sons, 1976.
- [14] NARASIMHAN S G, NAYAR SH K. Chromatic framework for vision in bad weather [C]. IEEE Conference on Computer Vision and Pattern Recognition, 2000, 1: 598-605.
- [15] NARASIMHAN S G, NAYAR SH K. Vision and the atmosphere[J]. International Journal of Computer Vision, 2002, 48(3): 233-254.
- [16] CHEN Z H, ZHOU Y, LI P, et al. DCNet: Dual-task cycle network for end-to-end image dehazing[C]. IEEE International Conference on Multimedia and Expo, 2021: 1-6.
- [17] GAO CH X, CHEN L J, ZHAO C D, et al. Lightweight image dehazing algorithm based on detail feature enhancement [C]. 2023 26th International Conference on Computer Supported Cooperative Work in Design, 2023.
- [18] ZHANG Y L, TIAN Y P, KONG Y, et al. Residual dense network for image restoration [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021: 2480-2495.
- [19] LYU H H, YAN H B, LIU K Y, et al. YOLOv5-ac: Attention mechanism-based lightweight YOLOv5 for track pedestrian detection [J]. Sensors, 2022, 22 (15): 5903.
- [20] HUSSAIN M, AL-AQRABI H, MUNAWAR M, et al. Domain feature mapping with YOLOv7 for automated edge-based pallet racking inspections[J]. Sensors, 2022, 22(18): 6927.
- [21] CORDTS M, OMRAN M, RAMOS S, et al. The cityscapes dataset for semantic urban scene understanding [C]. IEEE Conference on Computer Vision and Pattern Recognition, 2016: 3213-3223.
- [22] 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.
- [23] JOCHER G. YOLOv5 release v6.1 [EB/OL]. <https://github.com/ultralytics/yolov5/releases/tag/v6.1>, 2022.
- [24] WANG CH Y, YEH I H, LIAO H Y, et al. YOLOv9: Learning what you want to learn using programmable gradient information [C]. European Conference on Computer Vision. Cham: Springer, 2025: 1-21.

作者简介

谢祖华, 硕士研究生, 主要研究方向为深度学习、目标检测。

E-mail: 277589912@qq.com

李海涛, 硕士研究生, 主要研究方向为深度学习、目标检测。

E-mail: 2650941613@qq.com

胡建文(通信作者), 副教授, 主要研究方向为图像处理、深度学习。

E-mail: 117780631@qq.com