

doi:10.11835/j.issn.1000-582X.2020.027

雾霾天气车辆检测的改进 YOLOv3 算法

黄开启, 刘小荣, 钱艳群, 黄茂云

(江西理工大学 电气工程与自动化学院, 江西 赣州 341000)

摘要:针对传统 YOLOv3 算法中存在检测框定位不精确的问题,提出了一种改进的 YOLOv3 算法用来重新估计检测框位置,提高智能汽车在雾霾交通环境下的定位精度。首先运用图像去雾算法对采集到的图片进行预处理,然后构造定位置信度替代分类置信度作为参考项来选择估计检测框位置,并改进非极大值抑制(NMS)算法,引入软化非极大值抑制(soft-NMS),最后使用加权平均的方式来更新坐标位置,以达到提高定位精度的目的。实验结果表明,先经过单尺度 retinex 去雾算法处理图片,再通过改进的 YOLOv3 算法进行车辆检测,与使用原始的 YOLOv3 算法进行检测相比平均精度均值 mAP (mean average precision)提高了 0.44%,在满足检测实时性的同时,能够检测到更多的目标,对检测车辆的定位也更加精确。

关键词:检测框;去雾算法;车辆检测;改进 YOLOv3;定位置信度

中图分类号:TP391.4

文献标志码:A

文章编号:1000-582X(2021)12-095-08

Improved YOLOv3 algorithm for vehicle detection in hazy weather

HUANG Kaiqi, LIU Xiaorong, QIAN Yanqun, HUANG Maoyun

(School of Electrical Engineering and Automation, Jiangxi University of Science and Technology, Ganzhou 341000, Jiangxi, P. R. China)

Abstract: Traditional YOLOv3 algorithm has the problem of inaccurate detection frame positioning. In this paper, an improved YOLOv3 algorithm was proposed to re-estimate the detection frame position and improve the positioning accuracy of smart cars in hazy traffic environment. Firstly, the collected images were preprocessed using image dehazing algorithms. Then, to improve positioning accuracy, estimated test box position was selected by constructing positioning confidence as the reference to replace classification confidence, and the non-maximum suppression (NMS) algorithm was improved. Finally, the soft-NMS was introduced, and the coordinates were updated using the weighted average. The experimental results show that compared with the original YOLOv3 algorithm, the mAP(mean average precision) of improved YOLOv3 algorithm increased by 0.44%, which could detect more targets and locate the detected vehicles more accurately in real-time detection.

收稿日期:2020-07-31 **网络出版日期:**2020-09-11

基金项目:国家自然科学基金资助项目(61164009)。

Supported by National Natural Science Foundation of China (61164009).

作者简介:刘小荣(1992—),女,江西理工大学硕士研究生,研究方向为深度学习与模式识别,(E-mail)1808973401@qq.com。

通讯作者:黄开启(1969—),男,江西理工大学教授,研究领域为新能源汽车控制技术与机器人控制技术等,(E-mail)kaiqi.huang@163.com。

Keywords: detection frame; dehazing algorithm; vehicle detection; improved YOLOv3; positioning confidence

高级驾驶辅助系统(ADAS)旨在帮助驾驶员完全避免事故,进而大幅减少道路交通事故。车辆检测是 ADAS 的重要组成部分,在检测中准确知道并精准定位目标车辆,这对避免交通事故的发生起着极其重要的作用。

传统车辆检测算法采用滑动窗口的策略对整幅图像进行遍历,判断车辆可能出现的位置,接着人工选取某个特征,对目标进行识别。例如刘家旭等^[1]提出基于方向梯度直方图(HOG)特征的检测算法,周行等^[2]提出基于 HOG 特征和局部二值模式(LBP)特征融合检测算法。传统算法主要存在两个问题:一是基于滑动窗口的区域选择没有针对性,冗余窗口太多;二是由于目标的形态、光照、背景等多样性,手工设计的特征对于多样性的变化鲁棒性较差。随着深度学习的快速发展,卷积神经网络在计算机视觉领域取得了巨大成功,目前基于卷积神经网络的目标检测算法主要分为两类:一类是以 YOLO^[3]、SSD^[4]、RefineDet^[5] 等为代表的分类回归方法,另一类是以 R-CNN^[6]、Fast R-CNN^[7]、Faster R-CNN^[8] 等为代表的基于候选框的两阶段方法。其中 YOLOv3^[9] 作为性能较为出色的目标检测算法被广泛应用到车辆检测任务中。冯加明等^[10]提出通过删减部分残差以降低卷积层通道数并通过 k-均值聚类与手动调节相结合的方式提高检测的精度;杜金航等^[11]、崔文靓等^[12]和胡贵桂^[13]利用 k-均值重新聚类选取锚点预测边界框,提高检测速度与定位精度;吴仁彪等^[14]提出利用卷积神经网络的去雾算法对采集的图片进行预处理。汪昱东等^[15]通过在检测网络中加入物浓度判断模块来提高网络的适应性和鲁棒性。

以上方法主要存在以下两个问题:一是这些方法大多都是通过 k-均值重新聚类锚框的方式侧向提高定位准确性,而分类置信度与定位精度并不是强关联,所以它不能够保证检测框的可靠性;二是目标检测的场景均为晴朗的白天,少有针对雾霾天气进行检测的研究。针对问题一,以 YOLOv3 为基础网络,加入正态分布模型,将标准差作为定位置信度用来定位,并改进非极大值抑制,引进软化非极大值抑制降低漏检率从而提高定位精度。针对问题二,雾霾天气存在图像退化严重,对比度低,细节信息有损失等问题,严重影响室外视觉检测,笔者添加图像去雾预处理操作,采用单尺度 retinex 去雾算法对原图进行图像增强^[16],以提高图像对比度。

1 图像去雾处理

1.1 去雾算法

由于雾与霾有着本质的不同,霾的物理模型更加复杂,基于大气散射模型的图像复原方法不适用于雾霾环境下的去雾处理。针对此问题本研究中通过图像增强的方式共选取了单尺度 retinex 去雾算法、多尺度 retinex 去雾算法以及 McCann retinex 算法 3 种算法对图像进行去雾处理并对去雾效果进行对比。

1.2 去雾效果对比

首先采用单尺度 retinex 算法(single scale retinex, SSR)、多尺度 retinex 算法(multi scale retinex, MSR)和迭代 retinex 算法(McCann retinex)分别对图像进行增强处理,增强效果如下图 1 所示,可以看出 SSR 算法表现更好,所以本文中采用 SSR 算法对图像进行去雾处理。

2 YOLOv3 及其改进方式

2.1 YOLOv3 框架

YOLOv3 采用了全新的网络架构 Darknet-53(图 2),它包含了大量的 3×3 以及 1×1 卷积层,通过引用残差模块^[17]解决随着网络层数加深模型在训练过程中产生梯度爆炸以及梯度消失的问题。YOLOv3 使用了上采样操作,并将大特征图和小特征图上采样后的特征图进行拼接,使网络能够拥有既包含丰富的高层抽象特征又包含精确的位置信息特征的融合特征层。YOLOv3 借鉴了 FPN(金字塔网络)^[18]的思想,在 3 种不

同的规模上进行预测,检测结果分别会有 3 个不同尺度的输出,它包含了目标的坐标位置,目标是正样本还是负样本,目标属于那个类别的置信度,对于每个尺度的分支而言,在每个网格中会预测出 3 个结果(因为每个尺度下会有 3 个锚框)。最后将 3 个尺度的结果合并进行非极大值抑制(NMS)后输出最终的检测结果。



图 1 图像增强效果对比图

Fig. 1 Contrast of image enhancement effect

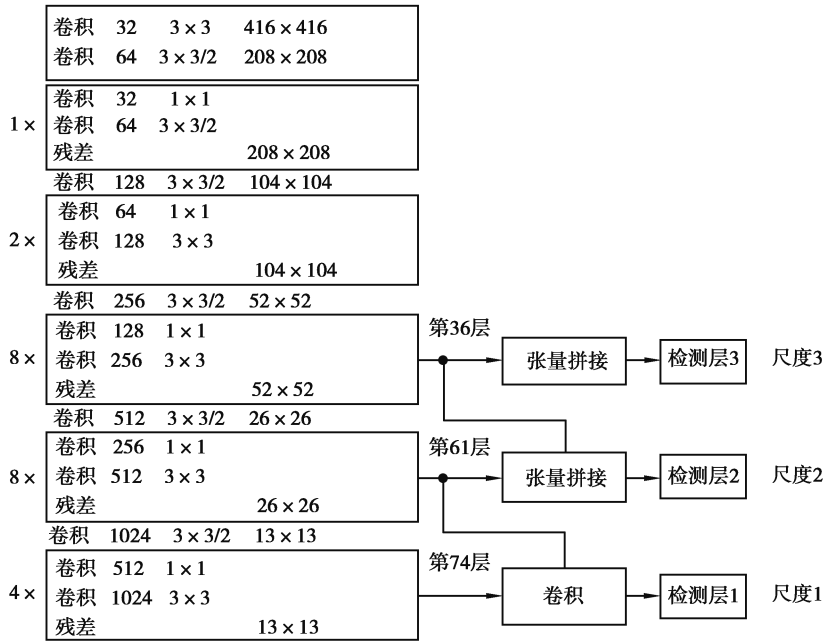


图 2 YOLOv3 网络结构图

Fig. 2 YOLOv3 network structure diagram

2.2 网络输出模型改进

原始算法中目标检测框的输出是以分类置信度为依据的,将拥有高的分类分数的候选包围框作为结果输出,实际上分类置信度与定位精度并不是绝对地成正比关系,选用分类置信度为依据并不能保证检测框的正确可靠。原始算法中目标框只有位置信息而没有概率值,检测框的输出无法以定位置信度作为依据。本研究中提出通过正态分布来对网络输出进行建模,将通过正态分布产生的标准差作为定位置信度。

引入正态分布后增加了网络的输出维度,坐标预测输出由原来的 t_x 、 t_y 、 t_w 、 t_h (t_x 、 t_y 表示边界框中心偏移量, t_w 、 t_h 表示宽高缩放比)共 4 个维度变为 $\hat{\mu}_{t_x}$ 、 $\hat{\sigma}_{t_x}$ 、 $\hat{\mu}_{t_y}$ 、 $\hat{\sigma}_{t_y}$ 、 $\hat{\mu}_{t_w}$ 、 $\hat{\sigma}_{t_w}$ 、 $\hat{\mu}_{t_h}$ 和 $\hat{\sigma}_{t_h}$ 共 8 个维度,其中 μ 代表均值, σ 代表标准差。对这 8 个维度参数做以下变换:

$$\sigma_{t_x} = \sigma(\hat{\sigma}_{t_x}), \sigma_{t_y} = \sigma(\hat{\sigma}_{t_y}), \sigma_{t_w} = \sigma(\hat{\sigma}_{t_w}), \sigma_{t_h} = \sigma(\hat{\sigma}_{t_h}). \quad (1)$$

$$\mu_{t_x} = \sigma(\hat{\mu}_{t_x}), \mu_{t_y} = \sigma(\hat{\mu}_{t_y}), \mu_{t_w} = \hat{\sigma}_{t_w}, \mu_{t_h} = \hat{\sigma}_{t_h}. \quad (2)$$

$$\sigma(x) = \frac{1}{1 + \exp(-x)}. \quad (3)$$

与原 YOLOv3 算法一样,这里需要对 μ_{t_x} 和 μ_{t_y} 进行 sigmoid 操作,将值限定在(0,1)范围内。如果不做约束,预测的边界框容易向任意方向偏移,会导致每个位置预测的边界框可能落在图片的任意位置,这就会导致模型训练不稳定,很难得到正确的偏移量; μ_{t_w} 和 μ_{t_h} 的处理与 YOLOv3 相同,不需要通过 sigmoid 操作,因为长宽的尺度变化可能为 1。此外,新增的标准差 σ_{t_x} 、 σ_{t_y} 、 σ_{t_w} 和 σ_{t_h} 也需要通过 sigmoid 将值限定在(0,1)范围内。标准差表明了坐标的可靠程度,0 表示可靠,1 表示不可靠,这是因为正态分布中,方差越大,分布的变化越大,包围框的估计结果越不可靠。

原始的 YOLOv3 进行边界框回归时,由于网络预测输出就是坐标本身,计算梯度时就利用了均方误差的方式。本研究中改进的 YOLOv3 输出的是均值和方差,所以使用了负对数似然损失函数。改进的边界框损失函数 L 为:

$$L = -\log\left(\prod_{i=1}^m \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x_i - \mu_i)^2}{2\sigma^2}\right)\right) = m \log \sigma + \frac{m}{2} \log(2\pi) + \frac{1}{2\sigma^2} \sum_{i=1}^m (x_i - \mu_i)^2, \quad (4)$$

式中 m 为网格数量。

2.3 NMS 改进

原算法中选择检测框时使用 NMS 保留得分最高的包围框并将与当前预测框重叠较多的预测框视作冗余,即将所有 IOU(intersection-over-union,交并比)大于阈值的框全部删除。但是对稠密物体进行检测时,如果同类的两个目标距离较近,检测时有可能出现误以为是同一个物体而造成漏检的问题。所以针对这个问题对原始的 NMS 进行了改善。不是将 IOU 大于阈值的框删除,而是将 IOU 大于阈值的框通过高斯加权,降低其置信度。改进后的函数如下:

$$s_i = s_i \exp\left(-\frac{\text{IOU}(M, b_i)^2}{\sigma}\right), \forall b_i \notin D. \quad (5)$$

式中: s_i 为第 i 个检测框的高斯权重函数; M 为当前得分最高框, b_i 为待处理框, b_i 和 M 的 IOU 越大,最后得到的分类分数就越低; D 为放最终检测框的集合。

预测框的输出要以定位置信度为依据,对所有与 M 的 IOU 大于一定重叠度阈值的候选包围框使用加权平均操作更新坐标位置,从而达到提高定位精度的目的。公式如下:

$$x = \frac{\sum p_i x_i}{\sum \frac{p_i}{\sigma_{x,i}^2}}. \quad (6)$$

改进后的 YOLOv3 检测数据如表 1 所示。

表 1 单类别目标 IOU 检测结果对比

Table 1 Comparison of IOU detection results of single category of targets

方法	IOU					
	平均值	小型汽车	自行车	摩托车	行人	公共汽车
YOLOv3	0.57	0.57	0.63	0.59	0.53	0.55
改进 YOLOv3	0.60	0.58	0.61	0.66	0.58	0.57

改进后的 YOLOv3 检测效果如图 3 所示。

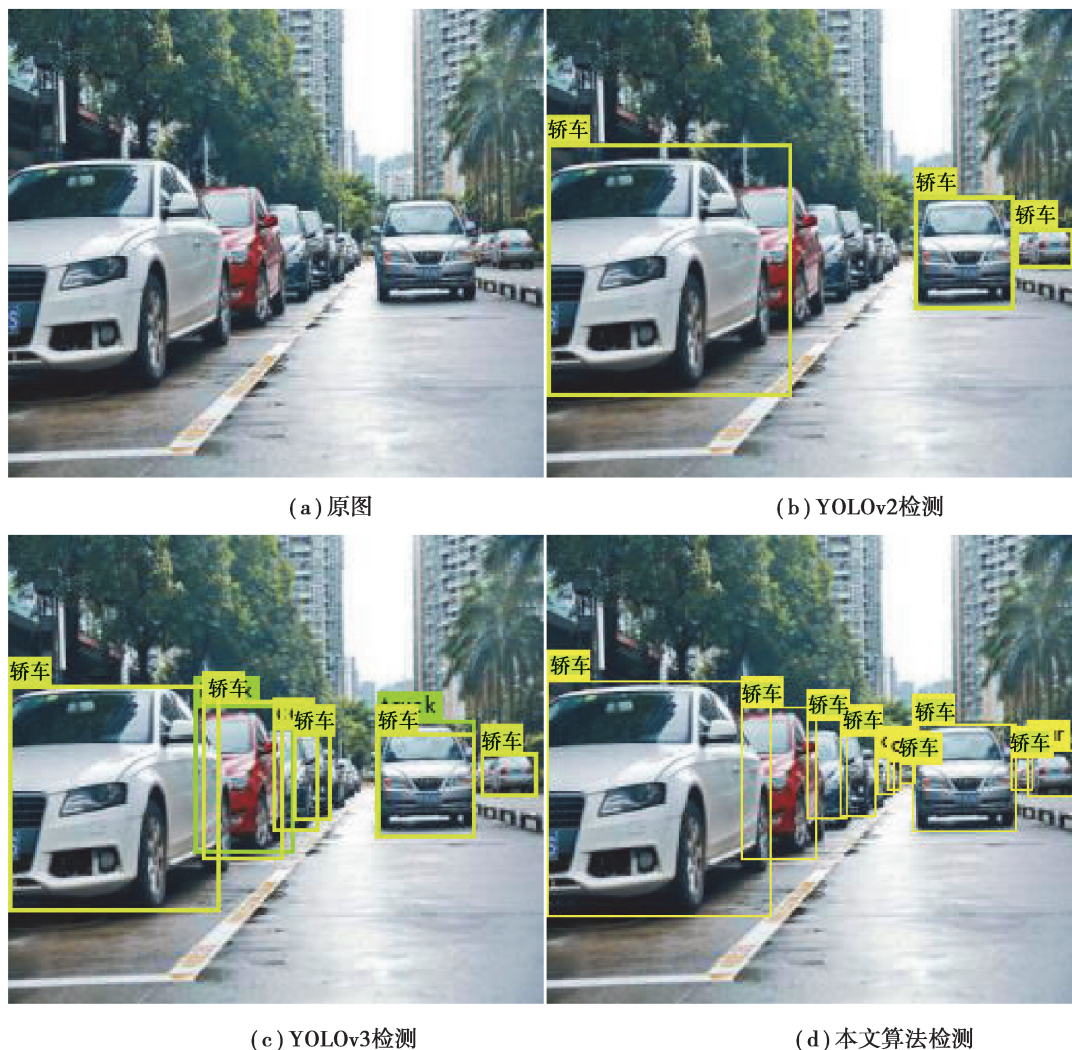


图 3 图像检测效果对比图

Fig. 3 Comparison of image detection effects

由表 1 可以看到改进的 YOLOv3 与原算法相比 IOU 提高了 0.03,表明改进后的算法在定位精度上有所提升。由图 3 可以看到改进的算法与 YOLOv2 算法相比可以检测出更多目标车辆,与原 YOLOv3 算法相比,检测框包围目标更加精确。由图 3 对比可以看到第一辆白色轿车的检测框,很明显改进的 YOLOv3 算法在定位上更加精确。原始的 YOLOv3 算法中多目标聚集在一起的时候会出现了误检的情况,如上图中的左边第 2 辆轿车,原 YOLOv3 算法中出现了 2 个检测框,显示此处有 2 辆车,很明显这是误检,而通过改进的 YOLOv3 算法准确识别出了它是一辆车。经过多次连续检测,改进前后的算法在检测速度上基本没有改变,检测一张图片的时间基本都在 0.19 s 左右,这表明了改进后的 YOLOv3 算法没有影响检测的速度。

3 实验

实验测试在 darknet 框架下进行,训练及测试的计算机硬件配置为 Intel(R) Core(TM) i5-8625U CPU@1.60GHz,GPU 为 NVIDIA GeForce MX250,操作系统为 Ubuntu18.04;mAP 的计算采用 voc2007 的方式,阈值设置为 0.5。

3.1 实验数据集

本实验中用于目标图像检测的标准化数据集是从 RESIDE 数据集中抽取的有雾图片与清晰图片共计

300 张作为训练集,图片主要针对交通道路上容易出现的人、小型汽车、摩托车、自行车以及大型公共汽车 5 大类进行训练。整个训练过程使用批量随机梯度下降法来优化损失函数,训练大约迭代了 55 000 次,实验动量系数设为 0.9,初始学习率设为 0.01,权重衰减值设为 0.000 5,批量大小设为 32。由图 4 可以看出迭代至 40 000 次之后,损失函数明显下降,40 000 次之后下降逐渐变得缓慢,最后损失函数稳定在 0.5 左右。

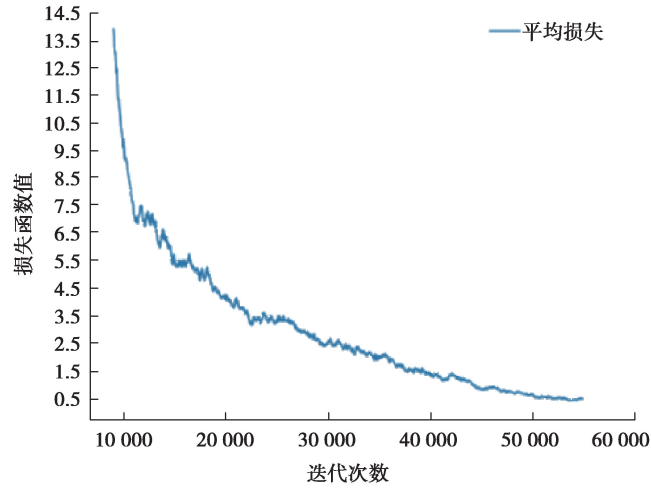


图 4 损失函数图

Fig. 4 Loss function diagram

3.2 检测结果与分析

为了验证本文算法的有效性,首先将改进的 YOLOv3 结合 NMS 算法与原 YOLOv3 算法进行对比,然后再分别将不同去雾处理算法结合本研究中改进的目标检测算法与原 YOLOv3 算法进行对比,实验结果如表 2 所示。由表 2 可以得出以下结论:在不改变分类精度的前提下通过引入定位置信度以后,目标的定位精度提高了 0.1%;经过本文算法-SSR 检测得到的平均精度相对于原 YOLOv3 模型提升了 0.44%,说明本文设计的检测算法较原 YOLOv3 算法检测效果更好。

表 2 不同方法在数据集上的检测结果

Table 2 Detection results of different methods on the data set

方法	mAP/%					
	平均值	小型汽车	自行车	摩托车	行人	公共汽车
YOLOv3	48.12	47.52	48.75	48.56	47.93	47.84
改进 YOLOv3-NMS	48.22	47.68	48.63	48.88	48.12	47.79
本文算法-MSR	48.23	46.98	49.56	47.44	50.26	46.89
本文算法-迭代 retinex	48.17	48.23	48.57	46.74	47.95	49.35
本文算法-SSR	48.56	49.64	48.11	48.05	49.74	47.26

为了测试实际雾天检测效果,设计了针对 3 种去雾算法的实际效果对比实验,实验得到的检测效果如图 5 所示。从图中可以看出,本研究中设计的改进算法与原算法相比,在雾霾天气情况下可以检测到更多目标,目标定位也更加精确。

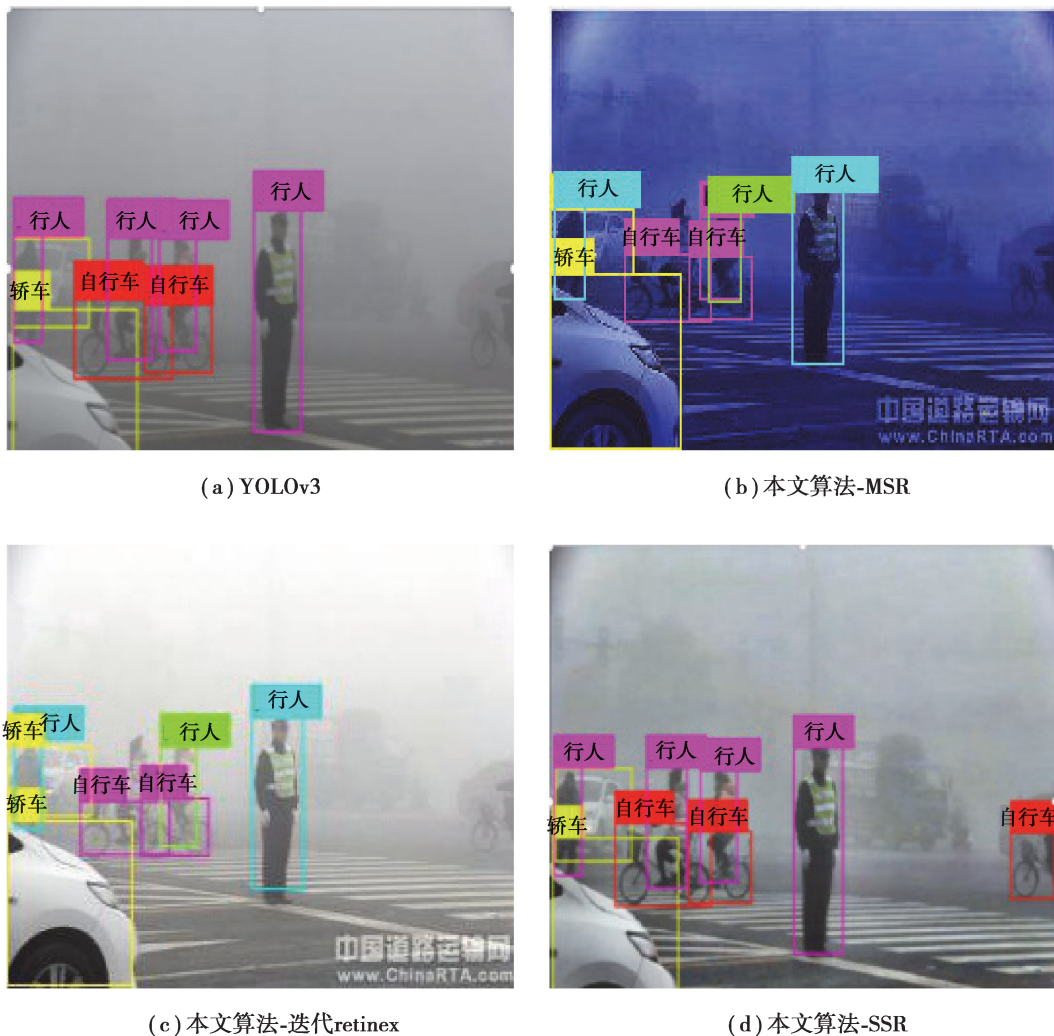


图 5 雾霾天气图像检测效果对比图

Fig. 5 Comparison of haze-weather image detection effects

4 结 语

为了提高雾霾天气车辆定位的准确性,通过改进 YOLOv3 的定位方法,使用标准差替代分类分数作为定位置信度,并通过单尺度 retinex 算法对图像进行预处理。实验结果表明,通过以上方法得到的检测结果对不同场景环境均可保证较高的检测准确率。但是本文中提出的算法仍然存在一些问题,如针对严重遮挡的目标的检测存在定位不精确,以及对经过去雾处理后的图片仍存在一些目标没有识别的情况,这将是后期工作需要改进的重点方向。

参考文献:

- [1] 刘家旭. 基于 HOG 特征提取的车辆检测方法研究[D]. 北京: 华北电力大学, 2017.
Liu J X. Research on vehicle detection method based on HOG feature extraction[D]. Beijing: North China Electric Power University, 2017. (in Chinese)
- [2] 周行, 陈淑荣. 一种基于多特征融合级联分类器的车辆检测算法[J]. 现代计算机(专业版), 2018(17): 38-43.
Zhou X, Chen S R. A vehicle detection algorithm based on multi-feature fusion and cascade classifier[J]. Modern Computer, 2018(17): 38-43. (in Chinese)
- [3] 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), June 27-30, 2016, Las Vegas, NV, USA. IEEE,

- 2016: 779-788.
- [4] Liu W, Anguelov D, Erhan D, et al. SSD: single shot multibox detector[M]//Computer vision - ECCV 2016. Cham: Springer International Publishing, 2016: 21-37.
- [5] Zhang S F, Wen L Y, Bian X, et al. Single-shot refinement neural network for object detection[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 18-23, 2018, Salt Lake City, UT, USA. IEEE, 2018: 4203-4212.
- [6] Girshick R, Donahue J, Darrell T, et al. Region-based convolutional networks for accurate object detection and segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016, 38(1): 142-158.
- [7] Girshick R. Fast R-CNN[C]//2015 IEEE International Conference on Computer Vision (ICCV), December 7-13, 2015, Santiago, Chile. IEEE, 2015: 1440-1448.
- [8] 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.
- [9] Redmon J, Farhadi A. YOLOv3: an incremental improvement [EB/OL]. [2020-07-15]. <https://arxiv.org/abs/1804.02767>.
- [10] 冯加明, 储茂祥, 杨永辉, 等. 改进 YOLOv3 算法的车辆信息检测[J/OL]. 重庆大学学报: 1-12 [2020-07-10]. <http://kns.cnki.net/kcms/detail/50.1044.N.20200619.1503.004.html>.
Feng J M, Chu M X, Yang Y H, et al. Vehicle information detection based on improved YOLOv3 algorithm [J]. Journal of Chongqing University: 1-12 [2020-07-10]. <http://kns.cnki.net/kcms/detail/50.1044.N.20200619.1503.004.html>. (in Chinese)
- [11] 杜金航, 何宁. 基于改进的 YOLOv3 道路车辆实时检测[J]. 计算机工程与应用, 2020, 56(11): 26-32.
Du J H, He N. Real-time road vehicles detection based on improved YOLOv3 [J]. Computer Engineering and Applications, 2020, 56(11): 26-32. (in Chinese)
- [12] 崔文靓, 王玉静, 康守强, 等. 基于改进 YOLOv3 算法的公路车道线检测方法[J/OL]. 自动化学报: 1-9 [2020-07-29]. <https://doi.org/10.16383/j.aas.c190178>.
Cui W J, Wang Y J, Kang S Q, et al. Road lane line detection method based on improved YOLOv3 algorithm [J/OL]. Acta Automatica Sinica: 1-9 [2020-07-29]. <https://doi.org/10.16383/j.aas.c190178>. (in Chinese)
- [13] 胡贵桂. 改进的 YOLOv3 算法在道路环境目标检测中的应用[J]. 汽车实用技术, 2020 (5): 117-121.
Hu G G. The application of improved YOLOv3 algorithm in road environment target detection [J]. Automobile Applied Technology, 2020 (5): 117-121. (in Chinese)
- [14] 吴仁彪, 冯晓赛, 屈景怡, 等. 雾霾环境下基于 PLATE-YOLO 的车牌检测方法[J]. 信号处理, 2020, 36(5): 666-676.
Wu R B, Feng X S, Qu J Y, et al. License plate detection method based on PLATE-YOLO in smog environment [J]. Journal of Signal Processing, 2020, 36(5): 666-676. (in Chinese)
- [15] 汪昱东, 郭继昌, 王天保. 一种改进的雾天图像行人和车辆检测算法[J]. 西安电子科技大学学报, 2020, 47(4): 70-77.
Wang Y D, Guo J C, Wang T B. Algorithm for foggy-image pedestrian and vehicle detection [J]. Journal of Xidian University, 2020, 47(4): 70-77. (in Chinese)
- [16] 段世杰, 黄华, 王鹏飞, 等. 一种单帧图像的快速去雾方法[J]. 软件, 2015, 36(5): 1-6, 15.
Duan S J, Huang H, Wang P F, et al. A fast defogging method for single image [J]. Computer Engineering & Software, 2015, 36(5): 1-6, 15. (in Chinese)
- [17] He K M, Zhang X Y, Ren S Q, et al. Deep residual learning for image recognition [C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 27-30, 2016, Las Vegas, NV, USA. IEEE, 2016: 770-778.
- [18] 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.