

DOI:10.19651/j.cnki.emt.2106547

自然场景下的交通标志检测与识别*

陈红 王相超 陈志琳

(西安工业大学 电子信息工程学院 西安 710000)

摘要: 针对我国自然场景下的交通标志检测误差大、检测速度慢等一系列问题,提出了一种对YOLOv4算法的改进的方法。首先在算法的输入端加入图像增强、图像降噪等处理,然后对算法的检测层进行修改,删除 19×19 检测层,增加 152×152 检测层。最后利用K-means II聚类算法对重建的交通标志数据集进行聚类分析,重新定义网络的初始候选框。实验结果表明,改进后的算法能够在自然场景下准确及时的检测到小交通标志。在以CCTSDb数据集为基础建立的交通标志数据集上取得了96 mAP,检测速度为26 FPS,比YOLOv4算法分别提高了1.7%和1.4。

关键词: 交通标志;YOLOv4;深度学习;图像增强

中图分类号: TP391 **文献标识码:** A **国家标准学科分类代码:** 510.5040

Traffic sign detection and recognition in natural scene

Chen Hong Wang Xiangchao Chen Zhilin

(School of Electronic and Information Engineering, Xi'an Technological University, Xi'an 710000, China)

Abstract: In view of a series of problems such as large error and slow detection speed of traffic signs under natural scenes in China, an improved YOLOv4 algorithm is proposed. Firstly, image enhancement, image denoising and other processing are added to the input end of the algorithm. Then, the detection layer of the algorithm is modified by deleting the 19×19 detection layer and adding the 152×152 detection layer. Finally, the K-means II clustering algorithm is used to carry out clustering analysis on the reconstructed traffic sign data set, and the initial candidate box of the network is redefined. The experimental results show that the improved algorithm can detect small traffic signs accurately and in real time in the natural scene. 96 mAP is obtained on the traffic sign data set based on the CCTSDb data set, and the detection speed is 26 FPS, which is 1.7% and 1.4 higher than YOLOv4 algorithm respectively.

Keywords: traffic signs;YOLOv4;deep learning;image enhancement

0 引言

随着社会经济水平和人们生活水平的提高,汽车的数量逐年攀升。随之带来的交通拥挤的现象不断加剧。而自动驾驶和智能交通系统对交通拥挤的现状带来了很大地缓解。对交通标志的检测与识别是对自动驾驶技术应用的前提。

传统的交通标志检测,主要利用交通标志的特点,但是检测的效果和速度都不明显。文献[1]提出将待测目标信息先进行区域的划分,然后再进行特征的提取。文献[2]提出先利用对圆形图像的处理,然后利用光学三原色对特征进行提取。文献[3]提出在车载视频检测的场景下,通过对交通标志的形状特性进行分析,从而进行检测。以上算法虽然对交通标志进行了精确的定位,但是易受到天气、光照等因素的影响,且算法实时性达不到卷积神经网络(convolutional neural networks,CNN)的效果。

随着人工智能技术在各行各业中应用日益广泛,深度学习算法在其中得到了很好地发展。目前比较流行的深度学习算法是一阶段和二阶段算法。二阶段算法是首先通过候选区域生成特征向量,然后用分类器对待测目标进行预测、识别。主要以R-CNN^[4]、Faster R-CNN^[5-6]、Mask R-CNN^[7]为主。2018年,Li等^[8]提出了一种新的交通标志分类器并对其进行实验测试。Yang等^[9]提出了新的深度神经网络模型,通过两阶段调整策略提取区域建议。由于二阶段算法检测速度很难达到工程应用的要求以及模型训练时间长,因此一阶段算法更受学者们的喜爱。一阶段算法是将整个图像作为算法的输入,将检测问题视为回归问题,直接对检测的信息进行定位和识别,由于取消了候选区域的提取,导致算法的精度不高,但是极大地提高了目标检测的速度达到了实际应用要求。主要以YOLO^[10-13]、SSD^[14]和

收稿日期:2021-04-28

* 基金项目:陕西省科技厅项目(2018GY-153)、陕西省西安市未央区科技局项目(201833)资助

RetinaNet^[15]为代表。2016 年,Zhu 等^[16] 在自建的数据集 TT100K(Tsinghua-Tencent 100K)进行实验,其召回率和精确度均达到了 80%以上,达到了实际应用的要求。2018 年,Wang 等^[17] 提出了一种级联掩码生成框架,使用多尺度图像为输入,按尺度升序进行处理提取感兴趣区域,来解决小目标低分辨率检测的问题。2019 年,Song 等^[18] 为了解决在移动边缘平台上部署神经网络具有较长的延迟和巨大的计算成本问题,提出了一种高效的卷积神经网络(CNN),它可以显著地减少冗余,减少参数,提高网络速度。其模型大小相同的情况下,准确率比清华-腾讯 TT100K 提高了 10%。2021 年,Liu 等^[19] 提出了一种新的交通标志检测网络(TSingNet),该网络通过学习尺度感知

和上下文丰富的特征,有效地检测和识别野外小交通标志和阻塞交通标志。

本研究在 YOLOv4 算法的基础上,通过对网络结构和候选框的改进,提高算法在自然场景下对小目标交通标志识别的准确率和实时性。

1 YOLOv4 算法

YOLOv4 算法是 Alexey 在 YOLOv3 算法的基础上进行改进而来的,不仅提高了算法的精确度同时提升了算法的实时性。如图 1 所示,其主要由 4 部分组成,主干网络(CSPDarknet53)、金字塔模型(SPP)、特征提取网络(PANet)、预测模型(Prediction)。

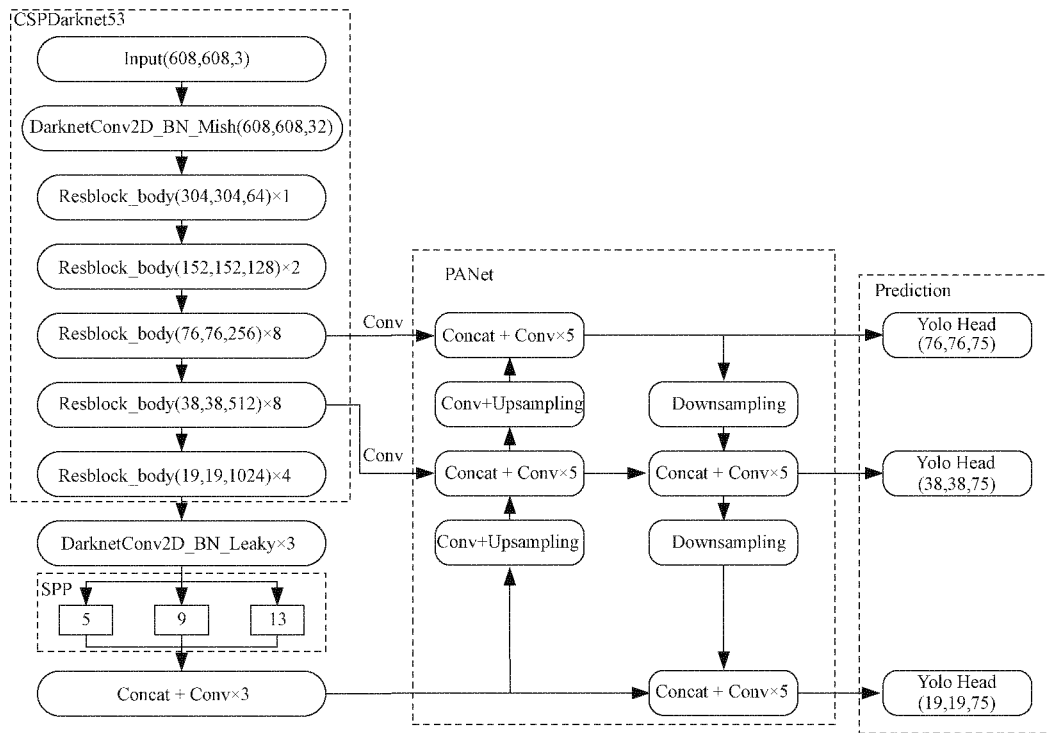


图 1 YOLOv4 算法网络模型

YOLOv4 主要从如下 3 个方面进行了算法的改进。

1) 特征提取网络

特征提取网络由 CSPDarknet53 替换 Darknet53 网络,不仅增强 CNN 的学习能力,而且降低了内存的消耗率,同时保证了准确性。

2) 激活函数

激活函数由 leaky_relu 替换为 Mish 函数,替换后的新的激活函数,在模型训练过程中不仅表现出来模型收敛速度较快,而且也节省了硬件成本。计算公式如式(1)、(2)所示。

$$Mish = x \times \tanh(\ln(1 + e^x)) \tag{1}$$

$$y_i = \begin{cases} x_i, & x_i \geq 0 \\ \frac{x_i}{a_i}, & x_i < 0 \end{cases} \tag{2}$$

3) MOSAIC 数据增强

为了增强训练样本集的丰富性,YOLOv4 算法每次随机在训练样本里面选取 4 张图片进行缩放、翻转、拼接等操作。这样不仅丰富了数据集的类型,而且在数据增强的过程中会出现很多小目标。

2 算法改进

总体改进思路:首先,为了对目标进行准确的定位,对数据集重新进行 K-means II 聚类分析得到候选框;其次,在 YOLOv4 算法的网络主干中删掉了大目标检测特征层,在原先小目标检测层的基础上,又增加了细小目标检测层;最后,在对模型进行测试之前增加数据增强、数据降噪等处理,将更清晰的图像输出给检测模型。改进后算法的训练和测试流程如图 2 所示。

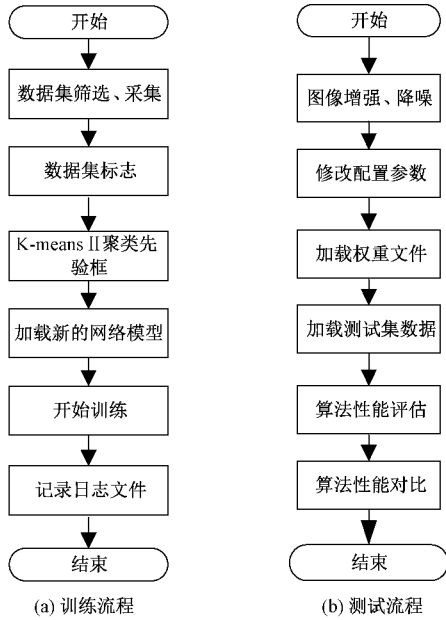


图 2 改进后的算法流程

2.1 先验框的设计

本文所实验的数据集主要以 CCTSDB 团队的数据集为基础,然后在不同天气下采集一部分数据集,因此需要重新对算法进行数据集进行聚类分析。YOLOv4 算法使用的是 K-means 算法对数据集进行聚类分析,因为其聚类算法需要随机确定初始的聚类中心,所以不同的聚类中心导致的结果不同。高健宸等^[20]使用 K-means++^[21]对数据集进行聚类,虽然较 K-means 算法略有提升,但是它明显的一个缺点是下一个聚类中心点的计算取决于上一个中心点。

为了提高模型的性能,本文采用 K-meansII 算法对实验的数据集重新计算,得到准确的先验框。与 K-means++ 算法不同的是 K-means II 使用了一个过程采样因子 $l = \Omega(k)$ 。

首先在训练的数据集中随机选择一个初始中心 c_1 ,然后计算距离初始中心最近的点的距离 d ,然后再对 $\log(d)$ 进行迭代,在每次迭代中,给定当前的中心集 C ,让它以概率 p_x (式 3)对每个样本进行采样。将采样得到的点加入到集合 C 中,然后更新集合 C ,再次对算法进行迭代。K-means II 算法的计算规则大致如表 1 所示。

$$p_x = \frac{l \cdot d^z(c_1, c_x)}{\phi_x(c)} \tag{3}$$

表 1 K-means II 算法的迭代原理

次数	流程
1	从样本中选取一个初始点 c_1
2	计算距离 c_1 最近点的距离 d
3	对计算出的距离 d 取对数后迭代
4	迭代的概率 p_x
5	将第 4 步迭代的点加入到中心集 c 中
6	更新中心点集合,重复第 2 步

2.2 多尺度预测框设计

YOLOv4 算法的目标检测框中使用了 3 个尺度对物体的信息进行提取,分别为 $19 \times 19, 38 \times 38, 76 \times 76$,这 3 个尺度分别负责大、中、小目标的检测,在检测之前会根据配置文件里面先验框的大小对边界框进行调整。这种方法相较以前的算法来说增加了检测的精度和速度。但是对于交通标志目标小,且在一张图像中所含的背景较多,导致需要检测信息所占的像素较小且交通标志的种类较多,在实际测试中,存在着漏检和错检的现象。

因此针对这一问题,通过分析待测物体的属性,本文提出在原算法模型的基础上增加了一个特征提取的尺度,同时为了保证算法的实时性,算法中大目标特征提取的尺度层去掉。改进后的算法 improve-YOLOv4 的网络结构如图 3 所示。

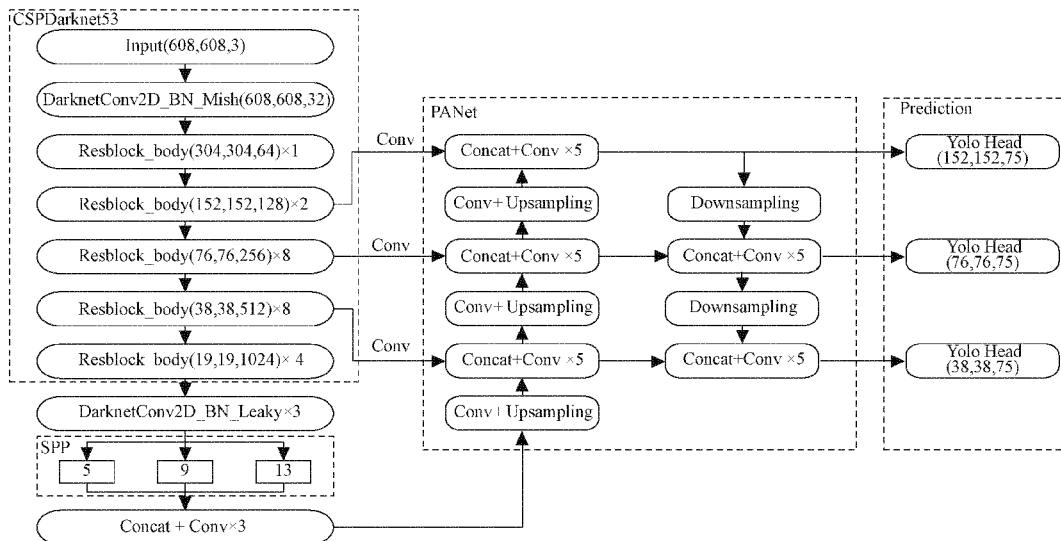


图 3 改进后的 YOLOv4 网络模型

1)首先将特征提取网络中的 76×76 层进行一个上采样和卷积的过程,使其尺度变成 152×152 ,然后和主干网络中的 152×152 输出层进行一个拼接过程,最后让其进行5次卷积输出到网络的预测层进行处理同时进行下采样。

2)首先将SPP网络处理后的输出层进行3次卷积后,直接进行上采样的过程与主干网络中 38×38 输出层进行一个拼接的过程,然后将PANet网络中的 76×76 进行下采样之后,直接和主干网络中的 38×38 进行拼接并进行5次卷积操作,最后直接输出到预测网络的 38×38 进行预测。

2.3 数据增强、降噪

本文针对的自然场景下的交通标志检测,针对自然场景下需要检测的信息并不是在理想状态下输入到算法模型中。因此在对模型进行实际使用之前,将要检测的图像进行增强、降噪处理,然后将处理后的图像输入到算法模型中。

1)考虑到实际场景下,图像的输入伴随着噪声的存在,因此在数据集中选取一张warning标签的图像,对其进行均值滤波处理,如图4所示。



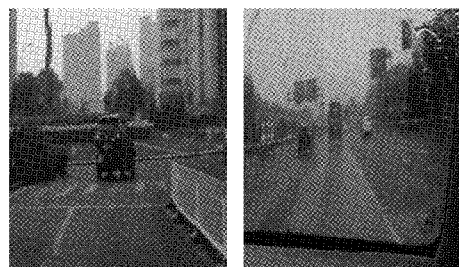
图4 降噪

2)考虑到夜晚场景下,模型的输入端的图像光线较暗、画面模糊,在测试的样本中随机选取一张图片,对其进行色彩度处理,处理后的结果如图5所示。

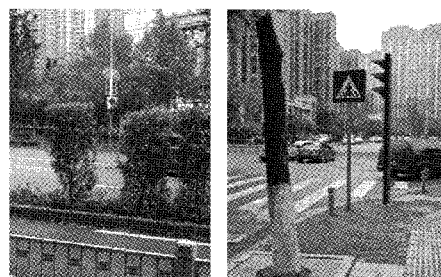


图5 增强

晚场景下在城市环境中采集了2000张数据集,然后在CCTSDDB数据集中筛选出11850张清晰度较高的数据集,最后组成新的数据集,该数据集共含有图片13850张,包含了3个类别:指示标志、警告标志、禁止标志。用python脚本对实验的数据集随机按照8:2的比例进行分类,实验中部分采样的数据如图6所示。



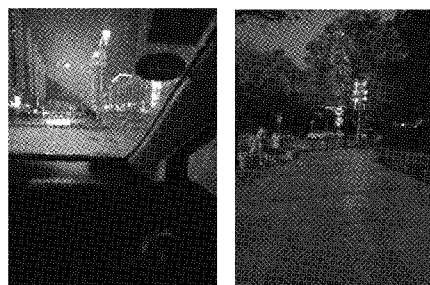
(a) 雨天



(b) 阴天



(c) 晴天



(d) 夜晚

图6 部分采样图

3 实验结果与分析

3.1 实验数据集及实验环境

1) 实验数据集

针对本文所研究的问题,分别在雨天、阴天、晴天、夜

2) 数据集标注

根据本文算法用到的模型要求需要将数据集标注成VOC格式,使用官方的标注工具Labeling图片进行标注,将35种黑色边黄色底的三角形警告标志标注为warning标签。将29种蓝色底的方形、圆形指示标志标注为

mandatory 标签。将 42 种白底红边的圆形禁令标志标注为 prohibitory 标签。具体标注如图 7 所示。



图 7 数据标注

3) 实验环境

本文实验中所用到的算法以 darknet 深度学习框架为基础进行环境搭建 python3.7, Visual Studio 2019, Opencv3.4.0, cuda11.0.3, cudnn8.0.3。硬件设备如表 2 所示。

表 2 实验硬件环境

参数	PC
CPU	Intel(R) Core(TM) i7-4790 HQ CPU @ 2.60 GHz(2 601 MHz)
GPU	NVIDIA GeForce GTX 1080 Ti
内存	16.00 GB

4) 实验参数设置

本文为保证训练结果的准确性,所对比的算法均在训练参数一样的情况下,对模型进行训练,具体的参数值设置如表 3 所示。

表 3 实验参数

参数	大小
迭代次数	20 000
学习率	0.001
衰减系数	0.005
动量	0.949
批量	16

3.2 性能评价指标

本实验模型评价的指标以 mAP、P-R、FPS 3 个参数为主,损失函数为辅。其中, mAP 是本实验中 3 个类 (average precision, AP) 的平均值如式(4)、(5)所示, P-R 是召回率(recall, R)和精确度 (precision, P) 所画的曲线图与

坐标轴的面积,计算公式如式(6)、(7)所示, FPS 是处理一张图片需要的时间。损失函数的计算如式(8)所示。

$$AP = \int_0^1 P(R) dR \tag{4}$$

$$mAP = \frac{1}{n} [AP_1 + AP_2 + \dots + AP_n] \tag{5}$$

$$P = \frac{T_p}{F_p + T_p} \tag{6}$$

$$R = \frac{T_p}{F_p + F_n} \tag{7}$$

$$L_{CoU} = 1 - IoU + \frac{\rho^2(b, b^{gt})}{c^2} + \frac{4}{\pi^2} \left(\arctan \frac{w^{gt}}{h^{gt}} - \arctan \frac{w}{h} \right)^2 \cdot \frac{4}{\pi^2} \left(\arctan \frac{w^{gt}}{h^{gt}} - \arctan \frac{w}{h} \right)^2 \tag{8}$$

$$(1 - IoU) + \frac{4}{\pi^2} \left(\arctan \frac{w^{gt}}{h^{gt}} - \arctan \frac{w}{h} \right)^2$$

其中, T_p 表示训练集中标注图片被检测为正样本; F_p 表示数据集中的负样本被检测为正样本; F_n 表示训练集中标注图片被检测为负样本; c 代表对角线的距离; ρ 表示的是欧氏距离; b 表示预测框与中心点的距离; b^{gt} 表示真实框与中心点的距离。

3.3 结果与分析

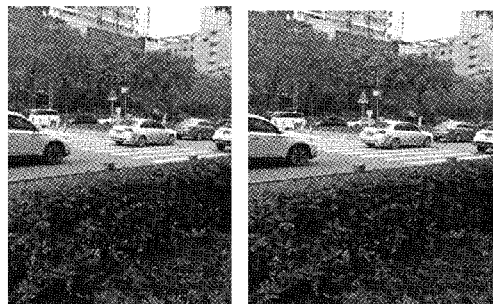
1) K-means II 算法实验结果

由于 YOLOv4 的先验框是根据 coco 数据集,用 K-means 算法聚类得到的,对其他数据集不具有适用性。本文用 K-means II 聚类算法,对新数据集聚类得到的先验框如表 4 所示。

表 4 K-means II 先验框

特征层	特征提取	先验框
中目标层	38×38	(34,72), (56,93), (107,46)
小目标层	76×76	(16,45), (28,42), (22,60)
细小目标层	152×152	(11,27), (13,36), (18,29)

在城市道路环境下用 K-means II 算法聚类后相对于 K-means 聚类算法实际测试如图 8 所示



(a) K-means 算法 (b) K-means II 算法

图 8 聚类算法对比

由图 8 可以看出,图 8(b)算法对目标检测框的定位和 warning 交通标志的位置完全重合,而图 8(a)算法对目标

检测框的定位不仅偏小而且位置也发生了偏移。实验表明 K-means II 算法对数据集进行重新聚类后,对目标检测的定位更加准确。

2)YOLOv4-improve 算法精确度实验

按照实验参数设置,分别对 YOLOv4、YOLOv4-tiny 以及改进后的算法(YOLOv4-improve),在测试集上对 3 种交通标志的精确度 AP 进行统计,统计结果如图 9 所示。

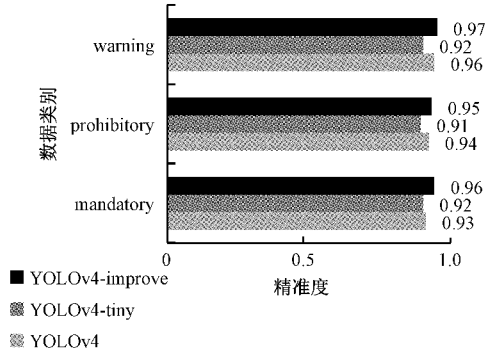


图 9 算法精度对比

由图 9 可知,在相同的测试集的情况下,改进后的算法 YOLOv4-improve 精确度(AP)比 YOLOv4 分别提高了 1%、1%、3%。比 YOLOv4-tiny 算法的精度(AP)分别提高了 5%、4%、4%。由此不难看出,改进后的算法有效地提高了在自然场景下对交通标志检测的精度。

3)模型损失函数实验结果

实验参数中设置的最大迭代次数为 20 000 次,在训练过程中查看日志文件得知模型迭代到 19 000 次左右后,损失函数基本不再下降,表明模型不再学习,因此终止训练,算法迭代 150 次之前的损失函数值较大,实验结果不明显,因此将日志文件中的迭代次数 150~19 000 的数据进行统计,实验结果计如图 10 所示。

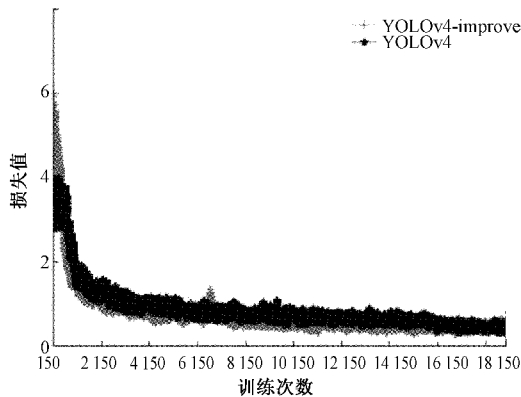


图 10 损失函数对比

由图 10 实验的数据可以得知,改进后的 improve-YOLOv4 算法从总体趋势来看,较 YOLOv4 算法下降的较快,表明了改进后算法的有效性。

4)模型 mAP 等参数的实验数据

由 YOLOv4 improve 算法精确度实验可以计算出算法模型的 mAP,通过对训练日志文件的处理,现将算法的 mAP、Recall、FPS 模型评价指标进行对比,实验数据如表 5 所示。

表 5 模型评价指标对比

算法名称	平均精度 (mAP/%)	召回率 (Recall/%)	检测速度 (FPS/s)
YOLOv4	94.3	97.2	24.6
YOLOv4-tiny	91.6	96.5	24.3
YOLOv4-improve	96.0	99.4	26.0

由表 5 可以看出,由于 YOLOv4-tiny 网络结构比较小,所以模型检测精确度略有下降,改进后的 YOLOv4-improve 算法在平均精度、召回率、检测速度上较 YOLOv4 算法分别提升了 1.7%、2.2%、1.4。比 YOLOv4-tiny 算法在平均精度、召回率、检测速度上分别提升了 4.4%、2.9%、1.7。

5)P-R 曲线实验

对 3 种算法的日志进行分析,用官方的脚本文件对数据进行统计,由于绘制 3 个曲线与横纵坐标的阴影部分,会互相重叠,不能展示实验效果,因此只绘制了 3 种算法的曲线,便于分析比较。实验结果如图 11 所示。

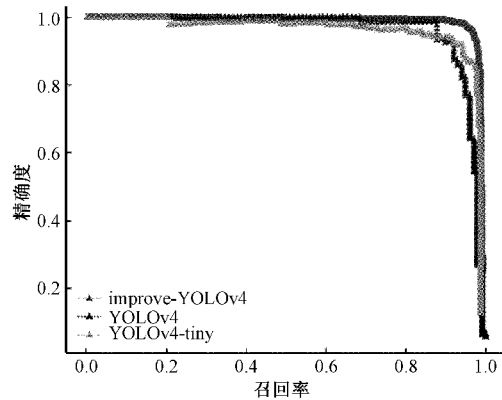


图 11 P-R 曲线对比

从图 11 中可以看出,改进后的算法 improve-YOLOv4 与横纵坐标轴的面积是最大的,表示模型的性能最好。而 YOLOv4 和 YOLOv4-tiny 曲线与横纵坐标轴的面积基本相差不大。

6)自然环境下算法检测实验

为了更好地验证改进好的算法模型与原有算法模型的效果,分别在阴天、雨天、夜晚场景下对 YOLOv4、YOLOv4-tiny、improve-YOLOv4 3 种算法进行实验,实验结果如图 12 所示。

在图 12 中可以看出,由于 YOLOv4-tiny 算法模型较

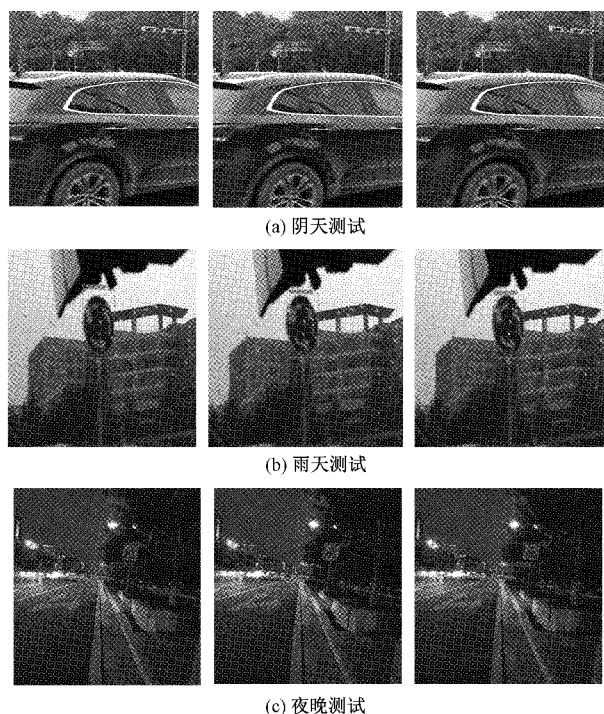


图 12 3 种算法测试结果

小,所以在实际测试中精度也比其他两种算法精度低。而改进后的 improve-YOLOv4 算法在实际测试中表现良好,能够在自然场景下对目标信息进行很好检测与识别,达到了实验的要求。

4 结 论

本文提出了一种改进后的 improve-YOLOv4 算法,针对在自然场景下交通标志尺度小、环境影响大导致的检测精度低、实时性差等问题。该算法提出在 YOLOv4 模型的基础上用 K-means II 聚类算法对数据集进行重新聚类生成新的先验框,此外还在网络的特征层加入了一个小尺度特征融合层来进一步提高模型对交通标志中小目标的检测精确度和检测速度。本文以长沙理工大学的数据集为基础,建立了新的数据集,并对其进行测试。实验结果表明,在自然场景下相对于 YOLOv4、YOLOv4-tiny 算法,改进后的 improve-YOLOv4 算法在准确率上分别提升了 1.7% 和 4.4%,在检测速度上分别提升了 1.4 和 1.7,其准确率提升的同时检测速度并没有下降,在自然场景下算法的性能表现良好。但是模型的网络较大,在嵌入式移动设备上较难部署,仍是一个待解决的问题。

参考文献

- [1] LI H, SUN F, LIU L, et al. A novel traffic sign detection method via color segmentation and robust shape matching [J]. *Neurocomputing*, 2015, 169: 77-88.
- [2] BERKAYA S K, GUNDUZ H, QZSEN O, et al. On circular traffic sign detection and recognition [J]. *Expert Systems with Applications*, 2016, 48: 67-75.
- [3] 杜影丽,贾永红,韩静敏.自然场景车载视频道路交通限速标志的检测与识别方法[J].*测绘地理信息*,2018,43(2):32-34,37.
- [4] GIRSHICK R, DONAHUE J, DARRELL T. Rich feature hierarchies for accurate object detection and semantic segmentation [C]. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2014:580-587.
- [5] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks[J]. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2017, 39 (6): 1137-1149.
- [6] 伊欣同,单亚峰.基于改进 Faster R-CNN 的光伏电池内部缺陷检测[J].*电子测量与仪器学报*,2021,35(1):40-47.
- [7] KAIMING H, GKIOXARI G, DOLLAR P, et al. Mask R-CNN[C]. *IEEE Conference on Computer Vision and Pattern Recognition(CVPR)*,2018.
- [8] LI J, WANG Z. Real-time traffic sign recognition based on efficient CNNs in the wild [J]. *IEEE Transactions on Intelligent Transportation Systems*, 2019, 20(3):975-984.
- [9] YANG T, LONG X, SANGAIAH A K, et al. Deep detection network for real-life traffic sign in vehicular networks[J]. *Computer Networks*, 2018, 136 (MAY8): 95-104.
- [10] REDMON J, DIVVALA S K, GIRSHICK R, et al. You only look once: Unified, real-time object detection[C]. *IEEE Conference on Computer Vision and Pattern Recognition(CVPR)*,2016:779-788.
- [11] REDMON J, FARHADI A. YOLO9000: Better, faster, stronger [C]. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017: 6517-6525.
- [12] REDMON J, FARHADI A. YOLOv3: An incremental improvement[C]. *2018 IEEE Conference on Computer Vision and Pattern Recognition(CVPR)*,IEEE,2018.
- [13] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. YOLOv4: Optimal speed and accuracy of object detection[J]. *ArXiv Preprint*, 2020, ArXiv:2004.10934.
- [14] LIU W, ANGUELOV D, ERHAN D. SSD: Single shot multibox detecto [C]. *IEEE Conference on Computer Vision and Pattern Recognition(CVPR)*,2016.
- [15] PEI D, JING M, LIU H, et al. A fast RetinaNetFusion framework for multi-spectral pedestrian detection [J].

- Infrared Physics & Technology, 2020, 105:103178.
- [16] ZHU Z, LIANG D, ZHANG SH, et al. Traffic-sign detection and classification in the wild[C]. 2016 IEEE Conference on Computer Vision and Pattern Recognition(CVPR), IEEE, 2016.
- [17] WANG G, XIONG Z, LIU D, et al. Cascade mask generation framework for fast small object detection[C]. 2018 IEEE International Conference on Multimedia and Expo(ICME), San Diego, CA, USA, 2018: 1-6.
- [18] SONG SH J, QUE ZH Q, HOU J J, et al. An efficient convolutional neural network for small traffic sign detection[J]. Journal of Systems Architecture, 2019, DOI: 10.1016/j.sysarc.2019.01.012.
- [19] LIU Y Y, PENG J Y, XUE J H, et al. TSingNct: Scale-aware and context-rich feature learning for traffic sign detection and recognition in the wild [J]. Neurocomputing, 2021, DOI: 10.1016/J. NEUCOM. 2021.03.049.
- [20] 高健宸,张家洪,李英娜,等. 基于 YOLOv4 的绝缘子爆裂故障识别研究[J/OL]. 激光与光电子学进展, 2021;1-13[2021-04-23].
- [21] 邓聪颖,叶波,苗建国,等. 基于 K-means++ 聚类与概率神经网络的数控机床变位姿态动态特性模糊评估[J]. 仪器仪表学报, 2020, 41(12):227-235.

作者简介

陈红,副教授,硕导,主要研究方向为信号与信息处理、图像处理。

E-mail:120268288@qq.com

王相超,硕士生,主要研究方向为图像处理。

E-mail:1820841846@qq.com

陈志琳,硕士生,主要研究方向为图像处理。

E-mail:707226539@qq.com