

基于函数改进的YOLOv3车辆检测与识别算法

宋华杰^{1,2}, 周磊¹

(1. 西安铁路职业技术学院, 陕西 西安 710026;

2. 西安工程大学, 陕西 西安 710048)

摘要: YOLOv3算法在目标检测和识别方面具有检测精度高、速度快等优点,但其损失函数中高宽比部分的损失运算对大目标与小目标的区分不明显,导致损失计算结果精度较差,小目标预测框尺寸不准确。为此,对YOLOv3算法的损失函数进行了改进,将宽高坐标误差部分修改为与目标宽高真实值相关的形式,并将原模型的非极大值抑制函数中有重叠区域的检测框的分数由固定值改为衰减函数的形式,最后对改进模型进行了车辆检测实验。研究表明,改进模型在不影响车辆检测速度的同时,召回率 R 、精确率 P 、调和 $F1$ 值和平均准确度 mAP 分别提高了0.1、0.105、0.103、0.021,改进模型的整体性能优于原YOLOv3模型。

关键词: YOLOv3; 损失函数; 非极大值抑制; 车辆检测; 车型识别

中图分类号: TP14

文献标志码: A

doi: 10.11959/j.issn.2096-6652.202301

Vehicle detection and recognition algorithm based on function improvement of YOLOv3

SONG Huajie^{1,2}, ZHOU Lei¹

1. Xi'an Railway Vocational and Technical Institute, Xi'an 710026, China

2. Xi'an Polytechnic University, Xi'an 710048, China

Abstract: YOLOv3 algorithm is characterized by high detection accuracy and high speed in the aspect of target detection and recognition. It is outstanding among similar algorithms, but there are still obvious problems. The loss operation of the wide and high part of the loss function of this algorithm is not obvious to the distinction between the big target and the small target, which leads to the problem that the loss calculation is not accurate enough and the prediction box size of the small target is not accurate. In view of this defect, the loss function of YOLOv3 algorithm was improved, the wide-height coordinate error was modified into the form of proportion, the non-maximum suppression method of the original model was improved and the overlap threshold was changed from the fixed value to the form of attenuation function. Finally, the model was applied to vehicle detection. The experimental results showed that the improved model improved the accuracy rate, recall rate and F1 value to different degrees while basically not affecting the vehicle detection speed, making the overall performance of the model better than the original YOLOv3 model.

Key words: YOLOv3, loss function, non-maximum suppression, vehicle detection, vehicle recognition

0 引言

基于图像的车辆识别与检测是智能交通系统中

解决交通拥堵、乱停放、违法车辆追踪问题的重要方法,也是智能交通系统的研究热点^[1]。在车辆图像检测方面,利用光流法^[2]、帧差法^[3]、背景差分

收稿日期: 2021-11-01; 修回日期: 2021-12-23

通信作者: 宋华杰, 574277447@qq.com

基金项目: 西安铁路职业技术学院2023年度立项课题 (No.XTZY23K12)

Foundation Item: 2023 Project Initiation Project of Xi'an Railway Vocational and Technical College (No.XTZY23K12)

法^[4]、梯度直方图 (histogram of oriented gradient, HOG)^[5]、哈尔 (haar-like) 特征^[6]等方法均可实现对视频中运动车辆的检测,但这些检测方法在速度和精度方面都难以满足实时检测的要求。而深度学习技术可获得车辆更多的细节特征,运算更快速、更精确、更智能,可大幅提高车辆识别准确性。

目前深度学习检测方法包括基于区域和回归的两种典型算法。基于区域的算法包括 RCNN^[7]、Fast R-CNN^[8]、Faster R-CNN^[9]、SPP-NET^[10]等。RCNN 首次提出选择性搜索方法,即用一定分割方法将图像分为多个小区域,通过纹理、颜色、面积等合并方式产生候选区域。RCNN 虽有很大突破,经过多次改进后精度基本满足要求,但基于区域的算法存在检测步数多、网络结构复杂、检测速度偏慢等问题,无法满足实时检测的要求。基于回归的算法主要包括 YOLO^[11]和 SSD^[12]。这两类算法将目标检测问题看作回归问题,直接输出目标的类别概率和位置信息,其中 YOLO 算法速度快、精度高,在各领域应用较广泛^[13]。目前 YOLO 算法的改进算法包括 YOLOv2^[14]、YOLOv3^[15]两种。在算法结构改进方面,阮激扬^[16]改进了 YOLO 算法的网络结构,在卷积层输入端添加批规范化层的操作,保证了每层网络具有相同分布的输入。杨眷玉^[17]把 YOLO 模型的第3、4层结构替换为 inception 结构,提高了网络能力,减少了网络参数,提高了检测速度。但 YOLOv3 算法在函数设计方面仍有其固有缺点,其损失函数中高宽比部分的损失运算对大、小目标的区分不明显,使模型对小目标宽高部分的损失不敏感,预测框尺寸不准确。另外,算法的非极大值抑制函数阈值设计为固定值,在候选框置信度小于阈值时,该候选框会被直接删除,导致模型

可能漏检距离近的两个物体。

为此,本文拟建立以 YOLOv3 算法为基础的自然场景下车辆检测与车型识别模型,将损失函数的宽高损失修改为与目标宽高真实值相关的形式,并将非极大值抑制函数中有重叠区域的检测框的置信度由固定值改为衰减函数的形式,以期在保持较好检测速度的前提下,提高模型对小数目样本和遮挡目标的检测精度,提高小尺寸目标检测框定位的准确性,改进模型整体性能。

1 YOLOv3 算法模型

YOLOv3 算法模型采用了 52 个卷积层,取消了 YOLOv2 中的全连接层和池化层,属于全卷积模型结构,其结构如图1所示。模型使用大量卷积层提取图像特征,并多次使用残差的跳层连接,将不同尺度的特征进行拼接,而降采样是通过改变卷积核的步长来实现的,避免了利用池化层进行降采样引发的梯度负面效果。

YOLOv3 算法模型对目标的检测是针对整个图片的,摒弃了滑动窗口方法而采用回归的思想,直接输出类别和位置信息。将输入图像缩放至 416×416 网格后再划分为 $S \times S$ 个网格,若某目标物体的中心点位于某网格中,则该网格负责检测这个目标,每个网格可输出 3 个检测框,即检测 3 个目标物体。每个检测框含有 5 个基本参数 (x , y , w , h , confidence) 和 C 个类别的概率信息。其中 (x , y) 为目标中心点坐标, (w , h) 为边界框的宽高, confidence 为有无目标的置信度, C 为该网格 3 个预测目标的类别概率信息。因此每个网格最后输出为 $3 \times (5+C)$ 的张量。YOLOv3 输出 3 个尺度上的预测结果,即在 3 个不同大小的感受野下对目标进行

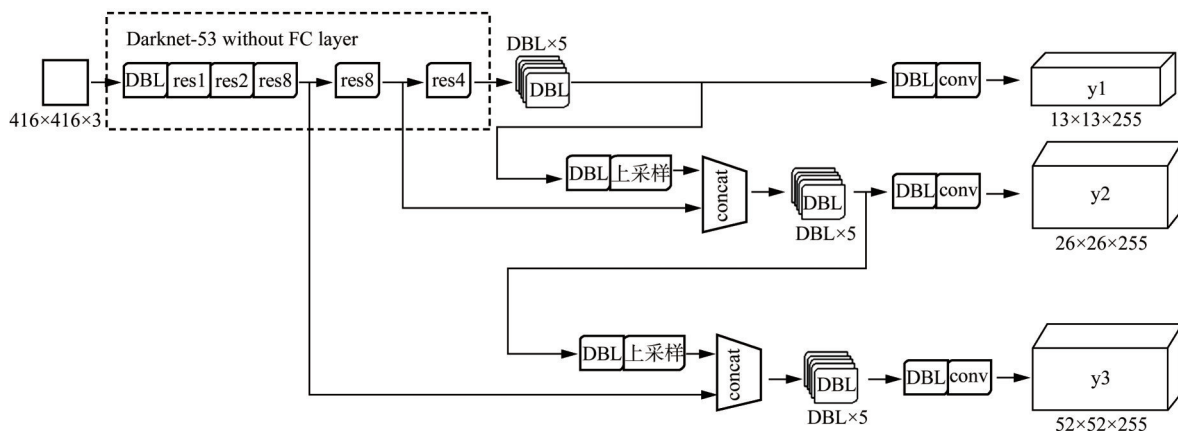


图1 YOLOv3 算法模型结构

预测，对不同大小的目标都有良好的检测效果。YOLOv3算法检测过程如图2所示。

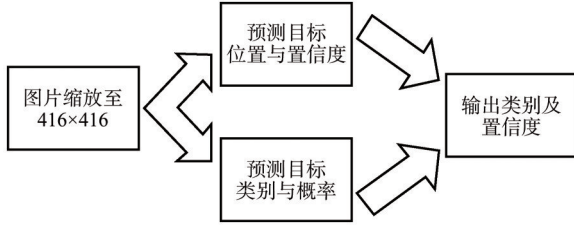


图2 YOLOv3算法检测过程

由图2可知，YOLOv3算法同时进行“预测目标位置与置信度”和“预测目标类别与概率”计算后输出类别和置信度，相比传统神经网络算法，提高了检测速度，是“one stage”（单阶段）算法的代表。

2 改进的YOLOv3模型

2.1 基于损失函数的改进

上述YOLOv3模型中，数据存在多次前向传播与反向传播，反向传播根据损失函数的计算结果对权重进行调整。YOLOv3模型的损失函数如式（1）所示：

$$\begin{aligned} \text{Loss} = & \lambda_{\text{coord}} \sum_{i=0}^{S^2} \sum_{j=0}^B I_{ij}^{\text{obj}} (2 - w_i \times h_i) [(x_i - \hat{x}_i^j)^2 + (y_i - \hat{y}_i^j)^2] + \\ & \lambda_{\text{coord}} \sum_{i=0}^{S^2} \sum_{j=0}^B I_{ij}^{\text{obj}} (2 - w_i \times h_i) [(w_i - \hat{w}_i^j)^2 + (h_i - \hat{h}_i^j)^2] - \\ & \sum_{i=0}^{S^2} \sum_{j=0}^B I_{ij}^{\text{obj}} [\hat{C}_i^j \log(C_i^j) + (1 - \hat{C}_i^j) \log(1 - C_i^j)] - \\ & \lambda_{\text{noobj}} \sum_{i=0}^{S^2} \sum_{j=0}^B I_{ij}^{\text{noobj}} [\hat{C}_i^j \log(C_i^j) + (1 - \hat{C}_i^j) \log(1 - C_i^j)] - \\ & \sum_{i=0}^{S^2} I_{ij}^{\text{obj}} \sum_{c \in \text{classes}} [\hat{P}_i^j \log(P_i^j) + (1 - \hat{P}_i^j) \log(1 - P_i^j)] \quad (1) \end{aligned}$$

损失函数由中心坐标误差、宽高坐标误差、置信度误差和分类误差4部分组成。 I_{ij}^{obj} 和 I_{ij}^{noobj} 表示第*i*个网格的第*j*个候选框是否负责该目标。若 $I_{ij}^{\text{obj}}=1$ ，则 $I_{ij}^{\text{noobj}}=0$ ；若 $I_{ij}^{\text{obj}}=0$ ，则 $I_{ij}^{\text{noobj}}=1$ 。 w_i^j 与 h_i^j 表示检测框的宽高， $\hat{w}_i^j$ 与 $\hat{h}_i^j$ 表示目标真实框的宽高。 $\hat{C}_i^j$ 表示置信度，若网格的第*i*个网格的第*j*个检测框负责检测某对象，则 $\hat{C}_i^j=1$ ，否则 $\hat{C}_i^j=0$ 。 P_i^j 与 $\hat{P}_i^j$ 分别表示预测物体的类别概率。 λ_{coord} 与 λ_{noobj} 为补偿常数。其中，宽高坐标误差部分

如式（2）：

$$B = \lambda_{\text{coord}} \sum_{i=0}^{S^2} \sum_{j=0}^B I_{ij}^{\text{obj}} (2 - w_i \times h_i) [(w_i - \hat{w}_i^j)^2 + (h_i - \hat{h}_i^j)^2] \quad (2)$$

其中，若第*i*个网格的第*j*个检测框负责某目标的检测，该检测框和目标真实框的宽高之差及模型中检测框与目标框的尺寸均经过归一化处理，故 w_i^j 、 h_i^j 、 $\hat{w}_i^j$ 与 $\hat{h}_i^j$ 的取值范围均为0~1。

相同误差对不同尺寸目标造成的影响不同，原模型中的损失函数为了体现这一不同，将预测框与目标真实框的宽高的误差乘以比例系数 $(2 - w_i \times h_i)$ ，但在实验过程中发现，仅乘该系数无法补偿这种不同。为体现同样尺寸的误差对大目标的影响较小，对小目标的影响较大，本节将损失函数中宽高坐标误差部分修改为比值的形式，即以误差占目标真实框的比例决定该误差的大小，而非仅依靠误差的数值。原模型中的比例系数 $(2 - w_i \times h_i)$ 是为减小该误差提出的，因此予以保留。修改后的宽高坐标误差部分如式（3）所示：

$$B = \lambda_{\text{coord}} \sum_{i=0}^{S^2} \sum_{j=0}^B I_{ij}^{\text{obj}} (2 - w_i \times h_i) \left[\left(\frac{w_i - \hat{w}_i^j}{\hat{w}_i^j} \right)^2 + \left(\frac{h_i - \hat{h}_i^j}{\hat{h}_i^j} \right)^2 \right] \quad (3)$$

损失函数如式（4）所示：

$$\begin{aligned} \text{Loss} = & \lambda_{\text{coord}} \sum_{i=0}^{S^2} \sum_{j=0}^B I_{ij}^{\text{obj}} (2 - w_i \times h_i) [(x_i - \hat{x}_i^j)^2 + (y_i - \hat{y}_i^j)^2] + \\ & \lambda_{\text{coord}} \sum_{i=0}^{S^2} \sum_{j=0}^B I_{ij}^{\text{obj}} (2 - w_i \times h_i) \left[\left(\frac{w_i - \hat{w}_i^j}{\hat{w}_i^j} \right)^2 + \left(\frac{h_i - \hat{h}_i^j}{\hat{h}_i^j} \right)^2 \right] - \\ & \sum_{i=0}^{S^2} \sum_{j=0}^B I_{ij}^{\text{obj}} [\hat{C}_i^j \log(C_i^j) + (1 - \hat{C}_i^j) \log(1 - C_i^j)] - \\ & \lambda_{\text{noobj}} \sum_{i=0}^{S^2} \sum_{j=0}^B I_{ij}^{\text{noobj}} [\hat{C}_i^j \log(C_i^j) + (1 - \hat{C}_i^j) \log(1 - C_i^j)] - \\ & \sum_{i=0}^{S^2} I_{ij}^{\text{obj}} \sum_{c \in \text{classes}} [\hat{P}_i^j \log(P_i^j) + (1 - \hat{P}_i^j) \log(1 - P_i^j)] \quad (4) \end{aligned}$$

2.2 基于NMS的改进

NMS即非极大值抑制（non-maximum suppression），该方法首先设置一个固定阈值，检测出的候选框也会计算出一个置信度，随后按如下顺序执

行。(1) 若候选框置信度小于阈值, 该候选框直接被删除, 不进入下一步运算, 这样可以减少算法运行过程中不必要的计算量。(2) 若候选框置信度大于阈值, 算法将所有的候选框按照置信度大小排序, 假设置信度最高的候选框为 X , 再计算其余每个候选框 A 和 X 的交并比 (intersection over union, IoU)。若IoU不小于提前设置好的重叠阈值 N_t (算法中给出), 则认为这两个框检测到的是同一个目标, 直接将候选框 A 删除, 即 $s_i=0$ 。若IoU小于 N_t , 则认为这两个框检测到的是不同对象, 保留候选框 A 。具体流程如图3所示。

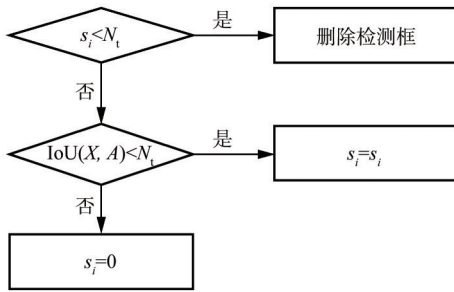


图3 NMS算法流程 (部分)

上述过程可防止同一个目标框被选多次的问题出现, 但在两辆车重叠区域较多的情况下, 容易漏掉一辆。原NMS方法可表示为式 (5):

$$s_i = \begin{cases} s_i, & \text{IoU}(X, A) < N_t \\ 0, & \text{IoU}(X, A) \geq N_t \end{cases} \quad (5)$$

原模型函数中重叠阈值 N_t 为固定值, 候选框 A 在经NMS函数处理后, 置信度 s_i 保持不变或变为0。本文为减少漏检情况, 将 $\text{IoU}(X, A) \geq N_t$ 的候选

框 A 的置信度设置为高斯加权函数 $f = e^{-\frac{x^2}{a}}$ 的形式, 而 $\text{IoU}(X, A) < N_t$ 时 s_i 保持不变。即两个检测框距离较近时, A 的置信度并非直接置为0, 而是要乘一个较小的权重, 权重大小视 A 与 X 的重叠情况而定。虽然检测分数会降低, 但该检测框仍在物体检测的候选集合序列中, 并将调整后的置信度与阈值比较, 若置信度小于设定的阈值才被删除。改进后的NMS算法可以表示为式 (6):

$$s_i = \begin{cases} s_i, & \text{IoU}(X, A) < N_t \\ s_i \times e^{-\frac{(\text{IoU}(X, A))^2}{0.3}}, & \text{IoU}(X, A) \geq N_t \end{cases} \quad (6)$$

3 实验结果与分析

3.1 实验数据集

本文采用公开数据集UA-DETRAC进行实验。该数据集拍摄于天津和北京道路, 涵盖晴天、阴天、雨天、夜晚等多个场景, 并标注了不同角度的行驶车辆, 但未标注路边停放车辆 (按背景处理)。标注车型12种, 但因其中7种车型样本少, 无法达到预期训练效果, 本文仅使用“Sedan”“SUV”“Taxi”“Bus”“Van”5种车型样本进行实验。数据集包括82 080张图片, 考虑到视频录制速度为25帧/秒, 相邻图片之间的差别非常小, 为节省训练时间, 随机选择10%的图片作为本文的数据集。在该数据集中选取50%作为训练和验证集, 其余50%作为测试集。数据集部分图片如图4所示。

3.2 实验平台

本文实验平台为win10环境下的Pytorch 4.0,



图4 数据集部分图片

基于此建立Pytorch框架下的YOLOv3模型，同时使用NVIDIA推出的运算平台CUDA支持GPU解决复杂运算问题，采用深度学习网络的GPU加速库CUDNN加速运算。计算机显卡为NVIDIA GTX 1650，满足运算要求。

3.3 评估指标与结果对比

选取目标检测算法中最常用的评价指标——召回率、精确率、F1值和mAP值，对实验结果进行评价。

3.3.1 召回率 R

召回率 R 也称查全率，是衡量目标检测模型性能的指标之一，表示检测出的正样本在所有正样本中所占的比例。召回率一般用式(7)表示：

$$R = \frac{TP}{TP + FN} \quad (7)$$

其中，TP指把正样本预测为正的数，FN指把正样本预测为负的数量。

3.3.2 精确率 P

精确率 P 也称查准率，表示检测出的正样本有多少是真正的正样本，一般用式(8)表示：

$$P = \frac{TP}{TP + FP} \quad (8)$$

其中，FP是把负样本检测为正样本的数量。

3.3.3 F1值

F1值为召回率和精确率的调和平均值，是综合查全率和精确率的综合指标，反应模型的整体性能。

$$F1 = \frac{2P \times R}{P + R} \quad (9)$$

3.3.4 mAP值

AP为平均精准度，mAP即所有类别的AP的平均值，也称平均准确度。

3.4 实验结果分析

3.4.1 模型性能分析

将原YOLOv3模型与改进YOLOv3模型在本文所述的环境中训练300轮(epoch)，两模型整体损失GIoU (generalized intersection over union)下降的曲线如图5所示。可以看出，原模型的损失下降稍慢，且不够稳定，存在波动，100 epoch之后曲线呈阶梯状，有升有降。而改进模型因修改了损失函数与非极大值抑制函数，模型在反向传播时能够迅速调整，其损失函数曲线下滑平滑稳定且下降速度较快，在训练接近300 epoch时基本收敛，改进模型的边框损失函数GIoU曲线非常平滑，收敛性较好。

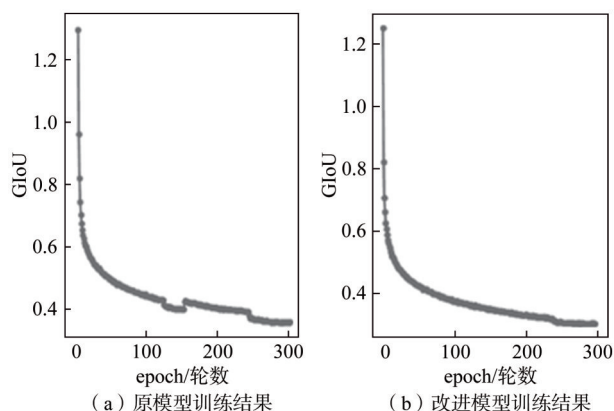


图5 原模型与改进模型训练结果

3.4.2 试验结果分析

用第3.4.1节训练的300 epoch的原模型和改进模型对样本中的测试集进行测试，部分类别测试结果见表1，其中All为所有类别样本的识别结果。

表1 原模型与改进模型的测试结果

类别	原模型				改进模型			
	R	P	F1	mAP	R	P	F1	mAP
All	0.81	0.762	0.786	0.934	0.91	0.867	0.888	0.955
Sedan	0.808	0.78	0.794	0.95	0.931	0.87	0.899	0.96
SUV	0.81	0.743	0.775	0.903	0.91	0.856	0.882	0.919
Taxi	0.803	0.776	0.789	0.95	0.899	0.89	0.894	0.95
Bus	0.829	0.763	0.795	0.966	0.917	0.907	0.912	0.985
Van	0.8	0.69	0.741	0.901	0.893	0.812	0.851	0.961

由表1可知，数据集中各类样本的原模型测试效果并不相同，这一结果与样本数量有很大关系，“Sedan”的样本数量最多，其测试效果在几种车型中较好，而“Bus”样本数量虽不多，但这几类车型中，“Bus”类别样本特征较明显，且外形较大，故测试效果要明显优于“Van”。而改进模型中，测试的召回率 R 、精确率 P 、F1值和mAP值均有所改善，且各种样本的测试参数均有提升，因在损失函数中增加了权重思想，加大了少数样本的权重，因而样本数目最少的“Van”和“Bus”类别的测试效果提升较大，其中“Van”的mAP较原模型提高了0.06，F1值提高了0.11。

为比较本文改进算法模型与其他算法在测试集中的表现，本节对改进模型、原YOLOv3模型、YOLOv3-spp、YOLOv3-tiny和Faster-RCNN等5种模型训练300轮，并在公开数据集进行测试，结果见表2。

表2 不同算法的测试结果

算法	R	P	F1	mAP	检测速度/fps
YOLOv3	0.81	0.762	0.785	0.934	17
YOLOv3-spp	0.878	0.812	0.843	0.951	16.8
YOLOv3-tiny	0.77	0.75	0.76	0.898	18.3
Faster-RCNN	0.879	0.84	0.859	0.953	2.4
改进模型	0.91	0.867	0.888	0.955	16.7

由表2可知, YOLOv3-tiny和YOLOv3检测速度较快, 但是检测效果较差。改进模型检测精确率分别比Faster-RCNN与YOLOv3-spp模型高0.027、0.055, 召回率分别比Faster-RCNN与YOLOv3-spp模型高0.031、0.032。改进模型检测效果优势明显, 且能满足检测实时性要求。

为验证改进模型对重叠或有遮挡的目标的检测能力, 将数据集中拥堵路段的图像分离出来, 记作数据集1, 并将训练好的原模型与改进模型分别在数据集1上进行测试, 测试结果见表3。

表3 拥堵场景下原模型与改进模型的测试结果

类别	原模型				改进模型			
	R	P	F1	mAP	R	P	F1	mAP
All	0.8	0.751	0.775	0.899	0.887	0.802	0.842	0.935
Sedan	0.811	0.762	0.786	0.915	0.91	0.83	0.868	0.953
SUV	0.799	0.757	0.777	0.921	0.902	0.811	0.854	0.94
Taxi	0.81	0.748	0.778	0.901	0.88	0.801	0.839	0.942
Bus	0.812	0.76	0.785	0.89	0.887	0.79	0.836	0.92
Van	0.768	0.728	0.747	0.868	0.856	0.778	0.815	0.92

由表3可知, 拥堵场景下, 原模型的整体表现不如表2混合场景的测试结果, 各类车辆在拥堵场景下的精确率和召回率都相对较低。因修改了函数, 改进模型在拥堵场景下的检测效果要优于原模型, 模型性能有较大提升。

另外, 在数据集中随机抽取部分图像, 用原模型与改进模型对图像进行检测, 检测结果如图6、图7所示。

由图6、图7可知, 在自然光线和夜间光线情况下, 改进模型可检测出原模型遗漏部分, 即远处车辆。故改进模型在一定程度上解决了YOLOv3模型检测时对小尺寸目标不敏感的问题。原模型平均检测速度为0.52 s, 改进模型平均检测速度为0.53 s。改进模型检测速度基本不受影响, 但对小尺寸目标检测效果更好。这对于实际场景中远处车辆的检测和分类识别具有重要意义。

4 结束语

本文基于YOLOv3算法, 改进自然场景下车辆检测与车型识别模型, 在公开数据集UA-DETRAC和自制数据集上进行训练测试, 得出以下结论。

(1) 对模型损失函数中的宽高损失部分进行优化, 使模型的宽高损失与目标尺寸线性度更高。结果表明, 改进模型在目标车辆的检测框定位效果及小数目样本的检测效果方面优于原模型, 其中



(a) 原模型检测结果一



(b) 原模型检测结果二



(c) 改进模型检测结果一



(d) 改进模型检测结果二

图6 光线较好时原模型与改进模型的检测结果



图7 夜晚场景下原模型与改进模型的检测结果

“Van”的mAP较原模型提高了0.06, F1值提高了0.11。

(2) 对原模型的非极大值抑制函数进行修正, 使模型不直接删除IoU较大的检测框, 而再次进行计算, 可减少自然场景下相互遮挡车辆的漏检率。结果表明, 这种改进对遮挡情况下的车辆检测与识别效果好, 且模型运行更为稳定。

(3) 改进模型在自然场景下取得较原模型更高的准确率, 改进模型检测速度基本不受影响, 平均检测耗时增加0.01 s。下一步将着重研究如何在保证检测精度的基础上提高检测速度。

参考文献:

- [1] 刘凯, 李滢东, 林伟鹏. 车辆再识别技术综述[J]. 智能科学与技术学报, 2020, 2(1): 10-25.
- [2] LIU K, LI Y D, LIN W P. A survey on vehicle re-identification[J]. Chinese Journal of Intelligent Science and Technology, 2020, 2(1): 10-25.
- [3] SEVILLA-LARA L, SUN D Q, JAMPANI V, et al. Optical flow with semantic segmentation and localized layers[C]//Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2016: 3889-3898.
- [4] 于晓明, 李思颖, 史胜楠. 混合高斯融合三帧差的运动目标检测改进算法[J]. 红外技术, 2019, 41(3): 256-261.
- [5] YU X M, LI S Y, SHI S N. An improved algorithm for moving target detection using a Gaussian mixture with three-frame difference[J]. Infrared Technology, 2019, 41(3): 256-261.
- [6] SONG Z Y, ALI S, BOUGUILA N. Background subtraction using infinite asymmetric Gaussian mixture models with simultaneous feature selection[J]. IET Image Processing, 2020, 14(11): 2321-2332.
- [7] LUO J H, LIN C H. Pure FPGA implementation of an HOG based real-time pedestrian detection system[J]. Sensors(Basel), 2018, 18(4): 1174.
- [8] 曾娟, 李守义, 张洪昌. 图像分区域多特征融合斜向车辆检测算法研究[J]. 公路交通科技, 2020, 37(8): 99-107.
- [9] ZENG J, LI S Y, ZHANG H C. Study on sub-region multi-feature image fusion oblique vehicle detection algorithm[J]. Journal of Highway and Transportation Research and Development, 2020, 37(8): 99-107.
- [10] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]//Proceedings of 2014 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2014: 580-587.
- [11] GIRSHICK R. Fast R-CNN[C]//Proceedings of 2015 IEEE International Conference on Computer Vision. Piscataway: IEEE Press, 2016: 1440-1448.
- [12] REN S Q, HE K M, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[C]//Proceedings of IEEE Transactions on Pattern Analysis and Machine Intelligence. Piscataway: IEEE Press, 2016: 1137-1149.
- [13] HE K M, ZHANG X Y, REN S Q, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(9): 1904-1916.
- [14] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 779-788.
- [15] 田思佳, 顾强, 胡蓉, 等. 一种基于深度学习的机械臂分拣方法[J]. 智能科学与技术学报, 2020, 2(3): 268-274.
- [16] TIAN S J, GU Q, HU R, et al. A robot sorting method based on deep learning[J]. Chinese Journal of Intelligent Science and Technology, 2020, 2(3): 268-274.
- [17] 刘肯, 何姣姣, 张永平, 等. 改进YOLO的车辆检测算法[J]. 现代电子技术, 2019, 42(13): 47-50.
- [18] LIU K, HE J J, ZHANG Y P, et al. Improved YOLO vehicle detection algorithm[J]. Modern Electronics Technique, 2019, 42(13): 47-50.
- [19] REDMON J, FARHADI A. YOLO9000: better, faster, stronger[C]//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2017: 6517-6525.
- [20] REDMON J, FARHADI A. YOLOv3: an incremental improvement[J]. arXiv preprint, 2018, arXiv: 1804.02767.
- [21] 阮激扬. 基于YOLO的目标检测算法设计与实现[D]. 北京: 北京邮电大学, 2019.
- [22] RUAN J Y. Design and implementation of object detection algorithm based on YOLO[D]. Beijing: Beijing University of Posts and Telecommunications, 2019.
- [23] 杨春玉. 基于卷积神经网络的物体识别研究与实现[D]. 成都: 电子科技大学, 2016.
- [24] YANG J Y. Research and implementation of object detection based on convolutional neural networks[D]. Chengdu: University of Electronic Science and Technology of China, 2016.

[作者简介]



宋华杰（1990- ），女，西安铁路职业技术学院助教，主要研究方向为城轨车辆与车辆信息识别。



周磊（1981- ），男，西安铁路职业技术学院助教，主要研究方向为城轨车辆检修与驾驶。