

基于 YOLOv5s 融合 SENet 的车辆目标检测技术研究

赵璐璐¹, 王学营², 张翼¹, 张美月¹

(1. 长安大学信息工程学院, 陕西 西安 710064;
2. 内蒙古自治区交通建设工程质量监测鉴定站, 内蒙古 呼和浩特 010050)

摘 要: 针对交通监控视频的 车辆目标检测技术在早晚高峰等交通拥堵时段, 车辆遮挡严重且误、漏检率较高的问题, 提出一种基于 YOLOv5s 网络的改进车辆目标检测模型。将注意力机制 SE 模块分别引入 YOLOv5s 的 Backbone 主干网络、Neck 网络层和 Head 输出端, 增强车辆重要特征并抑制一般特征以强化检测网络对车辆目标的辨识能力, 并在公共数据集 UA-DETRAC 和自建数据集上训练、测试。将查准率、查全率、均值平均精度作为评价指标, 结果显示 3 项指标相比于原始网络均有明显提升, 适合作为注意力机制的引入位置。针对 YOLOv5s 网络中正、负样本与难易样本不平衡的问题, 网络结合焦点损失函数 Focal loss, 引入 2 个超参数控制不平衡样本的权重。结合注意力机制 SE 模块和焦点损失函数 Focal loss 的改进检测网络整体性能提升, 均值平均精度提升了 2.2 个百分点, 有效改善了车流量大时的误检、漏检指标。

关 键 词: 车辆检测; 交通监控; 注意力机制; 焦点损失函数; YOLOv5 模型

中图分类号: TP 391

DOI: 10.11996/JGj.2095-302X.2022050776

文献标识码: A

文章编号: 2095-302X(2022)05-0776-07

Vehicle target detection based on YOLOv5s fusion SENet

ZHAO Lu-lu¹, WANG Xue-ying², ZHANG Yi¹, ZHANG Mei-yue¹

(1. School of Information Engineering, Chang'an University, Xi'an Shaanxi 710064, China;

2. Inner Mongolia Autonomous Region Traffic Construction Engineering Quality Monitoring and Appraisal Station, Hohhot Inner Mongolia Autonomous Region 010050, China)

Abstract: To address the problem that the vehicle target detection technology of traffic monitoring videos has high rates of false detection and missed detection due to serious vehicle occlusion in traffic congestion periods such as morning and evening peaks, an improved vehicle target detection model based on YOLOv5s network was proposed. The attention mechanism SE module was introduced into the Backbone network, Neck network layer, and Head output of YOLOv5s, respectively, thus enhancing the important features of the vehicle and suppressing the general features. In doing so, the recognition capability of the detection network for the vehicle target was strengthened, and training and tests were performed on the public data set UA-DETRAC and self-built data set. The results show that the

收稿日期: 2022-03-08; 定稿日期: 2022-05-09

Received: 8 March, 2022; Finalized: 9 May, 2022

基金项目: 2020年度陕西省交通运输厅科研项目(20-24K, 20-25X); 内蒙古自治区交通运输发展研究中心开放基金项目(2019KFJJ-003)

Foundation items: Scientific Research Project of Shaanxi Provincial Department of Transportation in 2020 (20-24K, 20-25X); Open Fund of Inner Mongolia Autonomous Region Transportation Development Research Center (2019KFJJ-003)

第一作者: 赵璐璐(1998-), 女, 硕士研究生。主要研究方向为基于深度学习的目标检测。E-mail: 2689797652@qq.com

First author: ZHAO Lu-lu (1998-), master student. Her main research interest covers object detection based on deep learning. E-mail: 2689797652@qq.com

通信作者: 王学营(1991-), 男, 博士研究生。主要研究领域为沥青路面新材料及检测。E-mail: 2020124099@chd.edu.cn

Corresponding author: WANG Xue-ying (1991-), Ph.D candidate. His main research interests cover new asphalt pavement materials and testing. E-mail: 2020124099@chd.edu.cn

three indicators were significantly enhanced compared with the original network, which was suitable for the introduction of the attention mechanism. The evaluation rate, the value, and mean average accuracy were evaluated, and the results showed that compared with the original network, the three indicators were significantly improved, suitable for the introduction of attention mechanisms. To address the imbalance between positive and negative samples and that between difficult and easy samples in YOLOv5s network, the network combined the focus loss function Focal loss and introduced two super-parameters to control the weight of unbalanced samples. Combined with the improvement of attention mechanism SE module and focus loss function, the overall performance of the detection network was improved, and the average accuracy was improved by 2.2 percentage points, which effectively improves the index of false detection and missed detection in the case of large traffic flow.

Keywords: vehicle detection; traffic monitoring; attention mechanism; focus loss function; YOLOv5 model

交通监控视频提供的对缓解城市交通拥堵、提高道路通行效率以及合理分配交通资源具有重要作用。基于交通监控视频的车辆目标检测是后续进行车辆跟踪、道路车流量统计的基础。基于交通监控视频的车辆目标检测,对检测的实时性要求较高^[1],且易受复杂环境背景及天气光线等干扰,尤其是城市交通拥堵路段,环境复杂、车流量,车辆互相遮挡严重,对车辆目标检测提出了挑战。

目标检测技术是基于交通监控视频的车辆检测核心技术。传统目标检测方法,输入一张待检测图片,首先采用滑动窗口的方式对图片进行候选框提取,接着提取每个候选框中的特征信息,最后利用分类器进行识别。典型的检测算法有 Haar+Adaboost^[2], Hog+SVM^[3]和 DPM^[4]。传统方法采用滑动窗口操作^[5]易导致算法产生大量冗余的候选框,使得检测速度慢、效率低、消耗资源多,而且传统算法采用基于手工设计特征提取方法的鲁棒性低、泛化效果差^[6]。

LEE^[7]提出的背景差分法利用当前图像与背景之间的差距识别车辆,在背景情况复杂时,难以应对车流量大的识别场景。TSAI 和 LAI^[8]提出的帧间差分法,其需要选定合适的时间间隔,且容易出现漏检、错检问题。

随着机器学习和 GPU 并行计算技术的蓬勃发展,基于学习的特征提取技术方兴未艾。目前基于深度学习的目标检测主要包括:基于区域建议的 Two-Stage 检测算法和基于回归思想的 One-Stage 检测算法^[9]。两阶段算法需要先生成预选框,再进行细粒度的物体检测,检测精度高,但效率低,其代表算法有: R-CNN^[10], Fast R-CNN^[11]和 Faster R-CNN^[12]。单阶段算法不必生成预选框,直接在网络中提取特征实现物体分类

和位置的预测。相比两阶段算法,单阶段算法检测速度快,代表算法有 RetinaNet^[13]和 YOLO^[14-15]。其中, YOLO 系列算法基于 PyTorch 框架,便于扩展到移动设备,属于轻量级网络。YOLOv5 包括 YOLOv5s, YOLOv5m, YOLOv5l 和 YOLOv5x 4 种网络结构,随着网络宽度和深度的增大,参数量依次增加。YOLOv5s 是轻量级网络的首选,便于部署到嵌入式设备。

YOLOv5 系列经 COCO2017 测试集测试,平均准确率达 72%,在 GPU Nvidia Tesla V100 的设备上检测速度为每张 2 ms,处于目标检测算法中领先水平。由于单阶段检测算法未生成候选框,所以检测速度快,但精度相比两阶段算法偏低,该系列算法在检测精度方面仍有改进的空间。

本文在 YOLOv5s 的基础上进行改进,以改善在车辆遮挡或小目标情况下检测准确率偏低的问题。首先在 Backbone 主干网络、Neck 网络层以及输出端分别加入注意力机制 SE,测试得到注意力机制的最优引入位置。其次,考虑到单阶段结构存在正负样本和难易样本不平衡的问题,引入 Focal loss 函数计算目标损失和分类损失,以优化训练过程。在公开数据集 UA-DETRAC 进行模型训练和验证,并且在自建数据集上测试算法的性能。

1 YOLOv5s 模型及改进

YOLOv5s 网络采用 One-Stage 结构,由 Input 输入端、Backbone 主干网络、Neck 网络层和 Head 输出端 4 个部分组成,如图 1 所示。Input 输入端具有 Mosaic 数据增强,自适应锚框计算以及自适应图片缩放功能。Backbone 主干网络包括 Focus 结构, CSP 结构^[16]以及空间金字塔池化(spatial pyramid pooling, SPP)^[17]结构,通过深度卷积操作提取图像中的不同层次特征。Neck 网络层由特征

金字塔(feature pyramid networks, FPN)和路径聚合网络(path aggregation network, PAN)组成。Head 作为最后的检测,在大小不同的特征图上预测不同尺寸的目标。

1.1 YOLOv5s 融合注意力机制

YOLOv5s 网络层次不断加深,输出端提取到的信息逐渐抽象,检测监控视频中远处小目标车辆更难以实现,本文在网络中融入注意力机制的方法有效改善了这一问题。

1.1.1 SENet 网络

SENet^[18]是典型的通道注意力网络,曾获得 ImageNet2017 分类比赛冠军。在深度学习神经网络中,并非所有提取的特征都是重要的。SE 注意力机制的作用是增强重要特征、抑制一般特征,并对卷积得到的特征图进行包含 Squeeze, Excitation 和特征重标定 3 步操作^[19],如图 2 所示。Squeeze 对卷积得到的 $C \times H \times W$ 进行全局平均池化,得到 $1 \times 1 \times C$ 大小的特征图。Excitation 使用一个全连接神经网络,对 Squeeze 之后的结果做一个非线性变换。特征重标定使用 Excitation 得到的结果作为权重,与输入特征相乘。SE 模块具有即插即用的便利特征,已经在一些网络中得到应用,但其融合在网络中的哪个部分效果更好,目前还没有完整的理论说明。

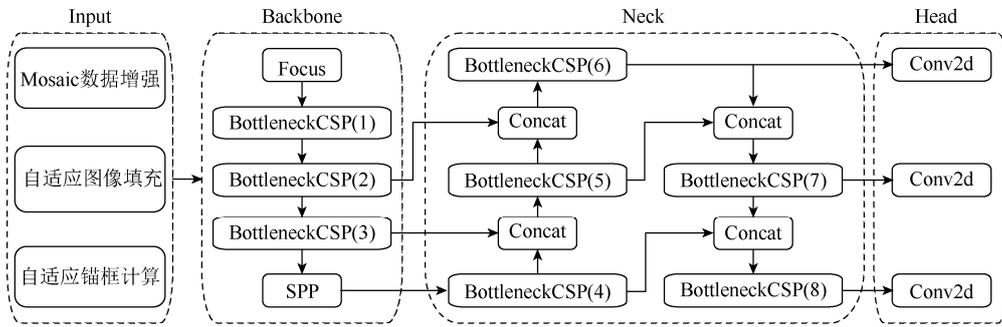


图 1 YOLOv5s 网络结构
Fig. 1 YOLOv5s network structure

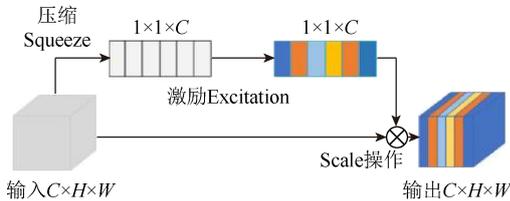


图 2 SENet 网络结构
Fig. 2 SENet network structure

1.1.2 3 种融合方法

针对 YOLOv5s,本文设计了 3 种不同位置的融合方法。将 SE 模块分别融合 Backbone, Neck 和 Head 3 个模块,未考虑 Input 的原因是其只对图像进行预处理,并没有特征提取的作用。由此产生 3 个网络模型,本文分别将其记为 YOLOv5s_A, YOLOv5s_B 和 YOLOv5s_C。

将 SE 模块融合在 Backbone 主干网络形成 YOLOv5s_A。Backbone 的主要作用是通过一个比较深的卷积网络提取图像中的深度特征,随着网络层数的加深,特征图宽度越来越小,深度越来越深,可以使用 SE 模块对不同位置的特征图进行通道注意力重构, BottleneckCSP 结构聚合不同层次特征,因此将 SE 放在 BottleneckCSP 之后,如图 3 所示。

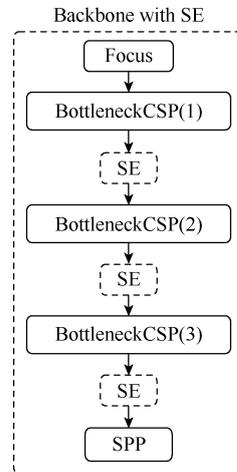


图 3 Backbone 主干网络融合 SE 模块
Fig. 3 Backbone network fusion SE module

将 SE 模块融合 Neck 中形成 YOLOv5s_B。Neck 中的 PAN 和 FPN 结构可以自上而下地传递语义信息,自下而上地传递定位信息,通过 4 个 Concat 操作将深层与浅层信息进行融合,因此将 SE 模块放在 Concat 之后,对融合的特征图进行通道注意力重构,如图 4 所示。

将 SE 模块与网络最后的 Head 融合,形成 YOLOv5s_C。YOLOv5s 通过 3 个尺度大小不同的特征图预测目标,在小特征图上预测大目标,大

特征图上预测小目标, 考虑在预测之前, 对每个特征图进行注意力重构, 如图 5 所示。

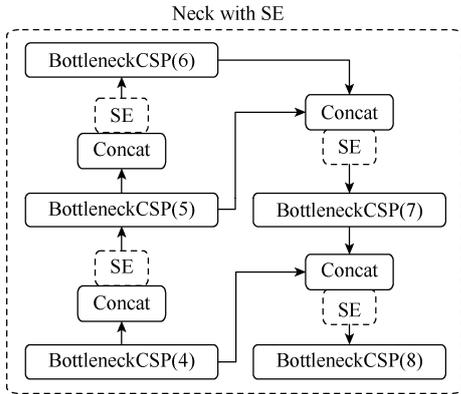


图 4 Neck 层融合 SE 模块

Fig. 4 Neck layer fusion SE module

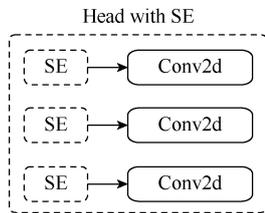


图 5 Head 输出端融合 SE 模块

Fig. 5 Head output fusion SE module

1.2 引入 Focal loss 函数

1.2.1 正负样本不平衡问题

YOLOv5s 的损失包含目标损失、分类损失和边界框回归损失。YOLOv5s 使用 BCE With Logits 作为目标损失函数和分类损失函数, 即

$$p = \frac{1}{1 + e^{-x}}$$

$$BCE(p, y) = \begin{cases} -\ln(p), & y = 1 \\ -\ln(1 - p), & y = 0 \end{cases} \quad (1)$$

其中, p 为经过 Sigmoid 激活函数输出的概率; y 为真实的样本标签, 取值为 0 或 1。

图像中包含车辆的部分为正样本, 其余部分为负样本。对于正样本, 输出概率越大则损失越小; 对于负样本, 输出概率越小损失越小。对于 One-Stage 目标检测算法, 正、负样本不平衡的问题较为突出, 在交通道路图中背景的占比明显大于车辆的占比, 损失函数得出的损失值绝大部分是负样本背景损失, 并且大部分负样本背景是简单易分的, 对于模型的收敛几乎没有作用。因此, 本文引入焦点损失函数 Focal loss, 使用参数平衡正、负样本对损失的影响, 将样本分为难分和易分样本, 降低易分样本对总损失的权重。

1.2.2 Focal loss 函数

对于正、负样本权重的控制, 需要降低大量

负样本对损失的影响, 利用平衡因子

$$\alpha_t = \begin{cases} \alpha, & y = 1 \\ 1 - \alpha, & y = 0 \end{cases} \quad (2)$$

因子 α_t 在样本标签不同时, 提供不同的权重, 例如 BCE With Logits 损失

$$BCE(p, y, \alpha_t) = \alpha_t BCE(p, y) = \begin{cases} -\alpha \ln(p), & y = 1 \\ -(1 - \alpha) \ln(1 - p), & y = 0 \end{cases} \quad (3)$$

通过改变 α_t 的大小控制正、负样本在损失的占比: α_t 在 [0.50, 1] 区间, 能够增加正样本损失的占比, 降低负样本损失的占比。 α_t 在 [0.25, 0.75] 范围, 能够取得较好的 AP 值。LIN 等^[20]分别取值为 0.25, 0.50 和 0.75, 因 0.25 不在 [0.50, 1] 范围内容, 将其舍弃并用后 2 个值进行实验。

α_t 是为了控制正、负样本对损失的贡献, 但不影响易分、难分样本的损失, 因此使用调制因子 $(1-p)^\gamma$ 和 p^γ 控制难分样本和易分样本的权重, 即

$$FocalLoss(p, y, r) = \begin{cases} -(1 - p)^\gamma \ln(p), & y = 1 \\ -p^\gamma \ln(1 - p), & y = 0 \end{cases} \quad (4)$$

其中, γ 取值范围 [0, 5], 通过 γ 控制调制因子的大小, 以控制难分、易分样本损失权重的大小。当 $\gamma = 0$, 式(4)是标准二分类交叉熵损失函数; 当 $0 < \gamma \leq 5$, 可以实现降低易分类样本对损失的贡献, 使得模型更加专注于难分类样本。

将平衡因子 α_t 和调制因子 $(1-p)^\gamma$ 和 p^γ 结合得到最终的 Focal loss, 即

$$FocalLoss(p, y, \alpha_t, \gamma) = \begin{cases} -\alpha(1 - p)^\gamma \log(p), & y = 1 \\ -(1 - \alpha)p^\gamma \log(1 - p), & y = 0 \end{cases} \quad (5)$$

其中, 平衡因子 α_t 可以平衡 One-Stage 检测模型中正、负样本不平衡的问题; 调制因子 $(1-p)^\gamma$ 和 p^γ 控制难易样本差异对损失的影响。

2 机器学习实验结果及分析

2.1 实验准备

(1) 实验平台。网络训练平台采用腾讯云服务器, 规格为 Tesla V100-NVLINK-32G GPU, 40 G RAM。模型框架采用 Python 3.8, PyTorch 1.9, CUDA 10.2。

(2) 数据集。UA-DETRAC^[21]是美国奥尔巴尼大学车辆目标检测和跟踪的数据集, 采集于北京和天津 24 个不同的地点, 包括 100 个具有挑战性的视频, 超过 14 万帧视频图像和 8 250 辆人工标

记的汽车目标, 共计 121 万标记过的目标检测框。由于数据中包含夜晚、晴天、阴雨天等不同天气场景, 以及城市公路、道路交叉口等丰富的交通场景并且拍摄角度接近于监控探头。马芸婷和乔鹏^[22-23]及本实验均选用该数据集。为避免相邻帧之间图像变化过小, 每 10 帧选取一帧的方式, 得到 14 000 张视频图像。从中选取 1 万帧车流量较大的图像作为本文的实验数据集。

(3) 数据标注。为验证本文算法的泛化性能, 模拟监控视频采集于西安市南二环文艺路天桥, 采集场景包括傍晚、阴雨天以及光照较强的晴天总量 1 500 张, 部分数据如图 6 所示。



图 6 数据实例((a) UA-DETRAC; (b)自建数据)

Fig. 6 Data example ((a) UA-DETRAC; (b) Self-built data)

自建数据集采用 LabelImg 标注工具进行人工标注, 生成的 Txt 标签格式如图 7 所示。

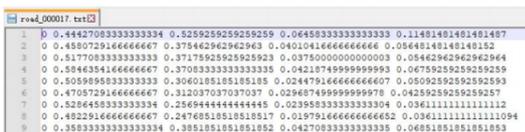


图 7 Txt 数据标签文件

Fig. 7 Txt data label file

标签文件中每行为一个车辆信息, 依次表示车辆类别, 车辆中心坐标 x , y , 以及标注车辆矩形框的宽度和高度。车辆类别 car, bus 和 van 分别对应数字 0, 1 和 2。坐标均被归一化, 将车辆中心坐标 x 和宽度除以图像宽度, 车辆中心坐标 y 和高度除以图像高度。

本实验采用查准率 Precision、查全率 Recall、均值平均精度 mAP 作为评价指标, 即

$$\begin{aligned} Precision &= \frac{Tp}{Tp + Fp} \times 100\% \\ Recall &= \frac{Tp}{Tp + Fn} \times 100\% \\ mAP &= \frac{\sum_{i=1}^k AP_i}{k} \end{aligned} \quad (6)$$

其中, Tp , Fp 和 Fn 分别为正确检测出的车辆数目, 错误检出的车辆数目以及未被正确检出的车辆数目。 AP 为单类别平均精度; mAP 为各类别 AP 的平均值, 用于对所有目标类别检测的效果取平均值, 可以代表检测性能^[24]。

2.2 SE 模块融合对比实验

本文将 UA-DETRAC 数据按照 9:1 的比例随机划分为训练集和测试集, 将自行采集并标注的图像作为测试集。实验均不采用预训练模型, 训练过程使用相同的参数配置, 输入图像大小为 640×640 , 优化器为 SGD, 初始学习率设为 0.01, 动量设为 0.937, 衰 1 减系数为 0.000 5。测试集上的实验结果见表 1。

表 1 融合 SE 的结果对比

模型	Precision(%)	Recall(%)	mAP
Origin YOLOv5s	85.7	95.6	0.952
YOLOv5s_A	86.2	94.8	0.967
YOLOv5s_B	87.8	97.2	0.970
YOLOv5s_C	84.9	94.2	0.940

并非所有融合 SE 的网络均能提升检测效果。YOLOv5s_A 相较于原始 YOLOv5s 网络查准率 Precision 和均值平均精度 mAP 有所提高, 但召回率 Recall 却有明显下降, 图像中车辆被检测到的情况较差。YOLOv5s_B 对于原始网络在 3 个评价指标上均有提升, 均值平均精度 mAP 也是 4 个网络中最高的, 达到 0.970。YOLOv5s_C 在 3 个评价指标中全面落后原始网络。

分析表 1 可知, 注意力机制并不是在网络中的任何位置均有作用, 在 Backbone 中, 网络提取的特征还不够充分, 因此只有 Precision 和 mAP 两项指标提高; 在 Neck 中, 网络对深层和浅层的特征图进行融合, 在此基础上对特征图进行注意力融合, 对不同的通道特征的重要性重新标定, 因此取得了最好地检测结果; Head 中是在不同特征图预测目标之前进行 SE 融合, 但此处的特征图已损失了很多低层的语义信息, SE 模块难以从这种高度融合的特征图中区分出重要的特征通道, 因此所有的指标均有下降, 是 4 个网络中效果最差的一个。基于以上分析本文将 YOLOv5s_B 作为车辆检测的基础模型。

2.3 Focal loss 实验

由式(5)可以看出, Focal loss 主要通过 2 个超参数 α 和 γ 控制正、负样本和难、易样本对损失的贡献, 为了更好地融合 Focal loss 函数与 YOLOv5s

网络适应车辆检测任务的需求, 确定一组最优的 α 和 γ 值成为进一步研究的内容。

实验中网络使用未融合 SE 模块的 YOLOv5s, 损失函数采用 Focal loss 函数, 其他各种参数均与 2.2 节对比实验保持一致。选择几组不同的 α 与 γ 进行对比实验, 即 α 取 0.50 和 0.75, γ 取值 [1,5] 范围的整数, 遵循控制变量法的原则, 各参照组仅有 α 和 γ 的取值不同, 测试结果见表 2。

表 2 不同参数组合的结果对比
Table 2 Comparison of results for different parameter combinations

参数		Precision(%)	Recall(%)	mAP
α	γ			
0.50	1	85.2	93.3	0.950
0.75	1	85.7	94.1	0.953
0.50	2	83.3	93.0	0.939
0.75	2	84.7	93.8	0.942
0.50	3	86.1	95.9	0.968
0.75	3	85.8	94.8	0.957
0.50	4	84.7	94.0	0.932
0.75	4	83.9	93.3	0.935
0.50	5	86.0	95.2	0.962
0.75	5	85.5	94.7	0.957

从表 2 可知, 并非所有超参数 α 和 γ 的组合能对结果产生好的影响。在实验中, 当 $\alpha=0.50$, $\gamma=3$ 时, 网络在测试集上取得了最好的检测结果, 3 项评价指标均高于其他组合, 并且高于原始 YOLOv5s 的结果, 证明了改进损失函数的有效性。

2.4 综合实验结果可视化

对原始 YOLOv5s 同时进行 SE 模块融合和 Focal loss 损失函数改进, 在 UA-DETRAC 训练集上进行训练并在本文自建测试集上进行测试。实验结果见表 3。可以看出, 同时引进 SE 和 Focal loss 之后, 网络的检测结果有了进一步提升, mAP 达到最高 0.974, 相较原始 YOLOv5s 提升了 2.2 个百分点。

表 3 原始网络和改进后网络的结果对比
Table 3 Comparison of the results of the original network and the improved network

模型	Precision(%)	Recall(%)	mAP
Origin YOLOv5s	85.7	95.6	0.952
Proposed	87.9	97.1	0.974

图 8 为自建数据集对部分晴朗白天、阴雨天和夜晚检测结果的可视化。可以看到, 原始网络对一些路段的目标和密集车辆出现了漏、误检现象, 而本文改进网络检测出了这些目标。白天组中标记 1 处, 车辆密集, 原始网络出现了漏检现象, 将多辆车识别为一辆; 在标记 2 处, 将一个车辆目标识别为了 2 辆。本文改进的网络模型解决了这样的问题, 没有出现漏、误检。阴雨天和夜晚监控图像受天气和光线影响, 增加了检测的难度, 原始网络和改进网络检测性能均受到影响。但图中对比显示, 改进网络对远处以及遮挡车辆的误、漏检率总体更低, 具有明显优势。



图 8 原始网络和改进网络的结果对比((a)原始网络; (b)改进网络)

Fig. 8 Comparison of the results of the original network and the improved network ((a) Original network; (b) Improved network)

3 结束语

本文使用改进的 YOLOv5s 网络检测交通监控中的车辆目标。针对检测中出现的误、漏检问题,

提出将注意力模块 SE 引入 YOLOv5s, 为判断 SE 位置不同对检测结果造成的影响, 在网络的 3 个不同位置 Backbone, Neck 和 Head 分别引入并进行对比实验; 使用焦点损失 Focal loss 替代原始的

损失函数, 改善网络的正、负样本和难、易样本不平衡的问题, 设置不同的参数并根据实验结果选择最合适的组合。分别在 UA-DETRAC 和自建数据集上训练、测试。实验结果表明, 本文改进方法相比原始 YOLOv5s 在评价指标 mAP 上提高 2.2%, 根据可视化结果, 本文方法可以有效降低漏、误检率。目前监控探头往往是算力较低的边缘设备, 因此, 在低算力设备上部署车辆检测模型并达到实时检测的要求是下一步的研究重点。

参考文献 (References)

- [1] 蒋镭圻, 彭月平, 谢文宣, 等. 嵌入 scSE 模块的改进 YOLOv4 小目标检测算法[J]. 图学学报, 2021, 42(4): 546-555.
JIANG R Q, PENG Y P, XIE W X, et al. Improved YOLOv4 small target detection algorithm with embedded scSE module[J]. Journal of Graphics, 2021, 42(4): 546-555 (in Chinese).
- [2] TIAN D X, ZHANG C, DUAN X T, et al. The cooperative vehicle infrastructure system based on machine vision[C]//The 6th ACM Symposium on Development and Analysis of Intelligent Vehicular Networks and Applications. New York: ACM Press, 2017: 85-89.
- [3] DROŹDŹ M, KRYJAK T. FPGA implementation of multi-scale face detection using HOG features and SVM classifier[J]. Image Processing & Communications, 2016, 21(3): 27-44.
- [4] FELZENSZWALB P F, GIRSHICK R B, MCALLESTER D, et al. Object detection with discriminatively trained part-based models[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2010, 32(9): 1627-1645.
- [5] ZHU H G. An efficient lane line detection method based on computer vision[J]. Journal of Physics: Conference Series, 2021, 1802(3): 032006-032014.
- [6] 李妮妮, 王夏黎, 付阳阳, 等. 一种优化 YOLO 模型的交通警察目标检测方法[J]. 图学学报, 2022, 43(2): 296-305.
LI N N, WANG X L, FU Y Y, et al. A traffic police object detection method based on optimized YOLO model[J]. Journal of Graphics, 2022, 43(2): 296-305 (in Chinese).
- [7] LEE D S. Effective Gaussian mixture learning for video background subtraction[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2005, 27(5): 827-832.
- [8] TSAI D M, LAI S C. Independent component analysis-based background subtraction for indoor surveillance[J]. IEEE Transactions on Image Processing, 2009, 18(1): 158-167.
- [9] 杨亚峰, 苏维均, 秦勇, 等. 基于语义标签的高铁接触网图像目标检测研究[J]. 计算机仿真, 2020, 37(11): 146-149, 188.
YANG Y F, SU W J, QIN Y, et al. Research on object detection method of high-speed railway catenary image based on semantic label[J]. Computer Simulation, 2020, 37(11): 146-149, 188 (in Chinese).
- [10] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]//2014 IEEE Conference on Computer Vision and Pattern Recognition. New York: IEEE Press, 2014: 580-587.
- [11] GIRSHICK R. Fast R-CNN[C]//2015 IEEE International Conference on Computer Vision. New York: IEEE Press, 2015: 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[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137-1149.
- [13] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42(2): 318-327.
- [14] REDMON J, FARHADI A. YOLOv3: an incremental improvement[EB/OL]. [2022-01-20]. <https://arxiv.org/abs/1804.02767?context=cs.CV>.
- [15] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. YOLOv4: optimal speed and accuracy of object detection[EB/OL]. [2022-02-05]. <https://arxiv.org/abs/2004.10934>.
- [16] WANG C Y, MARK LIAO H Y, WU Y H, et al. CSPNet: a new backbone that can enhance learning capability of CNN[C]//The IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. Washington New York: IEEE Press, 2020: 390-391.
- [17] 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.
- [18] HU J, SHEN L, ALBANIE S, et al. Squeeze-and-Excitation Networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2019, 42(8): 2011-2023.
- [19] 谭芳喜, 肖世德, 周亮君, 等. 基于改进 YOLOv3 算法在道路目标检测中的应用[J]. 计算机技术与发展, 2021, 31(8): 118-123.
TAN F X, XIAO S D, ZHOU L J, et al. Application in road target detection based on improved YOLOV3 algorithm[J]. Computer Technology and Development, 2021, 31(8): 118-123 (in Chinese).
- [20] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42(2): 318-327.
- [21] WEN L, DU D, CAI Z, et al. UA-DETRAC: a new benchmark and protocol for multi-object detection and tracking[J]. Computer Vision and Image Understanding, 2020, 193(C): 102907-102926.
- [22] 马芸婷. 基于深度特征的车辆检测与跟踪[D]. 兰州: 西北师范大学, 2020.
MA Y T. Vehicle detection and tracking based on depth feature[D]. Lanzhou: Northwest Normal University, 2020 (in Chinese).
- [23] 乔鹏. 基于深度学习和边缘任务卸载的交通流量检测研究[D]. 西安: 西安电子科技大学, 2019.
QIAO P. Research on traffic flow detection based on deep learning and edge task offloading[D]. Xi'an: Xidian University, 2019 (in Chinese).
- [24] 谢富, 朱定局. 深度学习目标检测方法综述[J]. 计算机系统应用, 2022, 31(2): 1-12.
XIE F, ZHU D J. Survey on deep learning object detection[J]. Computer Systems and Applications, 2022, 31(2): 1-12 (in Chinese).