

DOI:10.19651/j.cnki.emt.2210357

基于改进 YOLOX-S 的交通标志识别

刘凯 罗素云

(上海工程技术大学机械与汽车工程学院 上海 201600)

摘要: 交通标志是车辆在规范行驶过程中的一个重要指导,交通标志识别是无人驾驶汽车环境感知中必不可少的重要内容。本文基于 YOLOX-S 算法进行改进,通过在主干网络末端添加 CBAM 注意力机制模块,强化特征提取网络所得到的特征;使用 Focal Loss 函数,更好地消除正负样本不均衡问题,挖掘难例样本;使用 GIOU 损失函数,解决了原损失函数存在的优化不一致和尺度敏感的问题,进一步提高模型的识别准确率。本文基于 TT100K 数据集对提出算法进行了实验,对比了几种主流算法与本文算法的识别精度,实验结果表明在具有较高 FPS 的前提下,本文算法对大部分目标类别的检测精度都有所提升。相较于 YOLOX-S 模型,所提模型的 coco 精度评价指标 mAP₅₀ 提升 1.9%,mAP_{50:95} 提升 2.1%,FPS 为 35.6。证明了所做改进的有效性。

关键词: CBAM 注意力机制;Focal Loss;YOLOX-S;GIOU

中图分类号: TP391.4 **文献标识码:** A **国家标准学科分类代码:** 510.8020

Traffic sign recognition based on improved YOLOX-S

Liu Kai Luo Suyun

(College of Mechanical and Automotive Engineering, Shanghai University of Engineering and Technology, Shanghai 201600, China)

Abstract: Traffic sign is an important guide for vehicles in the process of standardized driving. Traffic sign recognition is an essential and important content in the environmental perception of driverless vehicles. Based on YOLOX-S algorithm, this paper strengthens the features obtained from the feature extraction network by adding CBAM attention mechanism module at the end of the backbone network. Utilizes Focal Loss function to better eliminate the imbalance between positive and negative samples and focuses on samples difficult to classify. Using the GIOU Loss function, the problems of inconsistent optimization and scale sensitivity of the original loss function are solved, and the recognition accuracy of the model is further improved. In this paper, the proposed algorithm is tested based on TT100k data set, and the recognition effects are compared with which of several mainstream algorithms. Experimental results show that under the premise of high FPS, the detection accuracy of most target categories is improved. Compared with the YOLOX-S model, the coco accuracy evaluation index Map₅₀ of the proposed model increased by 1.9%, Map_{50:95} increased by 2.1%, and FPS is 35.6. The effectiveness of the improvement is proved.

Keywords: CBAM attention mechanism;Focal Loss;YOLOX-S;GIOU

0 引言

交通标志的检测和识别是现代智能交通系统中不可或缺的一环,也是无人驾驶技术中重要的组成部分。交通标志识别就是通计算机视觉技术将检测的结果反馈给驾驶员或无人驾驶系统。目前学术界和工业界应用最广的方法是基于深度学习对交通标志进行识别。该方法的优点在于卷积网络提取的特征层级深、信息量大、成本低、易于部署等,因此使用深度学习进行交通标志识别已经成为无人驾驶感知的研究热点。

交通标志识别是目标识别的领域之一。基于深度学习

的目标识别按阶段主要分为两类,具体为单阶段(one-stage)算法和两阶段(two-stage)算法。两阶段目标识别算法中具有代表性的有 R-CNN^[1]、Fast R-CNN^[2]、Faster R-CNN 等,王国林^[3]使用 Faster R-CNN^[4] 模型进行交通标志识别,陈朋弟等^[5]提出基于 Mask R-CNN^[6] 提升推理精度做路面交通标志识别。但由于两阶段算法不能很好平衡推理速度与精度的关系,在高速行驶过程中的车辆进行实时检测的情况下并不适用。相较而言,单阶段目标检测算法运行速度快,准确度较高,更符合车辆行驶过程中的实际需求,经典算法有 YOLO 系列算法和 SSD 算法。江金洪

收稿日期:2022-06-17

等^[7]将深度可分离卷积引入 YOLOv3^[8]的算法的特征提取层,实现通道内卷积与通道间卷积之间的分离,从而极大地降低了用来检测交通标志模型的参数量和计算量。王子键^[9]通过基于 SSD^[10]的方法提出了一种基于距离因素的自适应负采样方法提高模型精度用于检测类似交通标志小目标物体,王文胜等^[11]使 ShufflenetV2^[12]的基础单元替换主干后的轻 YOLOV5 网络,并引入轻量级网络 Pelenet^[13]中的 Stem 模块,实现快速行驶中对与交通标识的实时检测,王泽华等^[14]利用空间注意力集中 (spatial attention pooling, SAP) 模块对 YOLOv4^[15]进行改进,再使用 Mobilenet^[16]轻量化网络,从而实现降低模型大小、提升检测速率的目标。

以上方法说明利用计算机视觉识别技术对交通标志进行快速准确识别的可行性,并且其可靠性较高,但是无论是 YOLO 还是 SSD,其算法本身都是一种 Anchor-based 算法,对于输入的一张图像,会生成数百上千的 Anchor boxes,其中只有很少一部分是包含真实目标并被标记为正样本,大量的负样本也分布在图像上,带来了正负样本不均衡的问题。无用的易分反例样本会使得模型的整体学习方向跑偏,导致无效学习并且增加计算量和显存的消耗。

针对以上问题,本文在许多优秀学者的基础上提出使用基于 Anchor-free 的 YOLOX 算法进行在车辆行驶过程中对交通标志牌的检测识别,同时引入卷积块注意模块 (convolutional block attention module, CBAM) 注意力机制强化主干提取到的特征,并使用 Focal Loss 消除正负样本不均衡问题,挖掘难例样本,最后利用并集上的广义交集

(generalized intersection over union, GIoU) 损失函数更好地反映预测框与真实框的重合度,解决了原损失函数存在的尺度敏感问题。实验在网上公开 TT100K 数据集上进行测试,并与 YOLO 经典视觉算法进行比较。

1 相关算法及相关改进

1.1 YOLOX-S 算法

YOLOX 是旷视科技研究院 2021 年提出的高性能检测器,目前其模型权重文件一直在持续优化更新中, YOLOX 在 YOLOv3 的基础上使用 CSPDarknet53 特征提取主干网络,引入 YOLOv5 的 Focus 通道增广技术,创新性地加入解耦头 Decoupled head 和动态正样本匹配方法 SimOTA,运用 Mosaic 和 Mixup 进行数据增强。YOLOX 系列包括 YOLOX-S、YOLOX-M、YOLOX-L、YOLOX-X、YOLOX-Darknet53,并提出两种轻量级网络 YOLOX-Nano、YOLOX-Tiny。本文基于参数量最少的 YOLOX-S 标准网络,对其改进用于对交通标志牌的检测识别。

图 1 是 YOLOX-S 的整体网络结构。其中特征提取网络 (Backbone) 使用的是 CSPDarknet, 加强特征提取网络 (Neck) 采用特征金字塔网络 FPN^[17] (feature pyramid network), 检测头 (Prediction) 运用了解耦头 (Decoupled head)。输入图片在 Backbone 网络中进行特征提取, 输出 3 种不同层级的特征传入 Neck 强化网络中进行上下采样结合的操作来强化输入特征, 之后将经过强化后的图片特征输入经过解耦的 Cls 检测头和 Reg 检测头进行目标框预测, 同时 Reg 分支中增加了 IOU 分支, 至此得到检测结果。

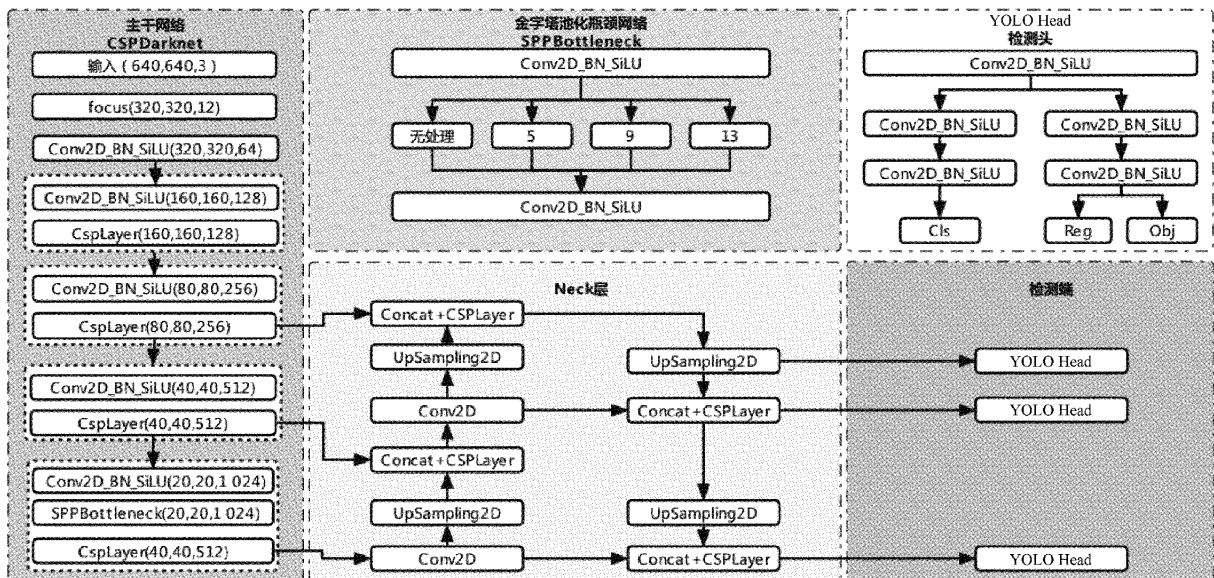


图 1 YOLOX-S 网络结构

1.2 CBAM 注意力机制

CBAM 是一种结合空间和通道的注意力机制模块。在给定一个中间特征图后,该注意力机制模块会沿着两个

独立的维度一次推断出注意力权重,然后与原特征图相乘,对特征进行自适应调整,这也使得 CBAM 更关注目标物体,相比于 SENet^[18]只关注通道上的注意力机制,

CBAM 可以取得更好的效果。由于 CBAM 模块是轻量级的通用模块,因此可以无缝地集成到任何 CNN 架构中,其

额外计算开销可以忽略不计,并且能够与 YOLOX 一起进行端到端的训练。CBAM 的结构概览如图 2 所示。

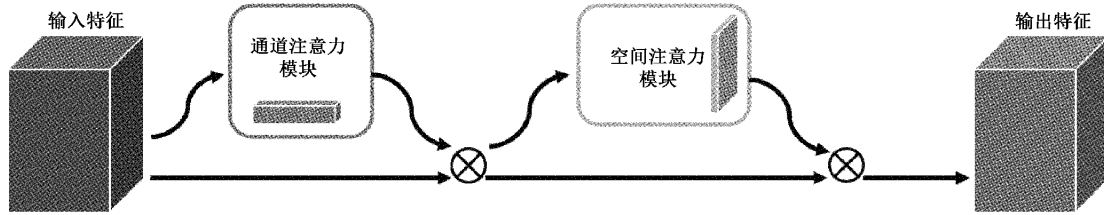


图 2 CBAM 注意力模块流程

CBAM 包含两个独立的子模块,分别是通道注意力模块(channel attention module, CAM)和空间注意力模块(spatial attention module, SAM)。通道注意力模块如图 3 所示,将输入的特征图,分别经过基于宽和高的全局最大池化(Max Pool)和全局平均池化(Avg Pool),得到两个

$1 \times 1 \times C$ 的特征图,然后分别经过 shared MLP,将 MLP 输出的特征进行基于逐像素的加和操作,而后经过 sigmoid 激活,生成最终的通道主义特征,最后将通道注意特征与输入特征做逐像素乘法操作,生成空间注意力模块 SAM 所需要的输入特征。

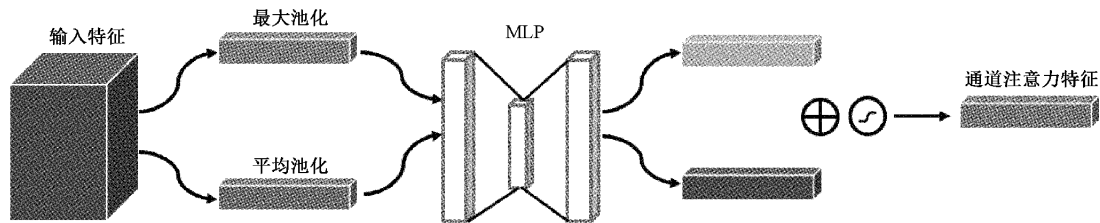


图 3 CAM 注意力模块流程

具体计算公式如下:

$$M_c(F) = \sigma(MLP(AvgPool(F)) + MLP(MaxPool(F))) \quad (1)$$

化简为:

$$M_c(F) = \sigma(W_1(W_0(F_{avg}^c)) + W_1(W_0(F_{max}^c))) \quad (2)$$

式中: σ 为 sigmoid 激活函数, F 为输入特征, F_{avg}^c 指经过平均池化的特征, F_{max}^c 表示经过最大池化的特征, W_0 、 W_1 表示感知机中的两层权重参数。

空间注意力模块如图 4 所示。将 CAM 输出的特征图作为本模块的输入特征图,经过基于通道的全局最大池化和全局平均池化,得到两个 $H \times W \times 1$ 的特征图,其次将两个特征图进行拼接,经过一个 7×7 卷积和一个 sigmoid 激活函数生成空间注意力特征,最后将该特征与输入该模块特征做乘法操作的到图 2 中的合并后的特征(refined feature)。

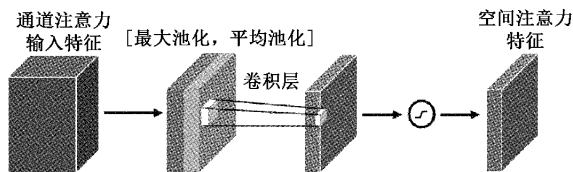


图 4 SAM 注意力模块流程

具体计算公式如下:

$$M_s(F) = \sigma(f^{7 \times 7}([AvgPool(F); MaxPool(F)]))$$

化简为:

$$M_s(F) = \sigma(f^{7 \times 7}([F_{avg}^s; F_{max}^s])) \quad (3)$$

式中: $f^{7 \times 7}$ 为 7×7 卷积层, F_{avg}^s 代表经过平均池化的特征, F_{max}^s 代表经过最大池化的特征。

1.3 Focal Loss 和 GIOU

随着目标检测算法的快速发展和以及划分规则完善,主流目标检测算法主要分为 two-stage 和 one-stage 两大类。前者可以达到很高的准确率,但是速度较慢,很难达到工业界实时性检测标准。后者代表算法为 YOLO 系列,其特点是检测速度很快但准确率却不如后者,因此学术界在不影响检测速度的前提下以此为针对点来提升 one-stage 算法的准确率。Focal Loss 的核心思想认为,one-stage 准确率之所以不如 two-stage 是因为前者正负样本类别不均导致。Focal Loss 是在交叉熵损失基础上修改得到的,通过减少易分类样本的权重,使模型在训练时更专注于难分类的样本。

函数公式如下:

$$FL(p_i) = -\alpha_i(1-p_i)^\gamma \log(p_i) \quad (4)$$

式中: α_i 为类别权重, $(1-p_i)^\gamma$ 为调制系数 γ 为聚焦参数。

GIOU 与 IOU 一样,具有非负性,尺度不变性等特性^[19],除此之外又具备对尺度不敏感的特性。GIOU 是 IOU 损失函数的下界,取值为 $[-1, 1]$,由于惩罚项的引入,在不重叠的情况下,预测框会向真实框移动。综上所述,只有当预测框与真实框完全重合时, GIOU Loss 才会为 0。

GIOU 计算公式如下:

$$GIOU_{Loss} = \frac{|A \cap B|}{|A \cup B|} - \frac{A^c - U}{A^c} \quad (5)$$

式中: A 为矩形框, B 为真实矩形框, A^c 为 A 与 B 的最小外接矩形面积, U 为 A 与 B 并集面积。

1.4 改进的 YOLOX-S 网络

尽管 YOLOX-S 目前更新的最新版本已经具有非常好的检测识别性能,但考虑到网络要应用的场景可能是在车辆高速行驶过程中的这种复杂场景。针对此类复杂场景下的目标识别任务,本文在目前官网最新版本上 YOLOX-S 网络上提出以下改进。

1) 增加 CBAM 模块

YOLOX-S 的主干网络在 4 个不同层级都使用了 CSPLayer 层,因其包含有较大的残差边,在进行深层特征提取的过程中避免梯度消失的同时,可能也会夹杂一些噪音输入给更深的网络层级,这可能会对主干提取特征产生影响。考虑到需要使用官网提供的预训练模型,不能破坏网络主干结构,本文利用 CBAM 模块是轻量级的通用模块,可以无缝地集成到任何 CNN 架构中的属性,在主干特征提取网络 3 个层级的输出端加入 CBAM 注意力模块来尽量抵消 CSPLayer 残差边所带来的噪音影响,以强化特征提取网络所得到的特征,提升网络对不同尺度特征的提取能力。

2) 使用 Focal Loss

YOLOX-S 是一种基于 Anchor-free 的网络,而其他单

阶段基于 Anchor-base 的算法在聚类时会产生数百上千的 Anchor boxes,这些 Anchor boxes 中正样本所占整体比例非常小,大量的负样本分布在图像上,带来正负样本不均衡的问题。无用的易分反例样本会使得模型的整体学习方向跑偏,导致无效学习,还会增加计算量和显存的消耗。而 YOLOX-S 的 Anchor free 机制将每个单元格位置的预测 Anchor boxes 由 3 个减少到 1 个,并且直接预测 4 个值,分别为 2 个偏置项以及预测框的长和宽。并且在最新版的 YOLOX 选用了 Multi positive 策略,将 Ground truth 中心点所落单元格周围区域划归为正样本区域,大大削弱了正负样本不均衡问题。但在某些检测远距离小物体场景下,图片中负样本的个数依旧是大大超过正样本个数,因此本文加入 Focal Loss 来抵消正负样本所带来的影响,挖掘难例样本,让网络更专注于难分类的样本。

3) 采用 GIOU 损失函数

YOLOX 中使用的是 IOU 损失函数,考虑到本次实验数据集属于密集小目标类别,本文使用能更好反映非重叠区域重合度的 GIOU 损失函数,并且 GIOU 损失函数不仅具有 IOU 损失函数的特性,还具备对尺度不敏感的特性。

改进后的 YOLOX-S 网络如图 5 所示。输入经过 CSPDarknet 主干提取 3 种层级特征,输入特征到注意力模块中,经过 CBAM 机制强化所提取特征,强化后的特征输入进 Neck 网络,通过上采样与上层特征进行融合,传达位置特征。最后将得到不同层级信息融合后的特征输入进预测端,从而得到预测结果。

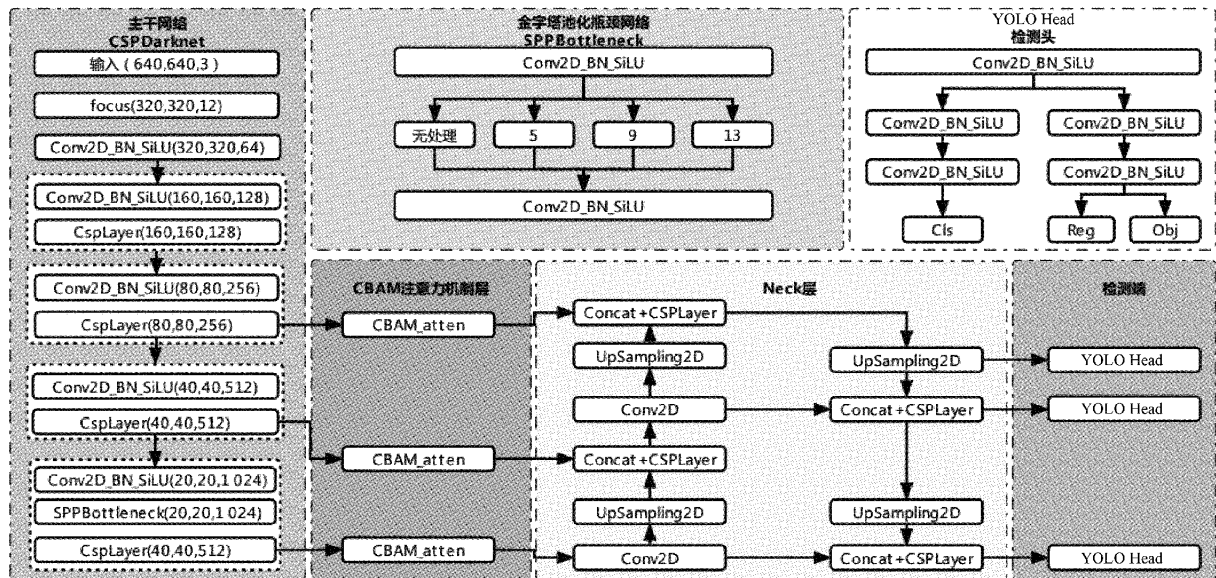


图 5 改进 YOLOX-S 网络结构

2 实 验

2.1 实验环境

本次实验在 Python3.6, CUDA11.1, PyTorch1.9.1 环境下进行。所有网络都在 NVIDIA RTX 3070 GPU 上进

行训练和测试, GPU 个数 2 块。

2.2 数据集及预处理

本文采用了由清华大学制作的 TT100K 数据集,全部数据集大小 19.2 G,因为数据集标签是 json 格式,需要转化为 VOC 格式。本次实验利用 TT100K_to_voc 代码转

化标签格式。按照 VOC 标签格式划分 trainval.txt、train.txt、test.txt、val.txt，划分比例设定为 6 : 2 : 2。由于数据集中存在类别不均衡的情况，其中实例数不超过 10 张的就有 143 类，实例数超过 50 张的只有 46 个类别，所以为了避免样本极度不均衡在对比实验中产生负面影响，本文仅选择数据集实例数排名前 45 的类别。

2.3 数据增强

因为部分常见交通标志做水平翻转、旋转类的数据增强会出现致命错误，如图 6 所示。所以在配置数据增强部分需要将这两部分关掉^[20]。

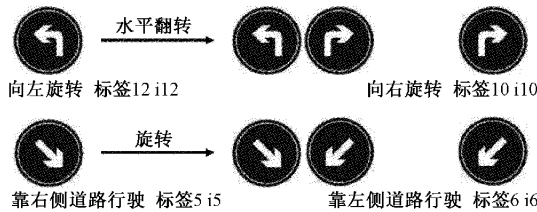


图 6 错误预处理效果图

本次实验采用了 Mosica 数据增强方法，该方法将 4 幅包含有检测目标的预处理图像拼接为一张大图，从而极大地丰富了被检测物体的背景，经 Mosica 数据增强后的效果图如图 7 所示。

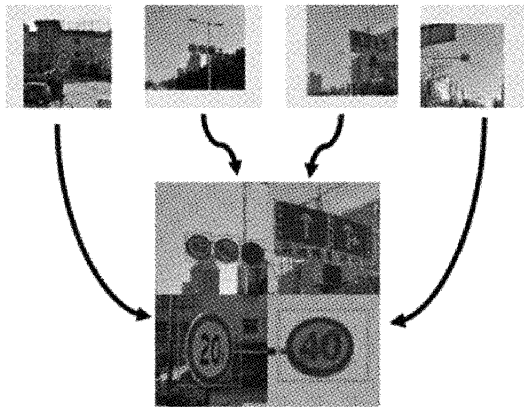


图 7 Mosica 数据增强

2.4 实验设计与参数设定

关于原始 YOLOX-S 网络实验，考虑到经过剔除后的 TT100K 数据集较小，为防止出现过拟合状况，暂定：Train epoch 为 120。图 8 为原始网络模型的训练损失函数。

可以观察到使用 YOLOX-S 与训练模型时在 epoch 达到 100 左右时，Loss 曲线已经趋于平稳状态，模型已经逐渐收敛，并且训练过程中也未出现过拟合状态。

根据以上分析，在改进 YOLOX-S 网络中可将训练 epoch 定为 120。虽然数据增强操作大大丰富了检测数据集，提升了网络的鲁棒性，但考虑到 Mosaic 数据增强生成的训练图片已经远远脱离了自然图片的真实分布，并且 Mosaic 大量的 crop 操作会带来很多不准确的标注框，导

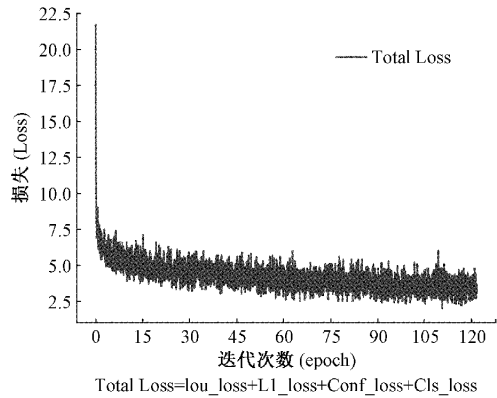


图 8 训练过程 Loss 曲线图

致模型的泛化能力变差，故尝试在最后 20 次迭代中关闭所有数据增强。改进后的 Loss 曲线如图 9 所示。

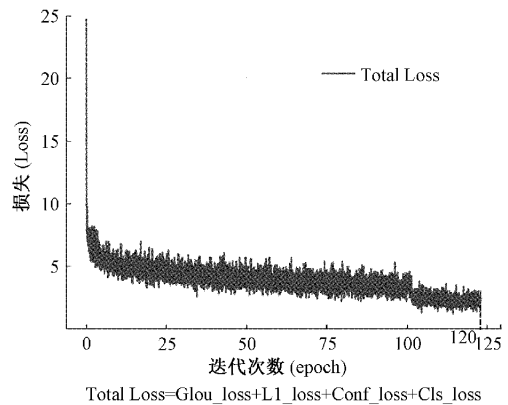


图 9 改进网络训练 Loss 曲线图

图 9 中表现最后 20 次关闭数据增强后 Loss 还有一部分的明显下降。由此证明上述猜想成立，则设定最后 20 次关闭数据增强对模型优化有一定的提升。

综上所述，训练参数设定如表 1 所示。

表 1 训练参数设定表

参数	设定值
max_epoch	120
no_aug_epochs	20
batch_size	16
device	2
input_size	640,640
data_num_workers	4
mosaic_prob	1.0
mixup_prob	1.0
wight_decay	0.000 5
momentum	0.9

本文共设计了一组对比实验和一组逐步改进 YOLOX-S 的消融实验，为符合模型部署与训练要求，本文

实验选用了“-s”轻量级预训练模型。通过消融实验检验本文不同改进部分对网络性能的影响,随后与主流网络 YOLOV5-S, YOLOV3-S 以及原始 YOLOX-S 网络做对比,综合分析改进网络模型的性能。

2.5 评价指标及 AP 曲线

本文采用目标检测模型常用评估指标 mAP(mean average precision)以及 FPS(frames per second)^[21]进行模型评估。AP 指 PR(precision-recall)曲线下面积、mAP 指每个类别 AP 的均值,AP 以及 mAP 值越大越好。Precision 和 Recall 的计算公式如式(6)~(7)所示。

$$Precision = \frac{TP}{TP + FP} \tag{6}$$

$$Recall = \frac{TP}{TP + FN} \tag{7}$$

式中: TP (true positives)表示被分为正样本且分对的; FP (false positives)表示被分为正样本但分错的; FN (false negatives)表示被分为负样本但分错的。

改进 YOLOX-S 对 TT100 验证集的 mAP 曲线如图 10 所示。可以观察到在 AP 50 与 AP 50_95 两个精度标准下,最后 20 次精度都有明显的提升,再次验证最后 20 次迭代过程中关闭数据增强的正确性。

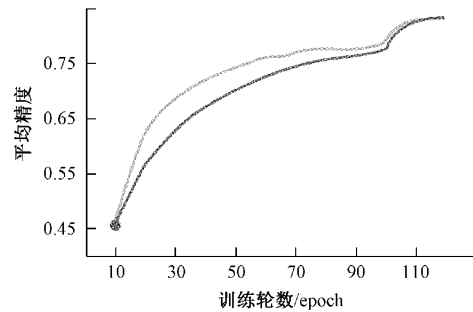
2.6 消融实验

为分析本文所提出改进方法对模型的影响,共设计了 3 组实验对不同改进进行分析。每组实验使用相同的训练参数在不同的模型内容上进行测试。模型性能检测结果如表 2 所示,其中“√”代表在改进模型中使用的策略,“×”代表在改进模型中未使用的策略。

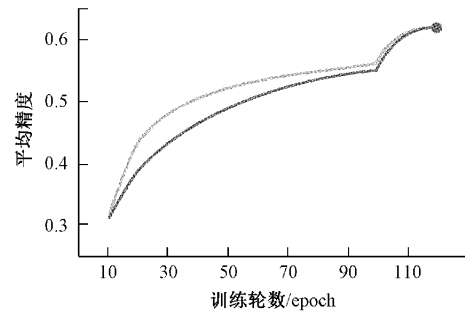
由表 2 分析可知,改进 1 添加了 CBAM 注意力机制,削弱了残差边所带来的噪音,mAP 有所提升。改进 2 在改进 1 的基础上使用 GIOU 损失函数代替原来的 IOU 损失函数来提高对非重叠区域反映重合度的能力。改进 3 在改进 2 的基础上,使用了 Focal Loss 来平衡正负样本,挖掘难例样本,mAP 继续提升。

2.7 模型对比

为证明改进 YOLOX-S 模型的检测性能,将其与主流 YOLO 系列模型做对比实验,针对 i2~wo 45 类交通标志的检测精度与推理速度(FPS)如表 3 所示。其中加粗类别



(a) AP50精度



(b) AP50_95精度

图 10 验证集 AP 曲线

表 2 消融实验结果

方法	CBAM	GIOU	Focal Loss	mAP 50	mAP 50 : 95
YOLOX	×	×	×	0.819	0.613
改进 1	√	×	×	0.824	0.619
改进 2	√	√	×	0.826	0.622
改进 3	√	√	√	0.838	0.634

数值代表改进 YOLOX-S 算法检测精度最高。有 26 类交通标志在对比中显示改进网络的检测精度最高,并且在 AP 50 与 AP 50_95 两种精度评价指标下,平均精度 mAP 都有所提升,其中在前者标准下 mAP 提升 1.9%,在后者评价标准下 mAP 提升 2.1%。改进算法的推理速度为 36.5 FPS,满足复杂交通情况下的实时性检测要求。改进 YOLOX-S 检测效果如图 11 所示。

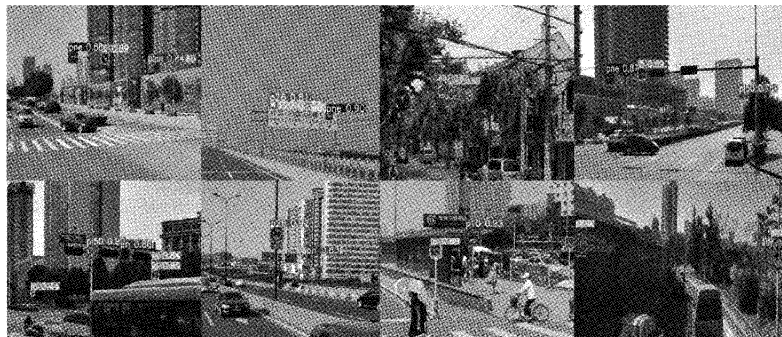


图 11 改进 YOLOX-S 检测效果图

表 3 主流目标检测模型精度对比表

模型	AP(IOU=0.5)											
	i2	i4	i5	il100	il60	il80	io	ip	p10	p11	p12	p19
YOLOV3-S	0.80	0.85	0.85	0.92	0.88	0.98	0.84	0.76	0.70	0.84	0.60	0.90
YOLOV5-S	0.84	0.88	0.87	0.94	0.88	0.96	0.84	0.79	0.70	0.84	0.58	0.90
YOLOX-S	0.86	0.88	0.90	0.98	0.90	0.99	0.88	0.78	0.74	0.84	0.60	0.92
改进 YOLOX-S	0.88	0.89	0.91	0.91	0.98	0.99	0.89	0.77	0.79	0.82	0.52	0.93
p23	p26	p27	p3	p5	p6	pg	ph4	ph4.5	ph5	pl100	pl120	pl20
0.77	0.81	0.98	0.80	0.85	0.73	0.82	0.80	0.81	0.51	0.83	0.83	0.72
0.78	0.81	0.98	0.80	0.85	0.80	0.86	0.80	0.89	0.53	0.88	0.85	0.73
0.79	0.82	0.99	0.80	0.85	0.76	0.84	0.82	0.89	0.54	0.89	0.88	0.74
0.86	0.86	0.99	0.81	0.86	0.73	0.93	0.82	0.80	0.53	0.91	0.89	0.77
pl30	pl40	pl5	pl50	pl60	pl70	pl80	pm20	pm30	pm55	pn	Pne	po
0.72	0.76	0.75	0.81	0.67	0.79	0.81	0.70	0.48	0.72	0.84	0.88	0.72
0.73	0.80	0.76	0.85	0.66	0.79	0.81	0.80	0.53	0.75	0.88	0.90	0.80
0.75	0.80	0.79	0.85	0.69	0.80	0.85	0.73	0.48	0.75	0.89	0.91	0.73
0.78	0.78	0.80	0.84	0.72	0.77	0.86	0.82	0.58	0.81	0.88	0.91	0.69
pr40	w13	w32	w55	w57	w59	wo	mAP_50		mAP_50:95		FPS	
0.91	0.90	0.49	0.81	0.82	0.81	0.79	0.763		0.559		26.3	
0.95	0.90	0.60	0.90	0.80	0.60	0.86	0.806		0.596		36.9	
0.98	0.96	0.55	0.90	0.87	0.91	0.77	0.819		0.613		35.2	
0.97	0.90	0.47	0.85	0.90	0.96	0.90	0.838		0.634		35.6	

3 结 论

本文针对车辆在行驶过程场景下对交通标志的识别提出了一种改进 YOLOX-S 的目标检测模型。为强化提取特征,并且不破坏原有的主干结构而能更好利用预训练模型前提下,该方法在 YOLOX 模型基础上,额外增加了注意力模块,引入了 CBAM 注意力机制;为消除正负样本不均衡采用了 Focal Loss 函数,另外使用了 GIOU 损失函数以更好地反映预测框与真实框的重合度。在数据增强方面,关闭了对数据集有致命影响的水平翻转、旋转类数据增强机制,并且利用损失曲线适应性调节 Mosaic 数据增强。

实验表明,本文所提出的检测方法针对车辆行驶途中等复杂场景下对交通标志识别表现良好,在保证推理速度的同时,检测准确度也具有不小地提升。其中相较于 YOLOX-S 模型,所提模型的 coco 精度评价指标 mAP_50 提升 1.9%,mAP_50:95 提升 2.1%,FPS 为 35.6。下一步将考虑如何在保证高准确度前提下,进一步提高推理速度。

参考文献

- [1] 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, Columbus: IEEE,2014: 580-587.
- [2] GIRSHICK R. Fast R-CNN[C]. Proceedings of 2015 IEEE Conference on Computer Vision, Santiago: IEEE, 2015:1440-1448.
- [3] 王国林.基于 Faster R-CNN 模型的交通标志识别方法[D].西安:西安电子科技大学,2021.
- [4] REN S Q, HE K M, GIRSHICK R, et al. Faster RCNN: Towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 39(6):1137-1149.
- [5] 陈朋弟,黄亮,夏炎,等.基于 Mask R-CNN 的无人机影像路面交通标志检测与识别[J].国土资源遥感, 2020,32(4):61-67.
- [6] HE K M, GEORGIA G, PIOTR D, et al. Mask R-CNN[C]. Proceeding of the IEEE Computer Society Conference on Computer Vision, Machine Intelligence, 2017:2961-2969.
- [7] 江金洪,鲍胜利,史文旭,等.基于 YOLOV3 算法改进的交通标志识别算法[J].计算机应用,2020,40(8): 2472-2478.
- [8] REDMON J, FARHADI A. YOLOv3: An incremental improvement[C]. IEEE Conference on

- Computer Vision and Pattern Recognition, IEEE, 2018;1-6.
- [9] 王子键. 基于 SSD 模型的交通标志识别[D]. 长春: 吉林大学, 2020.
- [10] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot multibox detector[C]. Proceedings of the 14th European Conference on Computer Vision. Amsterdam, The Netherlands: Springer, 2016;21-37.
- [11] 王文胜, 李继旺, 吴波, 等. 基于 YOLOv5 交通标志识别的智能车设计[J]. 国外电子测量技术, 2021, 40(10):158-164.
- [12] MA N, ZHANG X, ZHENG H T, et al. Shufflenetv2: Practical guidelines for efficient architecture design[C]. Proceedings of the European Conference on Computer Vision(ECCV), 2018; 116-131.
- [13] WANG R J, LI X, LING C X. Pelee: A real-time object detection system on mobile devices [J]. In Advance in Neural Information Processing Systems, 2018; 1963-1972.
- [14] 王泽华, 宋卫虎, 吴建华. 基于改进 YOLOv4 网络的轻量化交通标志检测模型[J]. 电脑知识与技术, 2022, 18(5):98-104.
- [15] BOCHKOVSKIY A, WANG C Y, LIAO H. YOLOv4: Optimal speed and accuracy of object detection[J]. ArXiv Preprint, 2020, ArXiv:2004. 10934.
- [16] SANDLER M, HOWARD A, ZHU M, et al. Mobilenetv2: Inverted residuals and linear bottlenecks[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Piscataway IEEE, 2018;4510-4520.
- [17] LIN T Y, PIOTR D, ROSS G, et al. Feature pyramid networks for object detection[C]. Proceedings of the IEEE Computer Society Conference on Computer Vision, IEEE, 2017;2117-2125.
- [18] JIE H, LI S, GANG S, et al. Squeeze and excitation networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 42(8): 7132-7141.
- [19] 王宸, 唐禹, 张秀峰, 等. 基于改进 EfficientNet 的锻件磁粉探伤智能检测方法研究[J]. 仪器仪表学报, 2021, 42(9):89-96.
- [20] 程换新, 蒋泽芹, 成凯. 基于改进 SegNet 模型的斑马线图像语义分割[J]. 电子测量技术, 2020, 43(23):104-108.
- [21] 张明路, 郭策, 吕晓玲, 等. 改进的轻量化 YOLOv4 用于电子元器件检测[J]. 电子测量与仪器学报, 2021, 35(10):17-23.

作者简介

刘凯, 硕士, 主要研究方向为多传感器感知融合模式识别。

E-mail: l15837070577@126.com

罗素云, 副教授, 硕士研究生导师, 主要研究方向为无人驾驶汽车环境感知。

E-mail: luosuyun@sues.edu.cn