

DOI:10.19651/j.cnki.emt.2312684

融合注意力机制的轻量化道路交通标志检测方法^{*}张建恒¹ 杨智宇¹ 夏利红¹ 梁志威²

(1.重庆工商大学机械工程学院 重庆 400067; 2.重庆清平机械有限责任公司 重庆 401120)

摘要:深度学习在自动驾驶环境感知领域发展逐渐成熟,但在城市道路上存在遮挡、重叠、残缺、小目标等情况下的多目标交通标志检测研究仍是目前关注重点。针对道路上的交通标志检测问题提出改进的YOLOv3目标检测方法:首先在YOLOv3的主干网络中引入深度可分离卷积层,优化卷积神经网络中的计算参数以减少计算量;其次在主干网络中的残差模块后引入CBAM空间通道注意力机制,增强网络对弱小特征信息的提取能力,提高对小目标交通标志的检测精度;最后改进原网络中的IOU交并比函数,引入CIOU交并比函数,减少候选框筛选不准的问题,提高目标检测准确率。实验基于CSTSDB开源交通标志数据集和部分自建数据集进行,实验结果表明,改进后的YOLOv3网络相比原YOLOv3检测算法对道路多目标交通标志的准确率提高了7%,并且对重叠、遮挡、小目标等交通标志漏检率更低,速度更快,有一定实际意义。

关键词:深度学习;环境感知;深度可分离卷积;注意力机制

中图分类号: TP391 **文献标识码:** A **国家标准学科分类代码:** 530.2040

Lightweight road traffic sign detection method with attention mechanism

Zhang Jianheng¹ Yang Zhiyu¹ Xia Lihong¹ Liang Zhiwei²

(1. School of Mechanical Engineering, Chongqing Technology and Business University, Chongqing 400067, China;

2. Chongqing Qing ping Machinery Co., Ltd., Chongqing 401120, China)

Abstract: Deep learning has been widely used in environment perception of autonomous driving field, Its application on multi-object traffic sign (such as shield, overlap, incomplete, and small goals) detection on urban roads is current research emphasis. This paper proposed an improved detection method based on YOLOv3. Firstly, it introduced the depthwise separable convolution layer into the backbone network of YOLOv3 to optimize the parameters and reduce the quantity of calculation in convolution neural network. Secondly, the spatial channel attention mechanism (CBAM) was introduced after the residual module in the backbone network to, aiming to Enhance the ability of the network to extract weak feature information, and improve the detection accuracy of small target traffic signs. Finally, the optimized intersection ratio function of IOU combining with the CIOU function, can improve the target detection accuracy of the candidate box screening. The experiments were conducted with CSTSDB open source traffic sign dataset and partial self-built dataset. The experimental results show that the improved YOLOv3 network improves the accuracy rate has been improved by about 7% than the original YOLOv3 detection algorithm, and has a lower leakage detection rate, faster speed, which has some practical significance.

Keywords: deep learning; environmental perception; depthwise separable convolution; attention mechanism

0 引言

道路的交通标志检测是自动驾驶和智慧交通中至关重要的环节,对智能汽车驾驶安全起着重要的作用。现代的道路发展迅速,交通标志也随之剧增。交通标志具有目标

小、形状各异,颜色不同等特点,相比道路中常规的车辆和行人更难检测。研究人员也对应提出不同的检测方法。传统的交通标志识别方法主要通过人为定义图像特征,如方向梯度直方图(histogram of oriented gradient, HOG)、尺度不变特征变换(scale-invariant feature transform,

收稿日期:2023-01-29

^{*} 基金项目:重庆市自然科学基金面上项目(cstc2020jcyj-msxmX0803)、重庆市教育委员会科学技术研究项目(KJQN20180812)、重庆市教委科技项目基金(KJQN202100845)、国家级大学生创新创业训练计划(920620009)项目资助

SIFT)、加速稳健特征(speeded up robust features, SURF)等,并结合支持向量机、模板匹配等方法进行识别^[1-4],这些方法的使用场景有局限性,识别精度有待提高,且在复杂、小目标、背景复杂的环境中,无法准确的检测图像中交通标志的位置,对于需要极高实时性自动驾驶来说是难以满足任务要求的。因此研究人员侧重于使用深度学习算法来解决智能汽车环境感知相关的识别任务。当前众多的深度学习目标识别算法可以分为两大类,一种是基于双阶段(two stage)的目标检测算法,主要代表算法包括由 Ross Girshick 等提出的基于区域的卷积神经网络(region-convolution neural network, RCNN)^[5]、快速的区域卷积神经网络(Fast-RCNN^[6])、更快的区域检测网络(Faster-RCNN^[7])、基于区域的卷积神经网络实例分割算法(Mask-RCNN^[8])等,相比于单阶段算法检测精度通常更高,但检测速度也相对较慢。另一种是单阶段(one stage)的目标检测算法,代表算法如 YOLO(you only look once)^[9-11]系列算法、单激发多框探测器(single shot Multibox detector, SSD)^[12]、RetinaNet^[13]等,这类算法不需要事先获取目标候选框,因此具有较高的检测速度。

Hasegawa 等^[14]提出了一种基于深度学习的高精度目标检测和识别方法,在识别多种交通标志任务上具有一定的鲁棒性。Kong 等^[15]提出了一种基于级联卷积神经网络的交通标志识别(TSR)算法,该方法较 YOLOv1 相比降低了约 55% 的计算复杂度^[15]。以上方法在面对图像中存在较大像素的交通标志的识别任务上有较好的识别效果,但当在图像上存在的较小像素的交通标志时,经过多层的卷积和池化计算后,这些小目标的交通标志特征几乎丢失,网络难以检测。为此有学者提出联系上下文信息来增强网络的特征提取能力,提高对小目标的识别准确度^[16-18],但在小目标的交通标志识别任务上提高并不明显。也有学者采用加入注意力机制的方式来提升网络的表征能力,也证明了增加注意力机制可以提升网络对图像中小像素目标的特征提取能力^[19-20],但实时性有所降低。基于传统或深度学习的交通标志检测研究已有较大的进展,识别定位精度也越高,但算法的参数量也随之增大,同时还需考虑检测对象的尺寸特点和位置特征信息。

综上所述,本文以 YOLOv3 主干网络 Darknet-53 作为基础架构进行改进。针对网络计算量大、目标定位不准等问题提出一种改进的 YOLOv3 道路交通标志检测网络,减小目标特征提取的区域,提高模型对交通标志检测和识别的速度和精度,主要改进如下:

1) 在主干网络 Darknet-53 中的残差模块引入深度可分离卷积层,通过多通道 3×3 卷积核及 1×1 的单点卷积,优化网络模型参数,降低计算成本;

2) 在主干网络的检测的 3 个分支中加入空间与通道注意力机制(convolutional block attention module, CBAM),避免了多层卷积后的图像特征信息丢失,增强在复杂道路

环境下交通标志的特征提取能力,提高重叠、残缺、遮挡等标志的检测精度;

3) 改进交并比(intersection over union, IOU)交并比函数,提高目标定位的准确度,同时以解决候选框在某些状态重叠或不相交的情况下无法优化的问题,并提高训练过程中目标检测框的定位精度。

1 YOLOV3 网络结构

YOLOv3 主干网络 Darknet-53 由 53 层卷积层组成,借鉴了残差网络(Resnet)近道连接的思想,同时采用多尺度分类思想,输入 416×416 的图像,输出 3 种不同大小特征图,分别为 13×13 、 26×26 、 52×52 对应检测大、中、小目标。在自然环境下的目标检测能达到 90% 的准确率,而对于高分辨率复杂环境下的道路交通标志的识别任务,YOLOv3 对道路中的不同交通标志的识别错误率较高,难以达到自动驾驶实际应用的要求。

2 交通标志检测网络模型

2.1 改进 YOLOv3 网络结构

为提高对道路交通标志检测精度和速度,提出一种改进 YOLOv3 的交通标志检测网络,如图 1 所示。首先将 $416 \times 416 \times 3$ 大小的图像,输入到主干残差网络中,其中图一虚线部分为改进的 YOLOv3 的主干网络,橙色模块 Res-DW, DW 表示深度可分离卷积层块,将主干网络 Darknet-53 中的每个残差模块中引入深度可分离卷积层,具体结构如图 1 左下角虚线框所示,通过 1×1 与 3×3 的卷积核分布提取图像信息,使特征提取网络在不降低精度的前提下减少主干网络在中的计算参数,具体实现细节如 3.2 节描述;其次将经过特征提取的卷积层输入到 3 个分支检测头里,为避免小目标在图像中经过多层卷积后易丢失特征信息的问题,在 3 个检测分支的前端加入 CBAM 空间与通道注意力机制,如图 1 中 3 个分支的红色模块,相比于其他注意力机制,双通道的 CBAM 注意力更能增强网络对小目标的特征表达能力,以提高对道路中小目标的交通标志的检测精度;为解决在道路交通标志定位不准的问题,引入 CIOU 交并比损失函数,加入中心点距离和最小矩形框参数提高候选框的定位精度问题。改进的 YOLOv3 主干网络提取的特征图颈部网络的 CBAM 模块和 SPP 等其他卷积模块输出 13×13 大小的特征图;输出的 13×13 的特征图经过 2 倍的上采样再与改进的 Darknet-53 网络中提取的第 43 层卷积层的特征图进行合并,得到 26×26 大小的特征图;将 26×26 大小的特征图经过 2 倍的上采样,和改进的 Darknet-53 中第 26 层的卷积层的特征图进行合并,得到 52×52 大小的特征图;由此形成的 3 种尺度的特征图,每个分支输出前加入 CBAM 模块,避免小特征信息丢失,实现对小、中、大不同大小的交通标志检测,具体改进的细节如下。

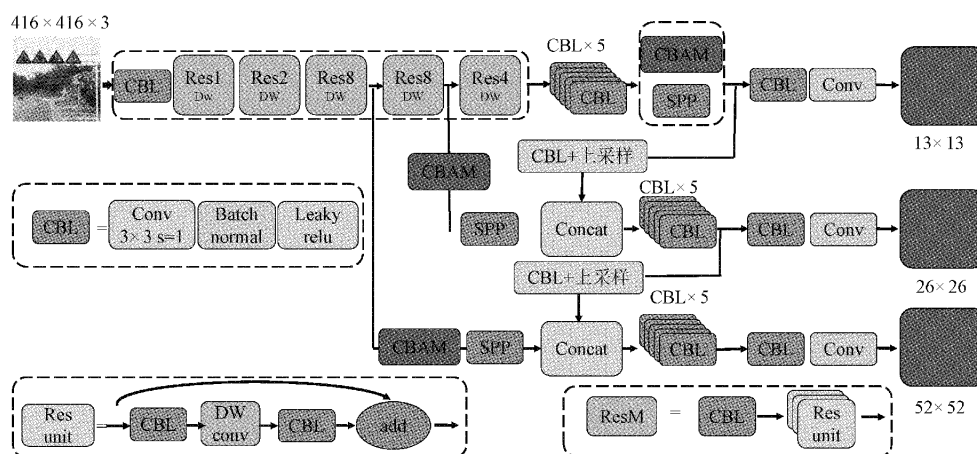


图1 改进的YOLOv3框架结构

2.2 深度可分离卷积

为了提高残差模块下采样后有更快的检测速度,本文将YOLOv3的主干网络中的残差模块中层引入深度可分离卷积层,如图1橙色模块。深度可分离卷积分为单通道卷积(Depthwise Convolution)和逐点卷积(Pointwise Convolution)两部分。单通道卷积是对输入图像进行单通道拆分,对同一通道的特征图只采用一种卷积核进行单通道卷积计算,只提取一个通道内的特征信息,最后将特征图进行合并。逐点卷积是使用和上一层相同通道数的 1×1 的卷积核对特征图进行卷积计算,最后输出多通道的特征图。将普通卷积和深度卷积参数量进行对比,其中普通卷积计算参数量为 $k \times k \times C_1 \times C_2 \times F \times F$,深度可分离卷积的参数量为 $k \times k \times C_1 \times F \times F + 1 \times 1 \times C_1 \times C_2 \times F \times F$,深度可分离卷积核普通卷积的参数量比值为:

$$\frac{k \times k \times C_1 \times F \times F + C_1 \times C_2 \times F \times F}{k \times k \times C_1 \times C_2 \times F \times F} = \frac{1}{C_2} + \frac{1}{k^2} \quad (1)$$

式中: C_1 表输入图像的通道数, C_2 表示输出图像的通道数,卷积核大小为 $k \times k$, $F \times F$ 表示特征图的大小,式(1)中可以看出,深度卷积的计算量为传统卷积的 $\frac{1}{C_2} + \frac{1}{k^2}$,本文输入尺寸为 $416 \times 416 \times 3$ 的图像,输出通道为24,卷积核大小为 3×3 的网络,参数减少约为原来的 $\frac{1}{7}$ 倍,参数量大大减小。

2.3 空间与通道注意力机制

注意力机制主要源于人类在理解图像内容过程中根据不同的注意力确认图像中的重点关注的区域,并且弱化其他区域内容的信息,从而更准确的获取图像信息。对于复杂的道路交通图像,交通标志通常只占较小的像素,容易出现漏检和错检等情况。CBAM注意力机制是将空间注意力机制和通道注意力机制相结合,如图2,空间注意力模块(spatial attention module, SAM)主要对输入图像使用平

均池化和最大池化聚合空间特征信息,得到一维的特征图,并通过卷积计算和sigmoid非线性处理得到空间注意力特征图 $M_s(F) \in R^{1 \times H \times W}$,其中 $H \times W$ 表示特征图的高和宽;通过最大池化和平均池化并聚合成一维特征图 $F_{max} \in R^{1 \times H \times W}$ 和 $F_{avg} \in R^{1 \times H \times W}$,计算空间注意力特征,计算公式如下:

$$M_s(F) = \sigma(f^{7 \times 7}([\text{Avgpool}(F); \text{Maxpool}(F)])) = \sigma(f^{7 \times 7}([F_{avg}^s; F_{max}^s])) \quad (2)$$

式(2)中 σ 表示sigmoid激活函数, $f^{7 \times 7}$ 表示大小为 7×7 的卷积核。

通道注意力模块(channel attention module, CAM)主要通过最大池化和平均池化获取图像的特征描述符: F_{avg}^c 和 F_{max}^c ,同时输入到多层卷积网络得到通道注意力特征图 $M_c(F) \in R^{C \times 1 \times 1}$,经过多层卷积网络和sigmoid非线性激活函数得到通道注意力特征图 $M_c(F) \in R^{C \times 1 \times 1}$,通道注意力特征图 $M_c(F)$ 计算如下:

$$M_c(F) = \sigma(\text{MLP}(\text{Avgpool}(F))) + \sigma(\text{MLP}(\text{Maxpool}(F))) = \sigma(W_1(W_0(F_{avg}^c))) + \sigma(W_1(W_0(F_{max}^c))) \quad (3)$$

式(3)中 σ 表示sigmoid激活函数,MLP表示多层感知机,权重分别为 W_0 和 W_1 , F_{avg}^c 和 F_{max}^c 分别表示平均池化运算和最大池化运算。

通常在检测网络中插入CBAM注意力模块有多种方式,通常在主干网络的最后加入CBAM模块,而对于小目标对象,在前层的卷积计算容易丢失掉小目标的特征信息。本文将CBAM模块添加在改进的Darknet-53网络中的3个分支输出前,避免经过CBL模块等卷积层对特征信息的丢失。

2.4 改进IOU交并比损失函数

对于网络提取出来的候选框,每个目标的位置信息是单独预测,每个预测出来候选框的坐标之间没有关联性,易造成大量冗余的候选框,影响识别效果。因此YOLOv3

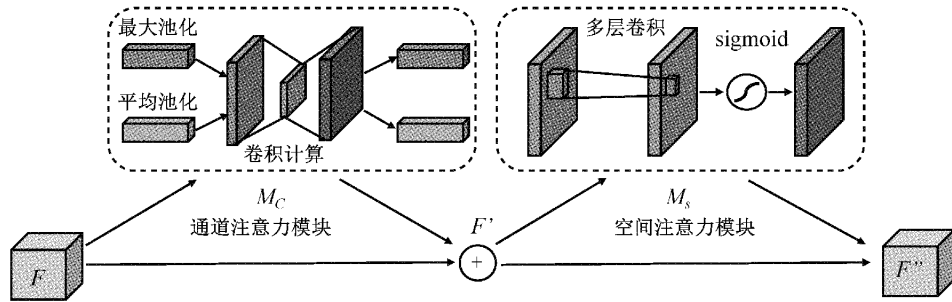


图 2 CBAM 注意力机制结构

中的目标框候选采用 IOU(intersect of union)算法进行筛选,以此来评价预测框与真实框的重合程度,如图 3 所示,其对应的 IOU 损失函数如式(4)。

$$IOU = \frac{|A \cap B^{gt}|}{|A \cup B^{gt}|} \quad (4)$$

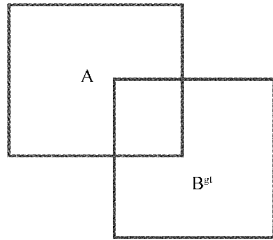


图 3 IOU 示意图

图 3 中 A 表示预测候选目标框, B^{gt} 表示物体所在的真实框,式中, $A \cap B^{gt}$ 表示候选框 A 和真实框 B^{gt} 相交的面积部分, $A \cup B^{gt}$ 表示候选框 A 和真实框 B^{gt} 相并集的部分。因此 IOU 就是一个比值的评价标准,当两个区域的重叠部分越多,交并比值就越大,预测的准确率就越高,目标定位精度更好。但 IOU 的候选框筛选没有考虑两个框完全不相交的情况,如图 4(a),此时损失函数梯度消失,无法更新预测框的位置。此外,当出现图 4(b)和(c)的情况,即预测框大小相同,且 IOU 也相等的情况,IOU 损失函数就不能区分两者情况。

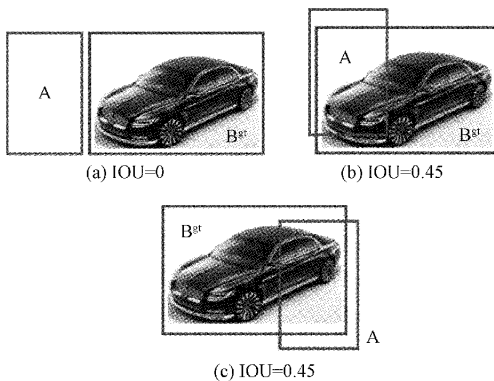


图 4 IOU 缺陷示意图

为解决以上问题,同时为了避免出现预测框在真实框

内部的情况,因此本文引入新的交并比函数, CIOU (complete intersection over union)^[21]。加入了框中心点的对角线距离,如图 5,公式为:

$$CIOU = 1 - IOU + \frac{\rho(b, b^{gt})}{c^2} + \alpha v \quad (5)$$

$$\alpha = \frac{v}{(1 - IOU) + v} \quad (6)$$

$$v = \frac{4}{\pi^2} \left(\arctan \frac{w^{gt}}{h^{gt}} - \arctan \frac{w}{h} \right)^2 \quad (7)$$

其中,式(6)中 $\rho(b, b^{gt})$ 为预测框中心点 b 与真实框中心点 b^{gt} 的欧式距离, c 为物体所在的真实框和预测框的外接矩形对角线长度,式(7)中 α 是正值的权衡参数, v 表示长宽比的一致性。式(8) h^{gt}, w^{gt} 为真实框的高和宽, t, w 为预测框的高和宽, CIOU 的损失函数考虑多个候选框之间的中心距离和长宽比一致性。

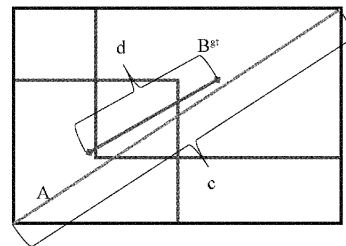


图 5 CIOU 示意图

3 实验结果与分析

本文实验使用设备配置为 linux 系统 18.04, CPU 为 I5, 内存 32 G, GPU 搭载的是 NVIDIA RTX2070 显卡, 显存 6 G, CUDA 版本为 10.1, CuDnn 版本是 8.5, 基于 Pytorch 深度学习框架进行训练, 采用动量梯度下降算法, 动量参数为 0.9

3.1 实验数据集

使用部分开源道路交通标志数据集和自建数据集, 开源数据集来源于中国交通数据集 (CSTSDB), 通过行驶车辆前方的摄像头拍摄的道路交通标志, 包括高速路、城市道路、交叉路口、步行路口等场景, 自建数据集是通过建立部分残缺、遮挡、暗光等环境的样本, 增强模型对特殊场景的检测能力, 一共 4 647 张样本, 包含警告 (warning)、禁止

(prohibitory)、强制(mandatory)3类标志,样本数量分别为1 213、2 403、1 031,并且按5:1的比例分为训练集和测试集,训练集分别为1 026、2 003、860,测试集205、400、171,数据集划分如表1所示。

表1 道路标志数据集划分

分类	列车数据集	测试数据集	类别编号
warning	1 026	205	1
prohibitory	2 003	400	2
mandatory	860	171	3

3.2 实验结果分析

针对原YOLOv3网络中的初始锚框的长宽比适用性问题,根据本文数据集中道路交通标志的形状大小特点,通过K-means聚类算法对数据集的锚框标签重新筛选,结合K-means聚类算法改进目标候选框的初选大小,分别得到不同尺度的锚框,K-means聚类得到锚框大小结果如表2所示。

表2 K-means聚类锚框尺寸

特征图的大小	大目标	中等目标	小目标
52×52	11,18	19,30	63,87
26×26	8,23	15,23	37,54
13×13	8,14	11,31	27,40

为验证改进后的网络对交通标志检测的有效性,设计一组消融实验对比分析原始1)YOLOv3网络、2)添加K-means聚类、3)改进CIoU交并比函数、(4)同时添加K-means聚类、CIoU和深度可分离卷积、5)同时添加K-means聚类、CIoU、深度可分离卷积、CBAM注意力机制。结果如表3所示,并通过模型大小,检测时间,准确率3个参数进行数据对比,由表3可知,仅添加K-means聚类算法相比于原YOLOv3网络的准确率提升3%,说明通过优化了候选框的尺度问题,能有效提高对目标位置的收敛速度;其次,用CIoU替代原来的IOU交并比函数,引入中心距离和长宽一致性也使准确率和检测时间都有较高的提升;引入深度可分离卷积层,模型大小也不到原YOLOv3网络的1/3,检测时间缩短至原网络检测时间的1/4,准确率相对于没有改进深度可分离卷积的网络提高约1.1%;最后在添加CBAM注意力机制后,提升了对小目标的特征提取能力,增加小目标交通标志的检测能力。虽然模型大小和检测时间有所增加,但检测准确率提高1.2%。

针对复杂环境交通道路情况,对正常道路环境的245张和复杂道路环境下276张样本进行测试。主要分析复杂环境下的识别结果,如图7(a)~(d)所示,结果表明小目标、标志残缺、弱光环境、遮挡等情况下,改进后的

表3 改进后网络的参数比较

方法	模型	检测	准确
	大小/MB	时间/ms	率/%
YOLOv3	24.3	15	83.8
YOLOv3+K-means	23.1	26	87.8
YOLOv3-CIoU+K-means	25.4	18	88.1
DW-YOLOv3-CIoU+K-means	16.7	9	89.2
DW-CBAM-YOLOv3-CIoU+K-means	15.9	11	90.8

YOLOv3算法都能检测到多类的交通标志。为增加应用场景,添加人为遮挡和植被遮挡等情况,检测结果如图7(e)~(f),也可看出改进后的网络模型后能有有效的识别残缺、弱光、远距离的小目标交通标志,表明改进后算法的可行性。



图6 改进后网络的检测结果

为进一步验证模型对小目标检测的有效性,增强模型对小目标的识别准确率,在YOLOv3的主干网络中的3个预测分支前添加注意力模块。将改进后的YOLOv3主干网络分别与SE、ECA、CBAM注意力机制结合进行对比实验,模型识别的精确度、loss曲线如图7,可以看出添加CBAM模块后相比SE、ECA注意力机制收敛更快。模型大小、检测时间、识别准确率对比如表4,由表4可知,添加CBAM注意力机制准确率分别比SE、ECA提升约0.7%和0.3%。

由图8(a)可知,加入CBAM注意力机制在准确率会

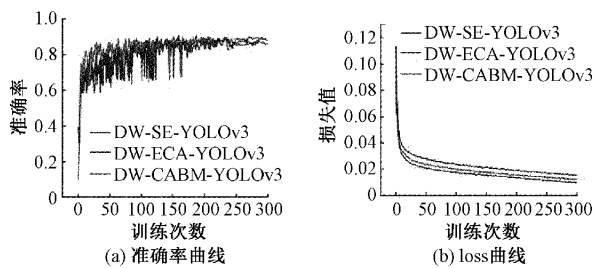


图 7 不同注意力机制的准确度曲线、loss 曲线

略高于 ECA 和 SE 注意力机制,由图 8(b)可以看出由于加入了空间与通道注意力模块,在 Loss 参数上 CBAM 模块下降较慢,收敛速度略低于 SE 注意力机制。由表 4 可知将 CBAM 注意力机制加入到 YOLOv3 模块中,其中 mAP 值相较于 SE 注意力机制提高 1.464%,比 ECA 注意力机制提高 1.42%,在检测精度上相较于 SE 和 ECA 注意力机制分别提高了 0.645% 和 0.22%。在不同的道路环境中,利用 YOLOv3_SE、YOLOv3_ECA 都会有不同程度的漏

检,误检等情况,如图 9。对于图 9 左侧图像,CBAM 注意力机制准确率高 SE 和 ECA 注意力机制;图 10 中间图像中,只有 CBAM 能识别到路边的小目标交通标志,SE 未识别出,ECA 错检广告牌为交通标志;图 9 右侧图像 SE 注意力机制漏检一个交通标志,ECA 漏检 2 个交通标志,CBAM 全部识别。本算法在引入 CBAM 注意力机制和深度可分离卷积以及 CIoU 交并比函数,提高了算法检测的准确率,并实现在不同环境下的识别效果,都表明 CBAM 注意力机制的引入对与小目标识别有较显著的提高。

表 4 不同注意力机制的参数比较

方法	模型	检测	准确
	大小/MB	时间/ms	率/%
YOLOv3	24.3	15	83.8
DW-SE-YOLOv3	15.8	10	90.1
DW-ECA-YOLOv3	15.8	9	90.5
DW-CBAM-YOLOv3	15.9	11	90.8

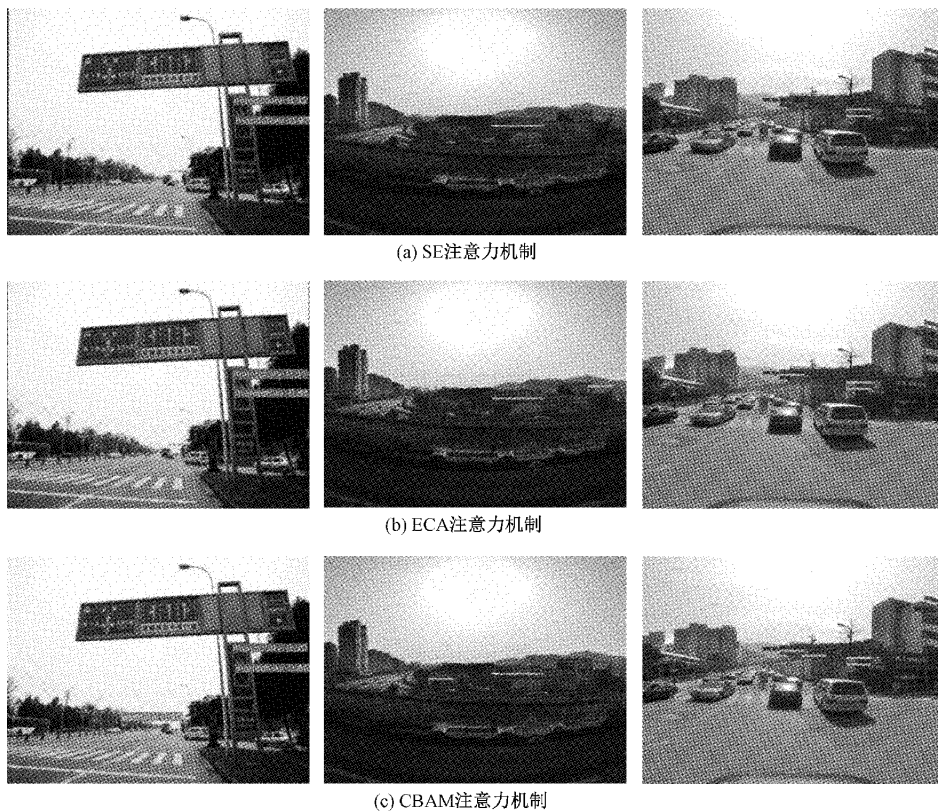


图 8 不同注意力机制识别结果对比图

为了更进一步验证本算法检测效果,与现有的主流检测算法在相同的数据集进行客观指标对比,检测结果如表 5 所示,由表 5 中数据可知,Faster-RCNN 的检测准确率为 84.6%,虽然准确率比单阶段的 YOLOv1 和 YOLOv2 更高,但检测耗时最长,由于采用了 RPN 网络的

双阶段检测方法,额外增加了目标候选区域的提取过程的计算时间;YOLOv1 到 YOLOv3 采用单阶段的目标检测网络,降低模型的检测时间,但检测精度也随之降低,SSD 的检测精度和时间都有一定程度的提升,但仍低于改进的 YOLOv3 网络检测精度,本文算法相比于原 YOLOv3 在

交通标志识别准确率上又进一步提升,达到90.8%,并且检测时间也优于原网络,满足实时检测的任务要求。

表5 不同算法实验结果对比

方法	模型大小/MB	检测时间/ms	准确率/%
Faster-RCNN	124	112	84.6
YOLOv1	57.6	41	78.2
YOLOv2	42.2	34	81.3
YOLOv3	24.3	15	83.8
SSD	78.1	14	85.6
本文算法	15.9	11	90.8

4 结 论

针对道路不同类别的交通标志在遮挡,弱光,残缺的环境中的识别准确率的问题,本文提出改进的YOLOv3检测模型,得出如下结论:

引入深度可分离卷积优化网络训练和检测速度,减少模型训练和推理时间,将YOLOv3中主干网络的分支输出加入CBAM注意力机制模块,提高了在复杂环境中目标特征较弱的检测能力。

通过改进新的候选框交并比函数CIOU,增加了目标候选框尺度的损失函数,通过k-Means算法对候选框的大小进行聚类,提高了对目标定位的准确率。

实际道路上的交通标志种类繁多,本文目前仅对开源数据集进行部分扩充,后续需进一步扩充数据集种类和数量,增加模型泛化性。

参考文献

- [1] JIANG Y, ZHOU S, YAN J, et al. Traffic sign recognition using ridge regression and OTSU method[J]. IEEE, 2011,11(6):613-618.
- [2] MEUTER M, NUNN C, GORMER S M, et al. A decision fusion and reasoning module for a traffic sign recognition system [J]. IEEE Transactions on Intelligent Transportation Systems, 2011, 12(4): 1126-1134.
- [3] NASTARAN Y E, MANUEL M J, DAVID J, et al. Robust vehicle detection in different weather conditions: Using MIPM[J]. Plos One, 2018, 13(3): 191-204.
- [4] DALAL N. Histograms of oriented gradients for human detection [C]. Computer Vision and Pattern Recognition. IEEE Computer Society, 2005,177(7): 886-893.
- [5] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation [J]. IEEE Computer Society, 2014(2):580-587.
- [6] GIRSHICK R. Fast r-cnn [C]. Proceedings of the 2015 IEEE International Conference on Computer Vision (ICCV), 2015:1440-1448, DOI:10.1109/ICCV.2015.169.
- [7] 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.
- [8] HE K, GKIOXARI G, DOLLAR P, et al. Mask R-CNN [C]. International Conference on Computer Vision. IEEE Computer Society, 2017:2980-2988, DOI:10.1109/TPAMI.2018.2844175.
- [9] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection [C]. Computer Vision and Pattern Recognition. IEEE, 2016:779-788, DOI:10.1109/CVPR.2016.91.
- [10] REDMON J, FARHADI A. YOLO9000: Better, faster, stronger [C]. IEEE Conference on Computer Vision and Pattern Recognition. IEEE Computer Society, 2017,440(56):6517-6525.
- [11] REDMON J, FARHAD I A. YOLOv3: an incremental improvement [C]. IEEE Conference on Computer Vision and Pattern Recognition, 2018, DOI:10.48550/arXiv.1804.02767.
- [12] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot multibox detector [C]. European conference on computer vision. Springer, Cham, 2016,9905(3): 21-37.
- [13] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2017, 99: 2999-3007.
- [14] HASEGAWA R, IWAMOTO Y, CHEN Y W. Robust detection and recognition of Japanese traffic sign in the complex scenes based on deep learning [C]. 2019 IEEE 8th Global Conference on Consumer Electronics (GCCE), IEEE, 2019,230(4):2307-2321.
- [15] KONG S, PARK J, LEE S S, et al. Lightweight Traffic Sign Recognition Algorithm based on cascaded CNN [C]. 2019 19th International conference on Control, Automation and Systems (ICCAS). 2019,84: 506-509, DOI:10.23919/ICCAS47443.2019.8971735.
- [16] YANG B, YAN J, LEI Z, et al. CRAFT objects from images [C]. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2016, 205:6043-6051, DOI:10.1109/CVPR.2016.650.

- [17] BELL S, ZITNICK C L, BALA K, et al. Inside-outside net: Detecting objects in context with skip pooling and recurrent neural networks[C]. 2016 IEEE Conference on Computer Vision and Pattern Recognition(CVPR). IEEE, 2016,314(8):2874-2883.
- [18] CHEN X, GUPTA A. Spatial memory for context reasoning in object detection [C]. 2017 IEEE International Conference on Computer Vision(ICCV). IEEE, 2017,440(9):6964-6973.
- [19] YY A, HC A, CHI Z B, et al. SARPNET: Shape attention regional proposal network for liDAR-based 3D object detection [J]. Neurocomputing, 2020, 379(8):53-63.
- [20] LI A, QI J Q, LU H. Multi-attention guided feature fusion network for salient object detection [J]. Neurocomputing, 2020, 411(5):416-427.
- [21] ZHENG Z, WANG P, REN D, et al. Enhancing geometric factors in model learning and inference for object detection and instance segmentation[J]. IEEE, 2020,52(8):8574-8596.

作者简介

张建恒, 硕士, 助理实验师, 主要研究方向为智能汽车环境感知, 机器视觉等。

E-mail: zhangjh@ctbu.edu.cn