

改进 YOLOv5 的轻量化交通标志检测算法

贾子豪¹, 王文青¹, 刘光灿²

(1. 南京信息工程大学自动化学院, 南京 210044; 2. 东南大学自动化学院, 南京 210096)

摘要: 随着当今时代科技和人工智能的高速发展, 人们越来越倾向于无人驾驶这项技术。考虑到安全问题, 针对驾驶过程中交通标志的实时检测问题, 在 YOLOv5 模型的基础上做出改进, 提出了一种轻量化的交通标志检测算法。在模型的特征融合部分加入了注意力机制, 可以使模型更加突出目标特征。在检测层前加入一种轻量化的亚像素卷积层, 在不增加计算量的基础上, 有效地提高检测特征图的分辨率。对损失函数 CIoU (Complete intersection over union) 加以改进, 加快了网络的收敛速度, 并且收敛效果较改进前有了一定提升。实验结果表明, 本文模型准确率可达到 90.6%, 较基础网络提高了 14.5%, 检测速度可达到 70 帧/s, 基本满足对交通标志的实时精准检测。

关键词: 目标检测; 轻量化; 注意力机制; 亚像素卷积; 特征融合

中图分类号: TP391.4 **文献标志码:** A

Improved Lightweight Traffic Sign Detection Algorithm of YOLOv5

JIA Zihao¹, WANG Wenqing¹, LIU Guangcan²

(1. College of Automation, Nanjing University of Information Science and Technology, Nanjing 210044, China; 2. School of Automation, Southeast University, Nanjing 210096, China)

Abstract: With the rapid development of science and technology and artificial intelligence, people are more and more inclined to driverless technology. Considering the safety problem, aiming at the real-time detection of traffic signs during driving, the algorithm is improved on the basis of YOLOv5 model, and a lightweight traffic sign detection algorithm is proposed. The attention mechanism is added to the feature fusion part of the model, which can make the model more prominent target features. Then a lightweight sub-pixel convolution layer is added in front of the detection layer to effectively improve the resolution of the detection feature map without increasing the amount of computation. Finally, the loss function CIoU (Complete intersection over union) is improved, which speeds up the convergence speed of the network, and the convergence effect is better than that before the improvement. The experimental results show that the accuracy of this model reaches 90.6%, which is 14.5% higher than the basic network, and the detection speed reaches 70 frames / s, which basically meets the real-time accurate detection of traffic signs.

Key words: object detection; lightweight; attention mechanism; sub-pixel convolution; feature fusion

引言

随着视觉导航和人工智能技术的飞速发展,在无人驾驶领域中,对路况的分析以及对道路上交通标志的实时检测^[1-2]便成了计算机视觉的重点研究方向。由于在驾驶过程中道路状况非常复杂,细微的问题可能会引起巨大的后果,因此要求对交通标志快速且精准地检测。

传统的交通标志检测主要源于两大方面:(1)基于颜色特征的检测,交通标志的颜色主要有红色、黄色、蓝色3种,分别表示禁令、警示、指示3种指令。而基于颜色特征的检测方法,常常因恶劣天气等因素而难以检测。(2)基于形状特征的检测,交通标志的形状通常有三角形、方形、圆形3类,然而基于形状特征的检测方法,通常因为标志的损坏或角度问题,导致检测错误。

随着基于深度学习在目标检测领域的优异表现,很多学者开始将重点转向通过深度学习完成交通标志的检测和识别。当前深度学习的目标检测算法主要有两类,一类是Faster-RCNN^[3]、Mask-RCNN^[4]、Cascade-RCNN^[5]等二阶段网络,该类算法通常先产生候选区域,然后再对边界框进行微调。另一类是以YOLO^[6-7]和SSD^[8]为代表的单阶段网络,单阶段网络不会产生候选区域,而是直接对目标进行分类和定位。两类网络各有不足,传统的二阶段网络虽然检测准确率较单阶段网络高,但由于其计算量较大、检测速度慢,所以很难满足对交通标志的实时检测。

通过以上所介绍的两类网络,为满足交通标志检测中的实时性,本文采用YOLOv5(You only look once version 5)单阶段网络作为基准模型,同时为了减少计算量,提高检测效果,在基准模型内特征融合部分加入注意力机制突出图像中的重要特征。本文的主要贡献如下:

(1)将注意力机制CBAM替换网络特征融合(Neck)中的部分模块,不仅使模型更加关注重点目标的特征,提高检测精度,而且降低了模型参数量以及计算量。

(2)在网络中检测层(Prediction)之前加入轻量型三角式卷积层(Light triangular convolution, LTC),从而可以在检测前再一次提高特征图的分辨率,更有利于小目标的检测。

(3)在训练过程中对损失函数CIoU(Complete intersection over union)做出改进,加快网络训练时的收敛速度,并且提高了网络的收敛效果。

1 YOLOv5模型简介

根据网络深度宽度的不同,YOLOv5可由参数量从小到大被分为YOLOv5n、YOLOv5s、YOLOv5m、YOLOv5l和YOLOv5x五种版本。由于考虑到交通标志检测的实时性,本文选择模型参数量最小的YOLOv5n作为基础网络。

YOLOv5的网络结构图如图1所示,网络分为主干(Backbone)、特征融合(Neck)、检测(Prediction)3个部分,该网络流程首先将输入图像进入主干网络,提取图像特征,第二部分特征融合网络采用的是特征金字塔(Feature pyramid network, FPN)和路径聚合网络(Path aggregation network, PAN)结构组合,以便更好地将主干网络和特征融合网络的语义信息和位置信息进行交互融合,最终形成3种不同尺寸大小的输出特征图进行检测。

保留主干网络(Backbone)进行特征提取,为了实现模型的轻量化,将特征融合部分(Neck)中计算量庞大且冗余的部分C3块替换为注意力模块(CBAM),考虑到提升目标检测的准确性,在检测层(Prediction)之前加入了轻量型三角式卷积层(LTC),增强了小目标的特征信息。各模块功能在下文具体介绍。

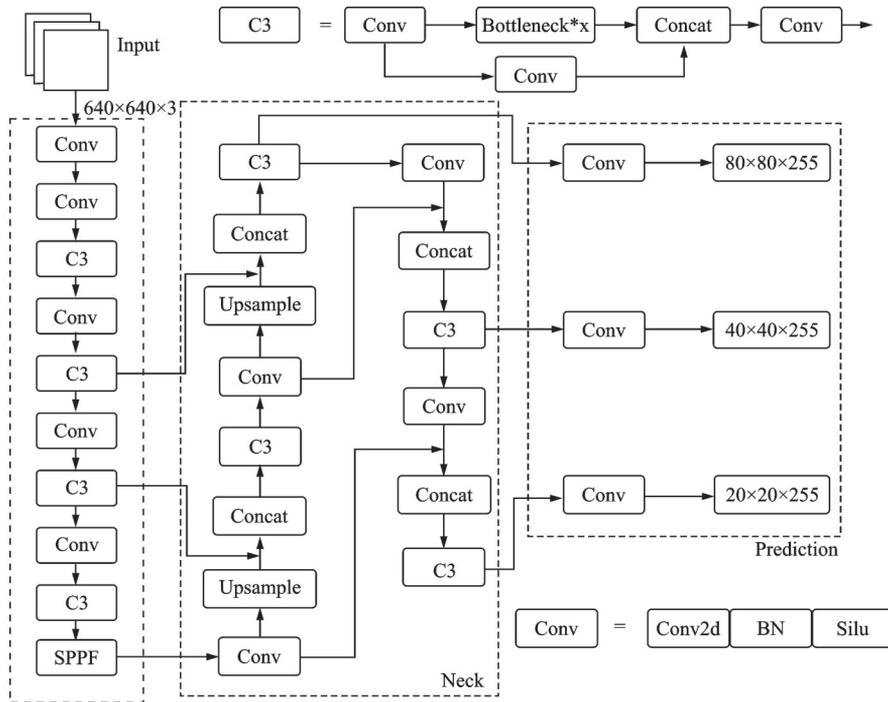


图1 YOLOv5网络结构

Fig.1 Network structure of YOLOv5

2 注意力机制

CBAM 模块^[9]由通道注意力模块(Channel attention module, CAM)和空间注意力模块(Spatial attention module, SAM)构成,这两个模块经由串行过程将权重信息重新分配,使得网络可以更有效率地关注到目标的重要特征信息。

通道注意力模块(CAM)将输入的特征图 F 分别进行基于宽度和高度的全局平均池化和最大池化,最终得到两个包含全局特征信息的特征矩阵($1 \times 1 \times C$),接着,再将特征矩阵分别送入一个两层的卷积神经网络^[10]。将输出的特征进行加和操作,再经过 Sigmoid 函数激活操作,最后生成通道注意力权重。

空间注意力模块(SAM)将 CAM 的输出特征图作为输入特征图,首先经过一个基于通道数的全局最大池化、全局平均池化的网络层,将特征图降维至 1 通道,得到特征图 ($H \times W \times 1$),然后将两个特征图做基于通道的拼接操作,再经过卷积层和 Sigmoid 函数激活操作,最终生成空间注意力权重。

本文采用通道注意力模块和空间注意力模块的串行组合,如图2所示,首先执行通道

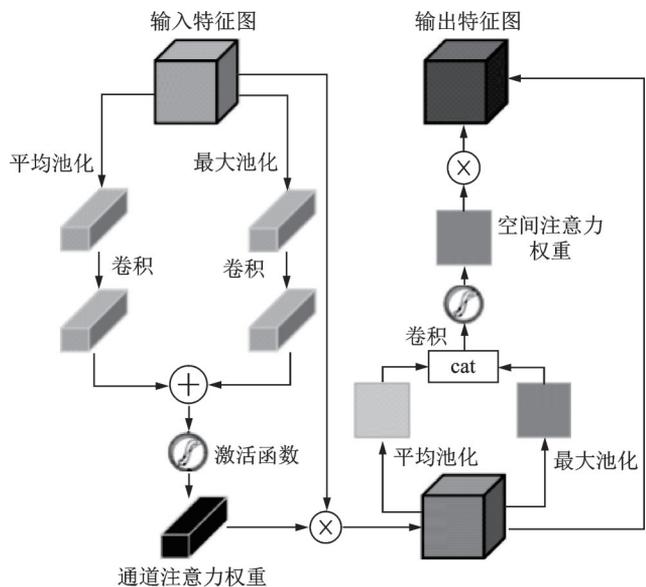


图2 CBAM 结构图

Fig.2 Structure diagram of CBAM

注意力模块,再执行空间注意力模块,可以在输出前更好地保留关注目标的特征信息。

CBAM注意力机制在深度学习中参数量和计算量都非常小,而特征融合网络中部分C3模块的计算量与参数量非常大,因此将CBAM注意力模块替换掉YOLOv5中特征融合网络(Neck)部分的C3模块,可以大幅度地减少参数量与计算量,同时使网络更加突出目标的特征信息。而且在特征融合网络(Neck)加入CBAM注意力机制,可以在融合过程中对目标通道和空间的特征信息建立更好的关联。改进CBAM后的特征融合模块图如图3所示。

3 轻量型三角式卷积层

YOLOv5的特征融合部分可以实现浅层位置信息和深层语义信息的相互补充,深层网络的特征图通常感受野相对较大,对应原图的区域也更大,对大目标的检测效果更好,然而对小目标的检测则容易出现漏检误检^[11],因此融合不同阶段网络的特征图,可以有效提高小目标的检测效果。

在YOLOv5基准网络中的特征融合部分,其特征融合实际上采用的是由一个自上而下的特征金字塔(FPN)^[12]与一个自下而上的路径聚合网络(PAN)^[13]组合结构。图4为特征金字塔(FPN)和路径聚合网络(PAN)的结构图,其中FPN是自顶向下,将高层的语义特征信息向下传递,对整个金字塔进行增强,然而只丰富了语义信息,PAN就是在FPN之后加入一个自底向上的金字塔,将底层的位置特征信息向上传递,这样既丰富了语义信息,同时也增强了位置信息。

考虑到YOLOv5会从特征融合网络(Neck)的PAN金字塔部分中选取3个不同尺度的特征图,进入Prediction检测层进行卷积和预测等操作。但同时PAN金字塔的结构为自底向上,会随着层数的上升,位置信息较语义信息逐渐减少。为提高检测时的定位准确性,在YOLOv5网络的检测层(Prediction)之前,也就是PAN的最后一层后加入一种轻量型三角式卷积层(Light triangular convolution, LTC)。

在该结构中,引入了亚像素卷积(sub-pixel convolution)代替传统的卷积层作为主要的卷积环节。亚像素(Sub-pixel)^[14]实际上就是对图像中相邻两像素之间进行细化,输入值通常为1/2或1/4。这意味着每个像素将被分为更小的单元从而对这些更小的单元实施插值算法^[15]。如果选择1/4,就相当于将每个像素在横向和纵向上插入3个值,也就是将一个像素放大为4个像素。通常经过亚像素插值之后,图像的像素值增加的同时,分辨率也有明显的提高。

同样亚像素卷积其实是一种传统卷积的简化形式,将一个低分辨率的图像,通过卷积和多通道之间的重组得到高分辨率的特征图。亚像素卷积过程图如图5所示,首先输入的低分辨率图像经过一系列隐藏层(Hidden layers),即经过一系列 $n \times n$ 卷积与池化操作。主要是为了将图像的特征进行提取,并且通过部分 1×1 卷积改变图像的通道数,在这里引入了一个参数为上采样倍率 r ,输入图像经过隐藏层操作通道数增加到 r^2 。然后亚像素卷积层的主要功能就是将这 r^2 个通道的特征图按像素顺序重新组合为新的 $W \times r, H \times r$ 的上采样图像结果。

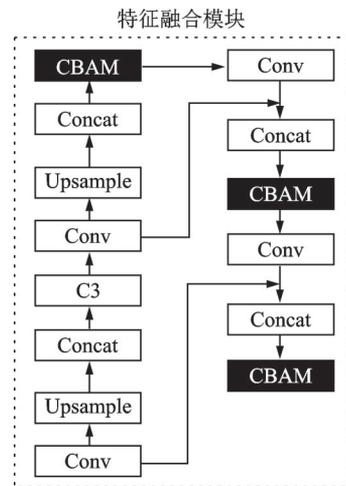


图3 改进CBAM后的特征融合模块
Fig.3 Feature fusion module after improving CBAM

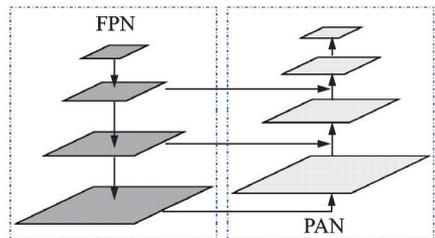


图4 FPN+PAN结构图
Fig.4 Structure diagram of FPN+PAN

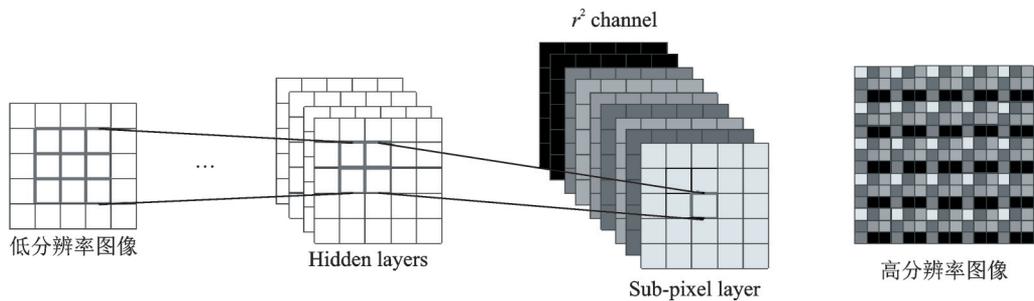


图5 亚像素卷积过程图

Fig.5 Sub-pixel convolution process diagram

轻量型三角式卷积层的结构图如图6所示,图中P1为特征融合网络最后一层的输出特征图,作为该卷积层的输入图,分别向两个方向作亚像素卷积和上采样,会生成P2大小的特征图,然后将P2大小的特征图与浅层网络中相同大小的特征图做特征融合,生成检测中等目标的特征图。然后将P1上采样得到的特征图做亚像素卷积得到的特征图和P2作亚像素卷积得到特征图做特征融合,最终生成3个不同尺寸大小的特征图(P1、P2、P3),且大小为等比增加(1:2:4),分别用来检测大、中、小目标。

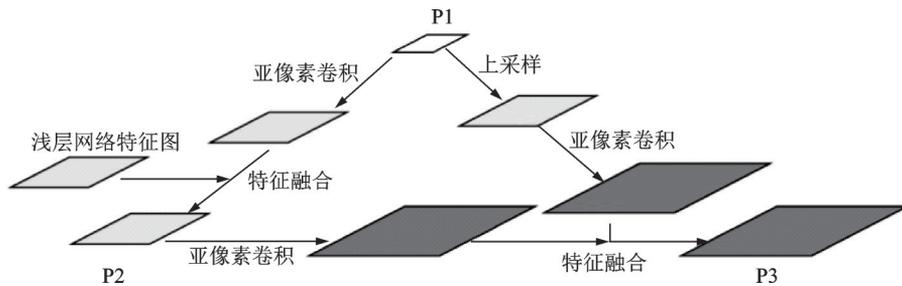


图6 轻量型三角式卷积层结构图

Fig.6 Structural diagram of light triangular convolution layer

由于在YOLOv5网络的基础上加以改进,不可避免地会增加一定的层数,所以为了保证不使参数量和计算量大幅增加,并且考虑到亚像素卷积层的计算量和参数量很少,该结构引入了亚像素卷积代替传统的卷积层作为主要的卷积环节。

4 改进的网络结构

在YOLOv5的主体结构上加入注意力机制(CBAM)并且在检测层之前加入轻量型三角式卷积层,改进后整体的网络结构如图7所示,网络主要分为主干网络、特征融合部分、轻量型三角式卷积层以及检测层。考虑到道路交通状况复杂,通常无人驾驶汽车会在较远距离来获取前方路况图像^[16],因此捕捉到的交通标志大多数是小目标^[17]。

传统的YOLOv5输入图像的尺寸设置为 640×640 ,经过原有网络操作后得到 80×80 、 40×40 、 20×20 三种尺寸大小的特征图,但是当检测特征图大小为 20×20 时,交通标志这类小目标的位置信息几乎难以检测。所以为了更有效地检测小目标,本文将模型的输入图像尺寸修改为 1280×1280 ,从而通过网络之后,得到 160×160 、 80×80 与 40×40 三种尺寸大小的输出特征图。其中 160×160 特征图的小目标位置信息比之前 80×80 特征图的位置信息要更加丰富,因此更有利于小目标的检测。 80×80 特

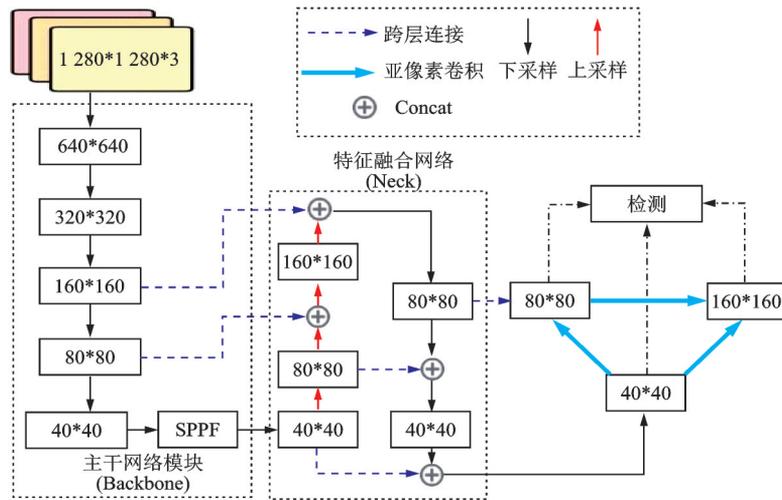


图7 改进的网络结构图

Fig.7 Improved network structure diagram

征图用于检测中等目标,而40×40特征图用于检测大目标。

为了使网络有效地检测较小的交通标志目标,首先通过主干网络(Backbone)提取图像特征,然后通过特征融合网络(Neck)丰富图像的语义信息与位置信息。为了使网络检测部分(Head)对目标更精准地检测,提高图像的分辨率是必要的,通过轻量型三角式卷积层(LTC)的亚像素卷积来提高特征图整体的像素值,从而增强图像的全局特征信息。网络进行LTC部分前后所生成不同通道的特征图对比如图8所示,在LTC部分前所生成的特征图已经可以良好地体现目标的位置信息,但仍存在目标信息模糊和不准确等问题。经过LTC生成的特征图,增强了整体的语义信息,并且更加突出了目标的特征。

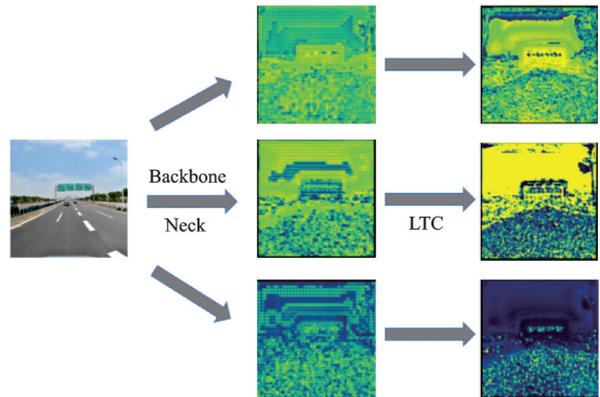


图8 特征图可视结果对比

Fig.8 Comparison of visual results of characteristic maps

5 损失函数改进

损失函数是用来衡量模型预测值和真实值之间的误差程度,极大程度上体现了模型的稳定性能。在YOLOv5中,损失函数由3部分组成,分别为分类损失(cls_loss),用来计算锚框与对应标定的类别是否一致,第二是定位损失(box_loss),用来计算预测框与真实框的位置误差,最后一个置信度损失(obj_loss),可以计算网络对目标预测的置信度。本文对损失函数CIoU Loss进行改进,实现了在交通标志检测中对目标位置信息更加准确的计算。

交并比(Intersection over union, IoU)函数是计算两个边界框交集和并集之比,用来衡量预测框的准确度

$$IoU = \frac{|B \cap B_{gt}|}{|B \cup B_{gt}|} \quad (1)$$

式中: B 、 B_{gt} 分别表示预测框和真实框, 通常 IoU 的值越高, 则网络的预测框越准确。但采用传统的 IoU 作为损失函数, 两个物体若不重叠, IoU 的值和梯度都几乎为 0, 不能反映两者之间距离大小, 因此函数无法进行优化。为此 YOLOv5 提出了 3 种改进损失函数的方法, 分别为 GIoU、DIoU 与 CIoU。

CIoU 是在 GIoU 与 DIoU 的基础上做出改进, 收敛速度和拟合的效果都优于前者, 增加了检测框大小的损失以及长和宽的损失, 可以使预测框更加与真实框重合, 收敛速度和拟合效果都优于前者。

$$CIoU = IoU - \left(\frac{\rho^2(b, b_{gt})}{d^2} + \alpha\nu \right) \quad (2)$$

在函数中加入了 $\alpha\nu$, 其中 α 为权重系数, 而 ν 用来衡量长宽比的一致性, 定义式为

$$\nu = \frac{4}{\pi^2} \left(\arctan \frac{w_{gt}}{h_{gt}} - \arctan \frac{w}{h} \right)^2 \quad (3)$$

在 CIoU 的基础上加以改进, 将式(2)分式上的欧氏距离进行平方根运算。在交通标志识别过程中, 会有大小不同的交通标志出现在图像中, 当目标越大时, 检测框和真实框的中心点距离也会越大, 从而在计算损失时, 会明显高于小目标的损失, 不利于小目标的损失计算。因此, 通过开平方根的方法, 可以减少大目标的损失权重, 更有效地计算损失, 并且通过实验分析发现, 损失函数改进前后算法的计算量基本没有增加, 从而并未增加模型的计算成本。加入开平方根的改进后可以使损失值比改进前的更低, 而且使损失值的收敛速度加快, 提高了损失函数的收敛效果。改进后的损失函数为

$$Loss = IoU - \left(\frac{\sqrt{\rho^2(b, b_{gt})}}{d^2} + \alpha\nu \right) \quad (4)$$

改进前后 Loss 收敛的趋势图如图 9 所示, 分别将改进前后的 box_loss、obj_loss、cls_loss 的收敛效果加以对比, 可以发现改进后的模型收敛速度和收敛效果都要优于初始的 CIoU 函数, 其中置信度损失 (obj_loss) 表现最为优异, 最终损失收敛值接近于 0.002, 小于 CIoU 的收敛值 0.004, 表示对目标检测的置信度也有一定提升。

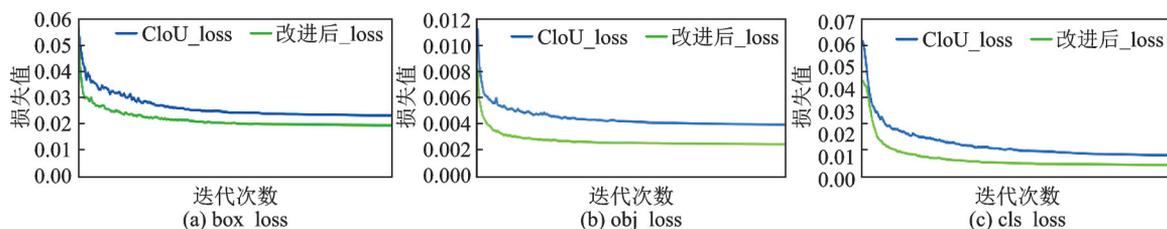


图9 收敛效果改进前后对比

Fig.9 Comparison of convergence effect before and after improvement

6 实验与分析

6.1 数据集介绍与实验环境

选择的数据集是 TT100K (中国交通标志数据集^[18]), 由清华大学和腾讯联合实验室所整理并公开。在中国 5 个不同的城市选择了多个区域, 从腾讯数据中心下载了 10 万张全景图, 每张图片被分割为 2048×2048 的大小。

其中含有交通标志的图片有 1 万多张, 经过对数据集的清洗和对图像损坏清理等处理, 最终本文采

用的数据集包含有 9 170 张图片,按 9:1 的比例划分训练集和测试集。训练集包含 8 253 张图片,而测试集有 917 张图片。

官方提供的 TT100K 数据集中共包含 151 个交通标志类别,其中大多数交通标志在数据集中的实例数量很少,并且接近一般的类别仅有个位数的实例,所以训练的效果不具有代表性。最终本文选择了 19 个在全国范围内常见且具有代表性的交通标志类别。分别是禁止驶入、常规提示标志、解除禁止速度、常规禁止标志、限速、限高、限重、限宽、限制轴重、禁止停车、常规警告标志、减速让行、最低限速、停车让行、人行横道、禁止通行以及其他禁止、提示、警告等标志,数据集中常见的交通标志类别如图 10 所示。



图 10 常见交通标志类别

Fig.10 Common traffic sign categories

本实验训练过程在 Windows 10、CUDA 10.2 环境下进行, GPU 配置为: NVIDIA RTX 2080ti, 以 Pytorch 为主要框架, 以 Python 为主要编程语言建立模型。

所有实验的训练迭代次数均设置为 300 轮, 每批次处理尺寸设置为 16, 学习率衰减策略为余弦退火方式, 并且初始学习率设置为 0.01, 周期学习率为 0.1。

6.2 评测指标

深度学习目标检测的重要评估指标主要有平均准确率 (Mean average precision, mAP)、准确率 (Precision, P)、召回率 (Recall, R), 它们的值越高表示模型的性能越好。每秒处理帧率 (FPS) 表示模型运行的速度, 其值越高表示模型检测速度越快。其中准确率 P 为预测为正的样本中实际为正的比例, 召回率 R 为正样本中被正确预测的比例。上述指标的计算公式分别为

$$P = \frac{TP}{TP + FP} \tag{5}$$

$$R = \frac{TP}{TP + FN} \tag{6}$$

$$AP = \int_0^1 P(R) dn \tag{7}$$

$$mAP = \frac{\sum_{i=1}^n AP_i}{n} \tag{8}$$

式中: TP (True positive) 为网络检测为真的正样本数量; FN (False negative) 为网络检测为非真的正样本数量; FP (False positive) 为模型检测为真的负样本数量; mAP 为所有类别的平均准确率。

6.3 实验结果分析

根据 YOLOv5 的基本框架做出了基于深度和宽度上一定比例的压缩和延伸, 深度表示网络中某一层模型的数量, 而宽度则表示某一层模型对通道数的改变。基准网络 YOLOv5n 的深度宽度比为 0.33:0.50。将深度宽度比分别作了两种改变, 一种是将深度延伸, 宽度压缩, 成为 0.67:0.25; 另一种是保持深度不变, 将宽度压缩, 成为 0.33:0.25。随着网络深度的压缩变化, 模型参数的计算量也随之减

少,可以使模型更加轻量化。

同时分别以这两种比例的模型加入CBAM注意力机制和轻量化三角式卷积层(LTC),由于在加入CBAM注意力机制,同时替换掉了特征融合网络的部分C3层,所以大幅地减少了计算量和参数量。

消融实验结果对比如表1所示,实验以YOLOv5n为基础模型,分别以深度宽度比为0.33:0.25和0.67:0.25两种框架逐次加入CBAM、LTC以及改进loss的实验,进行结果对比。

表1 消融实验结果对比
Table 1 Comparison of ablation experiment results

算法	0.33:0.25	0.67:0.25	CBAM	LTC	改进 loss	P/%	R/%	mAP	参数量/ MB	计算量/ GB
YOLOv5s						90.9	74.3	79.9	7.07	16.0
YOLOv5n						84.1	69.5	76.1	1.85	4.3
A	✓		✓			85.2	75.3	78.7	1.45	3.7
B	✓		✓	✓		88.1	75.7	81.4	1.49	4.0
C	✓		✓	✓	✓	86.5	77.2	82.2	1.49	4.0
D		✓	✓			91.5	81.8	86.6	1.80	4.7
E		✓	✓	✓		88.2	82.3	89.1	1.83	5.2
F		✓	✓	✓	✓	87.1	84.3	90.6	1.83	5.2

首先以0.33:0.25框架作消融实验,由于将基准模型在宽度上进行压缩以及对特征融合部分中的部分C3层替换为CBAM模块,所以计算量和参数量相对减少。加入LTC模块后使得网络的准确率P提升了2.9%,mAP也有显著提升。最后通过对loss进行改进,在参数量和计算量并未增加的情况下,mAP也得到一定的提升。

然后以0.67:0.25框架作消融实验,由于将基准模型在深度上作了延伸,所以网络中的模块数也有一定增加,因此参数量与计算量也随之增加。首先加入CBAM注意力模块,网络的准确率P可以达到91.5%,然后再依次加入LTC以及loss的改进,召回率R和mAP都在逐步增加。

通过实验结果表明,在0.67:0.25框架下,同时加入CBAM、LTC以及loss改进的网络mAP可以达到90.6%,较基准模型提升了14.5%。并且参数量为1.83 MB,较基准模型YOLOv5n减少了0.02 MB。

改进前后的检测效果图如图11所示,可以看出,初始的YOLOv5n对小目标检测时会存在漏检的问题,同时存在目标角度倾斜以及物体遮挡目标等问题导致误检或置信度较低。相比之下,本文算法则是在小目标可以检测到的基础上,对目标角度倾斜以及遮挡情况的检测保持了稳定的准确率。

YOLO系列的结果对比如表2所示,可以看出,在精度上YOLOv5l最高,可以达到91.6%,但由于其参数量过于庞大,导致其检测速度很慢,检测速度只有16.1帧/s,难以达到实时检测的效果,在速度方面YOLOv5n因其参数量较小,检测速度可以达到90.7帧/s,但精度方面仍有不足,只有76.1%。综合考虑,在YOLO系列中,本文算法无论在精度、速度和模型大小上都可以达到轻量化且精确的程度。

本文算法与其他框架的最新算法加以对比,如表3所示,加入了Faster-RCNN与FA-SSD算法的实验结果,他们模型的体积与参数量都较YOLOv5n大很多,并且难以实现轻量化,不便于嵌入式的装载。而本文的算法在精度和速度上都拥有着更良好的表现。

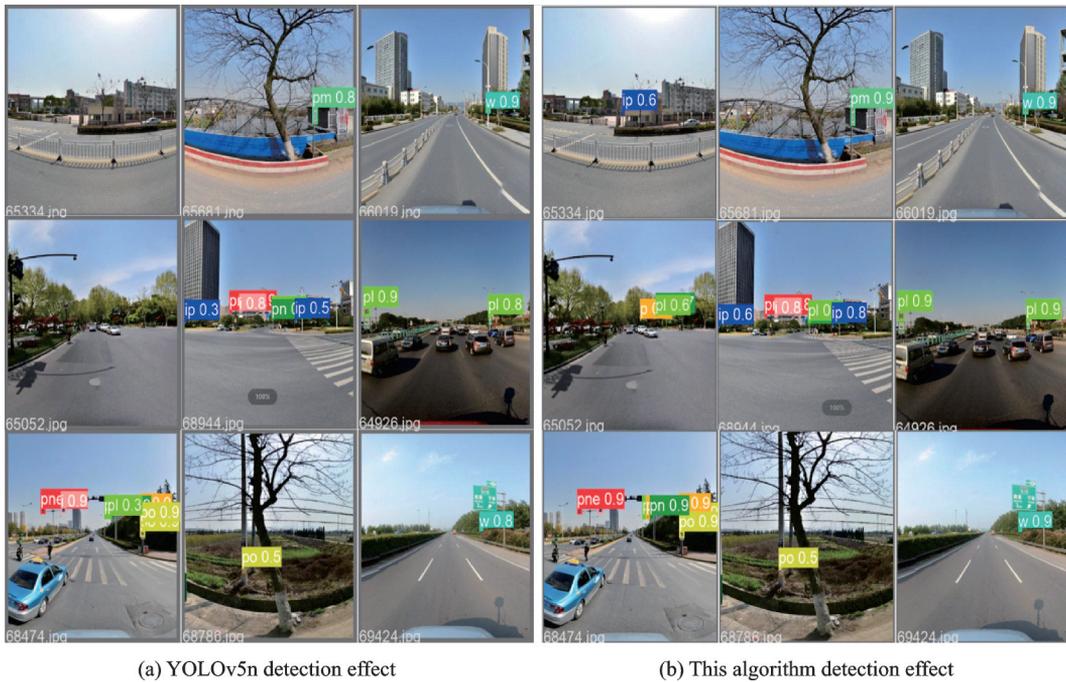


图 11 改进前后检测效果对比

Fig.11 Comparison of detection results before and after improvement

表 2 YOLO 系列算法对比

Table 2 YOLO algorithm comparison

算法	P/%	R/%	mAP/%	检测速度/(帧·s ⁻¹)	参数量/MB	计算量/GB
YOLOv5n	84.1	69.5	76.1	90.7	1.85	4.3
YOLOv5s	90.9	74.3	79.9	53.9	7.07	16.0
YOLOv5m	87.3	78.5	85.7	30.6	22.5	45.3
YOLOv5l	88.4	85.4	91.6	16.1	46.2	108.2
本文	87.1	84.3	90.6	70.5	1.83	5.2

7 结束语

针对交通标志检测的实时性问题展开了研究,提出了一种轻量化的交通标志检测算法。首先在特征融合部分加入了CBAM注意力机制,替换掉基准网络中部分复杂结构,减少了计算量,并且使模型更加关注重点目标,然后通过检测层之前的轻量化三角式卷积层来提升检测特征图的分辨率,最后对损失函数加以改进,提升了收敛速度,并且更加有利于对小目标位置进行定位。结合实验结果表明,本文算法与众多主流目标检测算法对比中可以同时表现出更高的准确率、更快的检测速度以及更少的参数量。

表 3 算法对比

Table 3 Algorithm comparison

算法	mAP/%	检测速度/(帧·s ⁻¹)	参数量/MB	计算量/GB
YOLOv5s	79.9	53.9	7.07	16.0
YOLOv5n	76.1	90.7	1.85	4.3
Faster-RCNN	62.6	11.8	52.6	91.1
Cascade-RCNN	64.1	14.7	68.96	118.86
FCOS	66.2	21.7	31.88	78.98
GFL	71.2	18.8	32.07	82.12
FA-SSD	80.2	13.6	4.5	8.91
本文	90.6	70.5	1.83	5.2

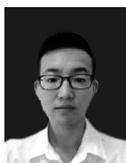
参考文献:

- [1] CHENG P, LIU W, ZHANG Y, et al. LOCO: Local context based faster R-CNN for small traffic sign detection[C]// Proceedings of International Conference on the Multimedia Modeling.[S.l.]: Springer, 2018: 329-341.
- [2] 韩松臣, 张比浩, 李伟, 等. 基于改进 Faster-RCNN 的机场场面小目标物体检测算法[J]. 南京航空航天大学学报, 2019, 51(6): 735-741.
HAN Songchen, ZHANG Bihao, LI Wei, et al. Small target detection in airport scene via modified Faster-RCNN[J]. Journal of Nanjing University of Aeronautics & Astronautics, 2019, 51(6): 735-741.
- [3] REN Shaoqing, HE Kaiming, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks[C]//Proceedings of Proceedings of Advances in Neural Information Processing Systems. Montreal: IEEE, 2015: 497-507.
- [4] HE K, GKIOXARI G, DOLLAR P, et al. Mask R-CNN[C]//Proceedings of the IEEE International Conference on Computer Vision. [S.l.]: IEEE, 2017: 2961-2969.
- [5] CAI Z, VASCONCELOS N. Cascade R-CNN: Delving into high quality object detection[C]//Proceedings of Computer Vision and Pattern Recognition. [S.l.]: [s.n.], 2018: 6154-6162.
- [6] REDMON J, FARHADI A. YOLO9000: Better, faster, stronger[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.[S.l.]: IEEE, 2017: 7263-7271.
- [7] REDMON J, FARHADI A. Yolov3: An incremental improvement[J]. arXiv Preprint arXiv, 2018: 1804.02767.
- [8] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot multi-box detector[C]//Proceedings of European Conference on Computer Vision. [S.l.]: [s.n.], 2016: 21-37.
- [9] WOO S, PARK J C. CBAM: Convolutional block attention module[J]. arXiv Preprint arXiv, 2018: 1807.06521.
- [10] GIRSHICK R, DONAHUE J. Rich feature hierarchies for accurate object detection and semantic segmentation[J]. arXiv Preprint arXiv, 2013: 1311.2524.
- [11] 唐磊. 基于双向特征融合的交通标志识别[J]. 汽车实用技术, 2022, 47(9): 1-3.
TANG Lei. Traffic sign recognition based on bidirectional feature fusion[J]. Practical Automotive Technology, 2022, 47(9): 1-3.
- [12] LIN T Y, DOLLAR P, ROSS G, et al. Feature pyramid networks for object detection[J]. arXiv Preprint arXiv, 2016: 1621.03144.
- [13] WANG Wenhai, XIE Enze, SONG Xiaoge, et al. PAN: Efficient and accurate arbitrary-shaped text detection with pixel aggregation network[J]. arXiv Preprint arXiv, 2018: 1805.10180.
- [14] SUN Y, CHEN J, LIU Q, et al. Learning image compressed sensing with sub-pixel convolutional generative adversarial network[J]. Pattern Recognition, 2019, 98(12): 107051.
- [15] JIANG Y, CHEN H. High-frequency feature learning in image super-resolution with sub-pixel convolutional neural network [C]// Proceedings of 2019 the 3rd International Conference on Video and Image Processing.[S.l.]: ICVIP, 2019.
- [16] 郭璠, 张泳祥, 唐璠, 等. YOLOv3-A: 基于注意力机制的交通标志检测网络[J]. 通信学报, 2021, 42(1): 87-99.
GUO Pan, ZHANG Yongxiang, TANG Jin, et al. YOLOv3-A: Traffic sign detection network based on attention mechanism [J]. Journal of Communications, 2021, 42(1): 87-99.
- [17] DONGTAO Z, JIE C, XING Y, et al. Traffic sign detection method of improved SSD based on deep learning[C]//Proceedings of the IEEE 4th International Conference on Computer and Communications.[S.l.]: IEEE, 2018: 1516-1520.
- [18] 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(学术版), 2022, 18(5): 98-104.
WANG Zehua, SONG Weihu, WU Jianhua. Lightweight traffic sign detection model based on improved YOLOv4 network[J]. Computer Knowledge and Technology: Academic Edition, 2022, 18(5): 98-104.

作者简介:



贾子豪(1999-), 通信作者, 男, 硕士研究生, 研究方向: 目标检测、模式识别, E-mail: 1299339316@qq.com。



王文青(1995-), 男, 硕士研究生, 研究方向: 目标检测、模式识别。



刘光灿(1982-), 男, 博士, 教授, 研究方向: 机器学习、计算机视觉与图像处理。

(编辑: 夏道家)