DOI:10.19652/j.cnki.femt.2305825

基于 FMS-YOLOv5s 的轻量化交通标志识别算法*

曹 立 康少波
(厦门理工学院电气工程与自动化学院 厦门 361024)

摘 要:针对目前的道路交通标志模型有着检测速度慢、模型大和参数多的缺点,提出了一种基于 YOLOv5s 算法的轻量化交通标志识别算法。首先引入轻量化 FasterNet 网络,利用该网络中的 FasterNet Block 结构与原主干网络的 C3 融合,形成一种全新的 C3Faster 结构;接着将原网络的损失函数修改为基于最小点距离(MPDIoU)的损失函数,来提高边界框回归的准确性和效率;最后结合高效且轻量的置换注意力机制(shuffle attention,SA),提高模型的泛化能力和稳定性。在 CCTSDB 2021 数据集上的实验结果表明,与原网络相比,改进后模型的参数量、模型大小、GFLOPs 分别减少了 17.5%、17.5%和 20%;同时 mAP@0.5、mAP@0.75、mAP@0.5:0.95 分别提升了 2.3%、3.4%和 2.4%。而且与 YOLOv3-tiny 等其他算法对比,所提出的算法有明显的优越性,能满足各种场景下移动端实时性的需求。

关键词:YOLOv5s;交通标志识别;轻量化;FasterNet;MPDIoU
中图分类号:TP3 **文献标识码**:A **国家标准学科分类代码**:520.604

Lightweight traffic sign recognition algorithm based on FMS-YOLOv5s

Cao Li Kang Shaobo
(Electrical Engineering & Automation, Xiamen University of Technology, Xiamen 361024, China)

Abstract: A lightweight traffic sign recognition algorithm based on YOLOv5s algorithm is proposed for the current road traffic sign model with the disadvantages of slow detection speed, large model and many parameters. Firstly, a lightweight FasterNet network is introduced, and the FasterNet Block structure in the network is fused with the C3 of the original backbone network to form a new C3Faster structure. Then the loss function of the original network is modified to MPDIoU to improve the accuracy and efficiency of the bounding box regression. Finally, the efficient and lightweight SA attention mechanism is combined to improve the generalization ability and stability of the model. The experimental results on the CCTSDB 2021 dataset show that compared with the original network, the number of parameters, model size, and GFLOPs of the improved model have been reduced by 17.5%, 17.5%, and 20%, respectively. Meanwhile, mAP@0.5, mAP@0.75, and mAP@0.5:0.95 have been improved by 2.3%, 3.4%, and 2.4%, respectively. And comparing with other algorithms such as YOLOv3-tiny, the proposed algorithm has obvious superiority and can meet the real-time demand of mobile in various scenarios.

Keywords: YOLOv5s; traffic sign recognition; lightweight; FasterNet; MPDIoU

0 引 言

目前网络技术和人工智能发展迅速,自动驾驶已经逐渐进入了人们的视野,在交通道路复杂且城市化集中的今天,交通标志成为了现代城市智能交通系统的重要组成部分,这些标志不但被用来传递一些引导、限制和警告等重要信息,还无时不刻保障着人们的出行安全和车辆安全,降低安全事故的发生。交通标志识别如今被广泛地应用于无人驾驶技术中,由于现实场景中交通标志的图形各种各样,同时常常会因为行人遮挡、光线、天气等因素受到限制,所以对于交通标志的识别精度以及速度成为了一项重要的研究方向。

收稿日期:2023-12-06
* **基金项目**:厦门市海洋与渔业发展专项资金青年科技创新项目(23ZHZB043QCB37)资助

目前的交通标志识别方法主要分为传统交通标志识别和基于卷积神经网络(CNN)的交通标志识别这两大类。传统的检测方法主要是通过颜色、纹理特征等来区分所识别的目标。例如,Guidigar 等[1]提出高阶光谱结合基于纹理特征的方法,利用线型辨别分析框架提高辨别能力,得到了较高的准确率。Ellahyani 等[2]利用方向梯度直方图(HOG)拓展到色调、饱和度和亮度(HSI)颜色空间,结合局部自相似特征,采用随机森林和支持向量机(SVM)获得了较好的结果。Liu 等[3]提出了一种使用群稀疏编码的特征提取方法,基于对数极坐标映射的编码方式识别交通标志,结果表明此编码方式优于当时的编码方式。而在卷积神经网络出现后,使得检测的速度和效果得到了巨大的提升。其中,深度学习被研究人员广泛使用,以 R-CNN(Regions with CNN features) 两阶段目标检测算法为例,该算法主要是利用卷积神经网络对目标进行分类来取得不错的结果。Liu 等[4]提出了基于多尺度的卷积神经网络(MR-CNN),在分类过程中利用多尺度联系上下文信息来加强特征融合,此方法在小目标的检测效果有较好的优势。而另外的一阶段目标检测算法以 SSD(single shot multiBox detector)和 YOLO(you only look once)系列为主,该系列算法的优点是检测速度和性能优秀,韩岩江等[5]提出了一种融合坐标注意力机制和双向加权特征金字塔(BIFPN)的交通标志检测方法,提高模型的抗光性和改善对小目标的检测能力。高涛等[6]采用 ResNet 残差网络构建骨干,为了提取更多信息,增加网络的浅层数,引入 K-means 聚类算法和 GIoU 损失函数提升精度,实验结果对比原网络有提升,且抗干扰能力强。霍爱清等[7]提出了一种改进 YOLOX 的弱光线道路交通标志检测方法,引入轻量级 Mobile ViT Block 模块,添加自适应特征融合金字塔和 Focal 损失函数,有效避免因弱光线导致的漏检和错检现象。俞林森[8]等设计了一种基于鬼影模块的轻量级骨干网络,引入前景注意力感知模块来抑制背景噪声,提升了模型的性能。

由于卷积神经网络存在复杂度过高、模型大小过大和实时性较差等问题,基于深度学习的交通标志检测算法在精度和速度上难以得到较好的平衡,不利于移动端的部署。而 YOLO 系列以速度快、效率高等特点在很多领域都体现了优秀的检测性能[9-11],所以本文选取 YOLO 系列作为研究对象,其中目前较为轻量的 YOLOv5 成为了首选,虽然 YOLOv5 网络有着较好的检测性能,但依然存在泛化能力较差、检测精度低等问题,因此在 YOLOv5s 的基础上,本文提出了一种 FMS-YOLOv5 模型,该模型在精度和轻量化之间取得了较好的平衡。

本文将 FasterNet[12]网络中的 FasterNet Block 模块融入 YOLOv5s 网络的 C3 部分,形成一种全新的 C3Faster 结构,来减少网络的参数量和计算量。将 CIoU 损失函数替换成 MPDIoU[13]损失函数,加快网络的收敛速度,提高模型的检测性能。将置换注意力机制(shuffle attention,SA)[14]模块嵌入到网络的颈部部分,在不增加参数量的前提下,提高模型的泛化能力。

1 YOLOv5 网络结构分析

在 YOLO 系列算法中,YOLOv5 相比前几代有着模型更小,更加灵活,选择性多等优势。YOLOv5 根据网络的深度(depth_multiple)和宽度(width_multiple)参数大小分为 YOLOv5s、YOLOv5m、YOLOv5l 和 YOLOv5x 4 个版本。考虑到轻便性,本文网络以 YOLOv5s 为基础,该网络主要由输入端(input)、主干网络(backbone)、颈部网络(neck)、头部输出(head)和输出端(output)组成。

YOLOv5 输入端与 YOLOv4[15]一样都采用了马赛克数据增强,即选取 4 张图片来随机缩放、随机裁剪和随机排布的方式进行拼接成一张新的图片,能够加强对小目标的特征提取。但是 YOLOv5 中新加入了自适应锚框计算和自适应图片缩放,在训练开始后,网络会在初始锚框的基础上输出预测框,与真实框进行对比计算差距,进行反向更新和迭代网络参数,提高网络准确性。自适应图片缩放是针对不同尺寸大小的图片,统一缩放到一种尺寸,如 608×608 大小,而原始比例不同则会在缩放后添加少量的黑边,不仅能够减少计算量,还能加快推理的速度。

YOLOv5 主干网络采用 CSP 结构,与 YOLOv4 不同的是,YOLOv5 中用了两种 CSP 结构,即主干网络为 CSP1_X,颈部网络为 CSP2_X。再是引入了 Focus 结构,该结构对图片进行切片操作,主要把高分辨率的图拆解为多个低分辨率的图,示意图如图 1 所示。例如输入的图片为 608×608×3,经过切片操作后,变成 304×304×12 的特征图,再经过 32 个卷积核的卷积后变成 304×304×32 的特征图。

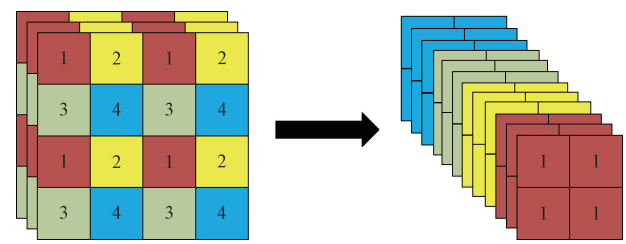


图 1 Focus 切片操作

Fig. 1 Focus slicing operation

YOLOv5 颈部网络与 YOLOv4 同样采用特征金字塔网络(FPN)和金字塔注意力网络(PAN)的结构,FPN 层自顶向下传递较强的语义特征,PAN 层自底向上传递较强的定位特征,二者联合加强对特征的融合提取。再者加入了空间金字塔池化结构(spatial pyramid pooling,SPP)模块,该模块置于主干网络后,能够更加有效的分离上下文特征和加强特征提取范围。而输出端采用 CIoU 函数作为目标定位框的损失函数,后处理过程中采用非极大抑制(non-maximum suppression,NMS)的方式筛选出最好

的检测结果。

2 模型改进

由于汽车在交通道路复杂的路口行驶容易受到光线、天气、行人等因素影响，对实时性的要求较高，所以对模型进行轻量化处理，对模型的 C3 卷积部分与 FasterNet Block 模块融合，使得能够搭载在小型设备上。同时为了减小误检和漏检的情况，在损失函数部分进行改进，用 MPDIoU 替换原网络的损失函数。最后在不对模型大小和参数有增加的条件下，引入轻量且高效的 SA 注意力机制，改进后的模型 FMS-YOLOv5 结构如图 2 所示。

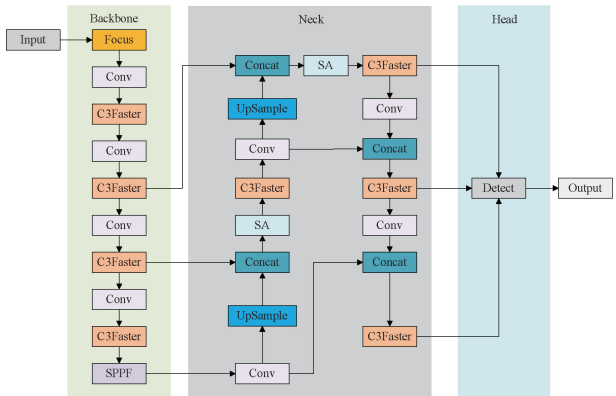


图 2 FMS-YOLOv5 结构

Fig. 2 Block diagram of FMS-YOLOv5 structure

2.1 FasterNet 网络融合的 C3Faster

在实际道路场景中对交通标志检测时，对便携性和实时性有一定的要求，要求模型尽可能的轻便和准确。目前卷积神经网络中较受欢迎的轻量化网络有 MobileNetv3[16]、ShuffleNet[17]、EfficientNet[18] 和 GhostNet[19] 等，这些都是利用深度卷积(DWConv)或组卷积(GConv)来提取空间特征。然而，这些卷积都致力于减少每秒浮点运算次数 (floating point operations，FLOPs)，这种减少不一定会使延时时间也相应的减少，尤其是在深度卷积中，模型的计算效率就不太理想，原因是该卷积的算子会频繁的对内存进行访问。延时时间和 FLOPs 之间的关系为：

$$Latency = \frac{FLOPs}{FLOPS} \tag{1}$$

式中：FLOPS 是每秒浮点运算次数，用来衡量硬件的性能；FLOPs 是浮点运算次数，可以用来衡量算法和模型的复杂度。

为了实现更快的网络，采用了一种全新的部分卷积 (partial convolution，PConv)，它利用了特征图中的冗余，只在一部分的输入通道上应用常规卷积，对其余的通道并不会产生影响，减少了计算冗余以及内存访问的数量，可以更有效地提取空间特征，PConv 与其他卷积对比如图 3 所示。在本质上，PConv 的 FLOPs 低于常规卷积，而 FLOPS 高于深度卷积或者组卷积。

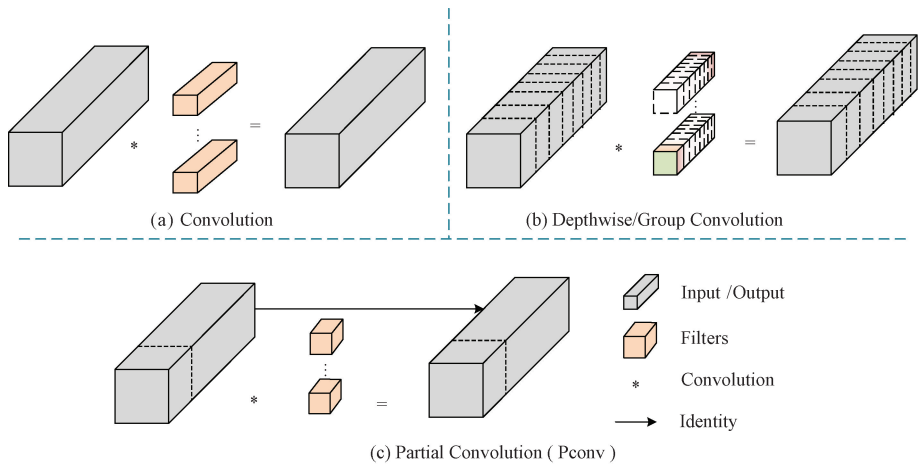


图 3 卷积对比

Fig. 3 Convolution comparison chart

基于 PConv 进一步搭建出 FasterNet 框架，它在分类、检测和分割任务方面实现了较为先进的性能，同时具有更低的延迟和更高的吞吐量。网络结构图如图 4 所示。该网络结构主要有 4 个 Stage，stage1 和 stage4 之前都有着一个嵌入层(步长为 4 的常规卷积 4×4)，stage2 和 stage3 之前都有着一个合并层(步长为 2 的常规卷积 2×2)，主要用于扩展空间下的采样和信道数，且每个 stage 都由一个 FasterNet block 组成。在 FasterNet block 中都有一个 PConv 层，后与两个 PWConv 或是 1×1 的常规卷积层相连，这 2 层一起显示为倒置残差块，将归一层和激活层嵌入这 2 层之间，以保持特征的多样性并实现较低的延迟。

本文针对轻量化的改进方法主要是利用 FasterNet block 的结构层对原网络的 C3 卷积层进行融合，形成一种全新的 C3Faster 卷积层，以此来减少模型的参数量。

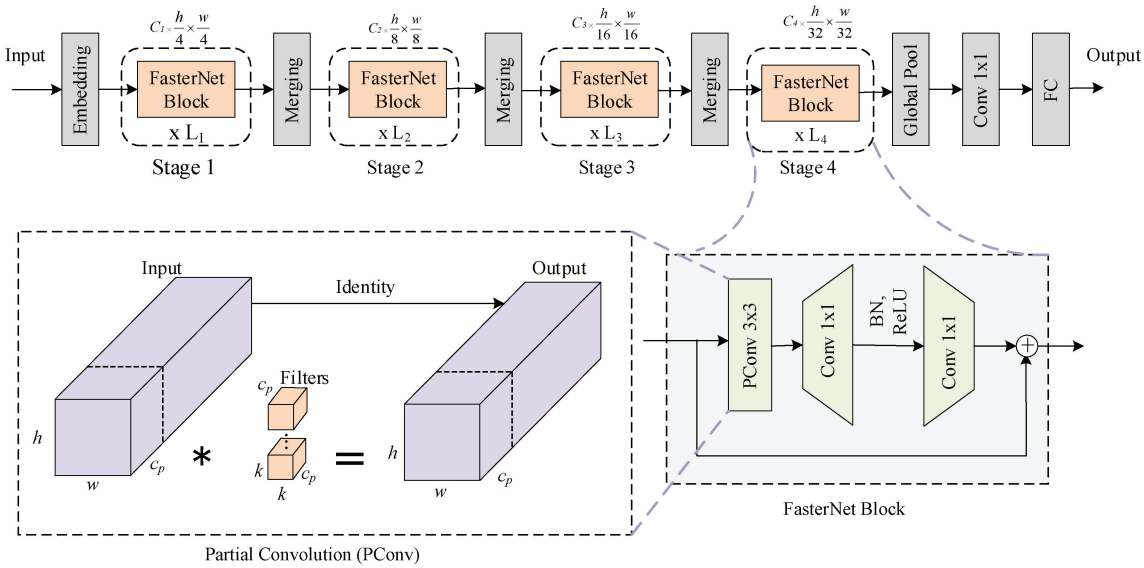


图 4 FasterNET 网络结构

Fig. 4 FasterNET network structure diagram

2.2 MPDIoU 损失函数

在目标检测和实例分割中，边界框回归（bounding box rgression，BBR）在其中起着目标定位的作用，大多数现有的 BBR 损失函数可以分为两类，基于 n 范数的损失函数和基于 IoU 的损失函数。然而当预测框和边界框具有同样大小的纵横比时，但宽度和高度完全不一致时，大多数现在的 BBR 损失函数都无法优化。现有的损失函数计算包括 GIoU[20]、DIoU[21]、CIoU 和 EIoU[22] 等，最初的 IoU 只计算两个边界框的并集面积，缺点是不能反映两个框是彼此靠近还是彼此很远，IoU 的公式如式(2)。后续提出的 GIoU 引入了惩罚因子，在不重叠的情况下，预测方会向目标方框移动，如式(3)，GIoU 被用于 YOLOv3 中，成为当时最先进的目标检测器，性能远高于 IoU。而 GIoU 存在一定的局限性，即当预测边界框被真实框完全覆盖时，性能就很差。后来 DIoU 的诞生考虑到了预测边界框与真实边界框之间的质心点距离，如式(4)，然而缺点是当预测边界框的中心点与真实边界框的中心点重合时，就成了 IoU。之后 CIoU 的提出，考虑到了中心点距离和纵横比，如式(5)～(7)，但是 CIoU 定义的纵横比是相对值，不是绝对值。针对此问题提出了 EIoU，如式(8)，当预测的边界框和真实框有相同的宽高比，而宽度和高度不相等时，性能将会下降，精度和速度都有损失。

$$IoU=\frac{B_{gt}\cap B_{prd}}{B_{gt}\cup B_{prd}} \tag{2}$$

$$GIoU=IoU-\frac{|C-B_{gt}\cup B_{prd}|}{|C|} \tag{3}$$

$$DIoU=IoU-\frac{\rho^2(B_{gt},B_{prd})}{C^2} \tag{4}$$

$$CIoU=IoU-\frac{\rho^2(B_{gt},B_{prd})}{C^2}-\alpha V \tag{5}$$

$$V=\frac{4}{\pi^2}\left(\arctan\frac{W^{gt}}{h^{gt}}-\arctan\frac{W^{prd}}{h^{prd}}\right)^2 \tag{6}$$

$$\alpha=\frac{V}{1-IoU+V} \tag{7}$$

$$EIoU=DIoU-\frac{\rho^2(w_{prd},w_{gt})}{(w^c)^2}-\frac{\rho^2(h_{prd},h_{gt})}{(h^c)^2} \tag{8}$$

式中：B_{gt} 为真实边界框；B_{prd} 为预测边界框；C 为覆盖 B_{gt} 和 B_{prd} 的最小方框，$|C|$ 为方框的面积；$\rho^2(B_{gt},B_{prd})$ 为预测边界框中心点和真实边界框中心点的欧氏距离；C^2 为最小的封闭矩形的对角线长度；w、h 分别为宽和高；α 为调节因子；V 为长宽比的相似系数；w^c 和 h^c 分别为最小闭合框的宽和高。

考虑了现有基于 IoU 的损失和 n 范数损失的优缺点，提出了一种新的基于最小点距离的 BBR 损失函数，即 MPDIoU 损失函数，如式(9)～(12)。原网络使用 CIoU 损失函数，CIoU 损失函数考虑到了回归框的宽高比和预测框的中心距离，但是存在一定的局限性，比如同时存在多个框的中心点与识别标签的中心点一致，它们的宽高比是一样的但是具体的数值不同，就可能导致识别的框与原标签不同。为了减少以上缺点带来的影响，引入 MPDIoU 来替换 CIoU 损失函数，以解决现有损失的问题，并获得更快的收敛速度和更准确的回归结果。

$$d_1^2=(x_1^B-x_1^A)^2+(y_1^B-y_1^A)^2 \tag{9}$$

$$d_2^2=(x_2^B-x_2^A)^2+(y_2^B-y_2^A)^2 \tag{10}$$

$$MPDIoU=\frac{A\cap B}{A\cup B}-\frac{d_1^2}{w^2+h^2}-\frac{d_2^2}{w^2+h^2} \tag{11}$$

$$L_{MPDIoU}=1-MPDIoU \tag{12}$$

式中：(x_1^A,y_1^A)、(x_2^A,y_1^A) 表示 A 的左上和右下点坐标，(x_1^B,y_1^B)、(x_2^B,y_2^B) 表示 B 的左上和右下点坐标。所提出

的 MPDIoU 简化了两个边界框之间的相似性比较，可以适应重叠或非重叠的边界框回归。

2.3 高效的 SA 注意力机制

近年来由于注意力机制能提高对特征的提取效率，能将最有限的资源分给最需要的部分，被众多的研究者深入研究。注意力机制主要分为空间注意力和通道注意力两类，它们分别用于捕获像素级成对关系和通道依赖性，尽管将它们融合在一起时会比单独作用取得更好的性能，但是最后会增加不少的计算量。为了解决注意力轻量化的问题，引入了高效且轻量的 SA 注意力机制，SA 注意力模块采用了 Shuffle 单元有效地结合了空间和通道注意力，SA 注意力模块的结构如图 5 所示。

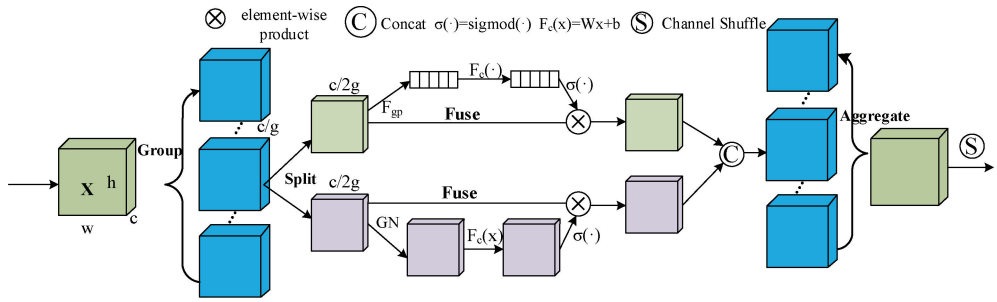


图 5 SA 注意力模块结构

Fig. 5 SA attention module structure

SA 注意力首先将得到的特征图进行分组，采用通道分割来并行处理每个组下的子特征。即对于给定的特征图 $\boldsymbol{X} \in \boldsymbol{R}^{C\times H\times W}$，其中 C、H、W 分别表示通道、空间高度和宽度，将 X 分为 G 组。$\boldsymbol{X}=[X_1,\cdots,X_G]$，$\boldsymbol{X}_k \in \boldsymbol{R}^{C/G\times H\times W}$，其中每个子功能 X_k 在训练时沿着通道维度生成两个分支，即 X_{k1}，$X_{k2} \in \boldsymbol{R}^{C/2G\times H\times W}$，一个分支通过利用通道的相互关系来产生通道注意图，另一个分支则利用特征的相互空间关系来产生空间注意图，使得到的模型可以着重关注“什么”和“哪里”是有意义的。

对于通道注意分支部分，由于完全捕获通道相关性的一种方法是采用 SE 块，但是会带来巨大的参数量，对于有轻量化需求的模型来说是不利的，所以此部分 SA 注意力模块采用了一种替代的方法，即通过使用全局平均池化(global average pooling，GAP)来生成通道统计信息，然后使用一对参数和移动通道向量；对于空间注意分支部分，该部分侧重于“哪里”这一信息部分，是对通道注意力的补充。首先采用群体范数来生成空间统计信息，然后建立一个与通道分支相似的紧凑特征。之后将这二者分支连接起来，得到所有的子特征被融合的部分，最后与 shuffle-Netv2 类似，使用通道洗牌(channel shuffle)操作符来实现不同子特征之间的通道信息得以融合。本文使用的 SA 注意力机制主要嵌入于颈部网络部分，即连接层(Concat)后一层，从实验的结果来看，不仅保留了模型的轻量化特征，还提高了模型的稳定性。

3 数据集与环境介绍

3.1 数据集

实验选用的数据集是由长沙理工大学所提供的一份全新的公开数据集(CCTSDB 2021)[23]，该数据集是通过车载记录仪所得到的，在 CCTSDB 2017 的基础上加了 500 张负样本，同时用更多困难的样本替换了之前易检测的图像，CCTSDB 2021 数据集包含城镇交通和高速公路交通总共 17 856 张图像，相比旧版本更能够提高网络的鲁棒性。

本文选择 LabelImg 软件来对数据集进行标注，生成 XML 的标注文件，再生成包含目标位置和类别适合 YOLO 格式的 txt 文件。本文数据集的分类主要分为指示(mandatory)、禁止(prohibition)和警告(warning) 3 类，如图 6 所示。其中包括 16 356 张训练集和 1 500 张验证集，用来训练的检测交通标志数总共 24 125 个，测试的交通标志总数 2 900 个。该数据集指示标签有 8 107 个，禁止标签有 14 430 个，警告便签有 4 488 个。

(a) 指示 (a) Mandatory　(b) 禁止 (b) Prohibition　(c) 警告 (c) Warning

图 6 CCTSDB 2021 数据集分类

Fig. 6 Classification of CCTSDB 2021 dataset

3.2 实验环境

本文的实验配置的 CPU 为 11th Gen Intel(R) Core(TM) i7-11800H@2.30 GHz，GPU 为 NVIDIA GeForce RTX 3060 Laptop(6 GB)，运行内存为 16 G，操作系统为 Windows10。编程语言版本为 Python 3.8.13，深度学习框架为 Pytorch GPU 1.10.0，GPU 开发环境为 CUDA11.3 和 CUDNN8.2.0。由 Anaconda3 软件提前配置好运行环境以及利用软件(Pycharm 2021.3)运行。输入图片的大小设置为 640×640，初始的学习率设置为

0.01，选用 SGD 优化器，批处理大小（batch size）设置为 16，训练的 epoch 为 200，且训练时不加载任何预训练权重（YOLOv5s.pt）。

3.3 评价指标

为了验证本文实验的有效性，采用精确率（precision），召回率（recall），平均精度（average precision，AP）、浮点计算数、平均精度均值（mean average precision，mAP）和帧率作为评价指标。精确率和召回率的公式如下：

$$Precision = \frac{TP}{TP + FP} \tag{13}$$

$$Recall = \frac{TP}{TP + FN} \tag{14}$$

式中：TP（true positive）为能够正确识别到对应的交通标志；FP（false positive）为错误地识别到交通标志；FN（false negative）为没有识别到交通标志。

其中，AP 表示的是检测的一种类别的 PR 曲线下的面积，通过 AP 值可计算 mAP，mAP 表示 IOU 取 0.5～0.95，间隔为 0.05 时，在所有类别上的对 AP 取平均值，能更好地评价模型的检测效果，如式（15）和（16）。浮点数用来衡量模型和算法的复杂程度，帧率也是检测实时性能的一个重要标准。

$$AP = \int_0^1 p(r)\mathrm{d}r \tag{15}$$

$$mAP = \frac{1}{m}\sum_{i=1}^{m} AP_i \tag{16}$$

4 实验与结果分析

4.1 轻量化网络对比实验

对网络进行轻量化时，既要保留原网络的高精确度，又要减少模型的参数，才能保证在车载设备上安全的运行。所以选取了以下几种轻量化网络进行实验对比分析，来证明本文所选用的轻量化网络的优秀性能。在 YOLOv5－7.0 版本上进行实验，训练的 epoch 为 100，结果如表 1 所示，MobileNetv3 和 PP-LCNet[24] 骨干网络相比原网络参数量降到了 1.38×10^6 和 2.96×10^6，浮点数降到了 2.3 和 5.9 GFLOPs，且精度为 41.7％和 53.8％；EfficientNet 和 GhostNet 骨干网络参数量分别降到了 3.5×10^6 和 5.85×10^6，但是精度也降到了 60.7％和 64.8％；而选用 FasterNet 后参数量降到了 5.78×10^6，精度为 74.4％，不仅比 GhostNet 更轻，精度也更高。可见轻量化骨干对本文数据集的精度影响较大，所以在轻量化和精度之间平衡，选用了 FasterNet 模块作为本文轻量化的基准，将此网络命名为 F-YOLOv5s。

表 1 轻量化模型对比

Table 1 Comparison of lightweight models

模型名称	参数量/($\times10^6$)	mAP@0.5/％	mAP@0.5:0.95/％	浮点数/GFLOPs
YOLOv5s	7.01	77.6	49.8	15.8
MobileNetv3	1.38	41.7	21.17	2.3
PP-LCNet	2.96	53.8	28.6	5.9
EfficientNet	3.5	60.7	33.2	7.5
GhostNet	5.85	64.8	37.3	12.4
FasterNet	**5.78**	**74.4**	**46**	**12.6**

4.2 损失函数对比实验

为验证损失函数改进的有效性，将本文选用的 MPDIoU 损失函数与原网络的 CIoU、目前较为先进的 EIoU 进行对比实验，训练的 epoch 为 100，实验结果如表 2 所示。从结果可知原网络 CIoU 的 mAP@0.5 为 74.4％，在参数条件不变的情况下，EIoU 损失函数的 mAP@0.5 为 75.6％，MPDIoU 损失函数的 mAP@0.5 为 78.4％。将 MPDIoU 与 CIoU、EIoU 相比，在模型的精度提升上有着明显的优势。此 3 种损失函数的收敛曲线如图 7 所示，横坐标为 epoch，纵坐标为损失值，可知 MPDIoU 损失函数和 EIoU 损失函数的收敛速度都快于 CIoU 损失函数。因此，选用 MPDIoU 作为损失函数的改进，将改进后的网络命名为 FM-YOLOv5s。

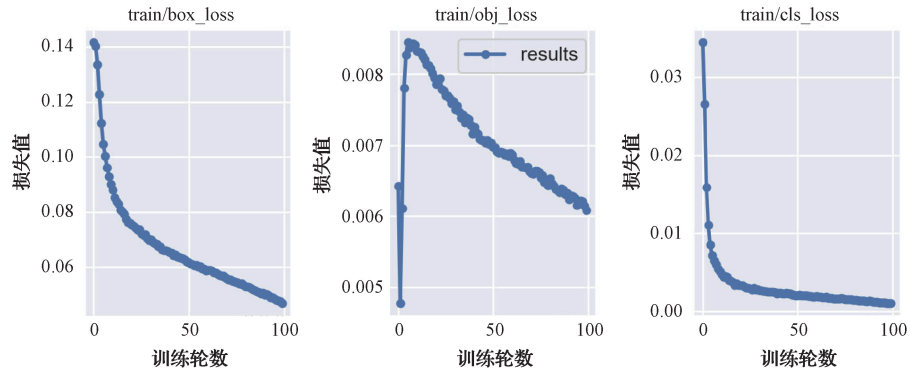
表 2 损失函数对比实验

Table 2 Loss function comparison experiments

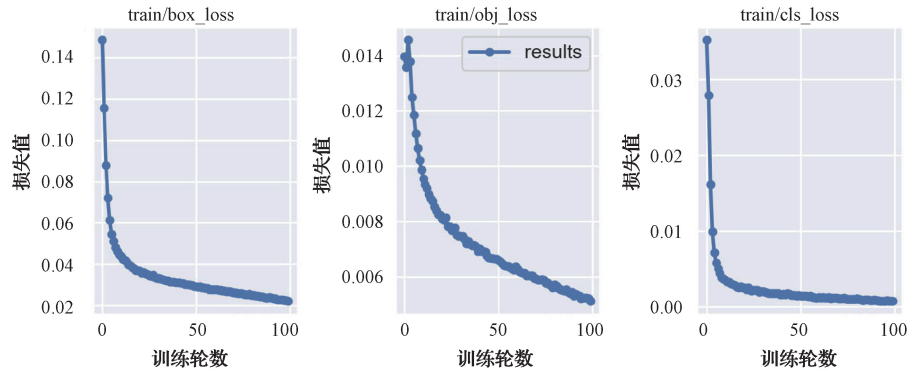
损失函数名称	参数量/($\times10^6$)	mAP@0.5/％	mAP@0.5:0.95/％
CIoU	5.78	74.4	46
EIoU	5.78	75.6	46.8
MPDIoU	5.78	78.4	48.9

4.3 注意力机制对比实验

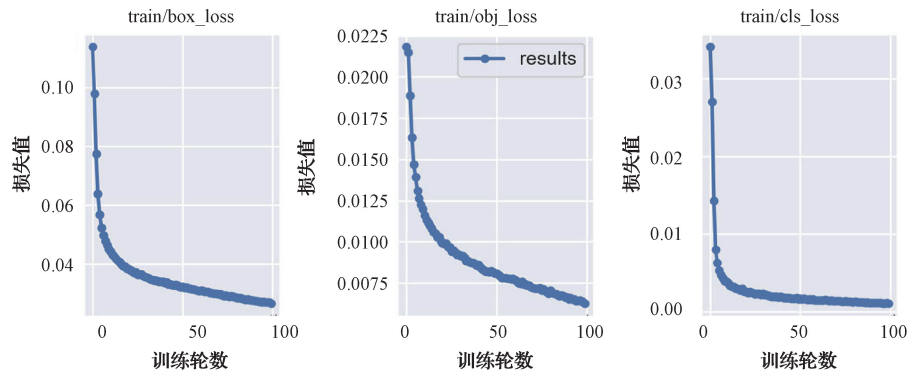
为了验证 SA 注意力机制对本文数据集的有效性，主要将该模块嵌入于 Neck 层，即上采样层之后，由于网络有两个上采样层，所以分为 FM-YOLOv5s-SA1 和 FM-YOLOv5s-SA2 两种嵌入方法，SA1 为在第 1 个上采样层的 Concat 层后嵌入，SA2 为第 1 个和第 2 个上采样层的 Concat 层后都嵌入该模块，训练的 epoch 设置为 200，再与较为先进的高效多尺度注意力（EMA）[25] 机制作对比，其中 EMA 注意力嵌入的位置与 SA1 的位置一致，实验结果如表 3 所示。



(a) CIoU损失函数
(a) CIoU loss function



(b) EIoU损失函数
(b) EIoU loss function



(c) MPDoU损失函数
(c) MPDoU loss function

图 7　损失函数收敛曲线

Fig. 7　Loss function convergence graphs

表 3　注意力机制对比实验

Table 3　Comparative experiments on attentional mechanisms

模型名称	参数量/($\times10^6$)	mAP@0.5/%	mAP@0.5:0.95/%	浮点数/GFLOPs
FM-YOLOv5s	5.78	80.3	51.9	12.6
FM-YOLOv5s-EMA	5.82	77.6	49.7	13.7
FM-YOLOv5s-SA1	5.78	80.3	52.1	12.6
FM-YOLOv5s-SA2	**5.78**	**80.4**	**52.1**	**12.6**

从以上结果可知，EMA注意力机制不仅增加了参数和计算量，而且对本文数据集的精度提升有负作用，远不如SA注意力机制。同时在不同位置嵌入SA注意力模块对模型的精度有不同的影响，其中FM-YOLOv5s-SA1的mAP@0.5:0.95提升了0.2%，FM-YOLOv5s-SA2的mAP@0.5提升了0.1%，mAP@0.5:0.95和FM-YOLOv5s-SA1相同。整体来说与FM-YOLOv5s相比有所提升，验证了SA注意力机制的有效性，综上所述，将改进后的网络命名为FMS-YOLOv5s。

4.4 消融实验

为了验证本文所添加的改进模块的有效性，对各个模块进行消融实验对比，实验结果如表4所示，epoch=200，分为A、B、C、D四组，“√”表示使用了该模块，“×”表示没有使用该模块。A组先对原网络进行轻量化处理，融合FasterNet block模块，从结果上可以看出参数量降低了1.23×10^6，浮点数减少3.2 GFLOPs，满足轻量化的要求。由于轻量化在精度上有所降低，所以对损失函数进行改进，B组即添加MPDIoU损失函数之后，精度有了明显的提升，mAP@0.5涨到了80.3%，相比YOLOv5s提升了2.3%，证明了MPDIoU损失函数的有效性。C组而单独添加SA注意力机制之后，涨到了79.8%，证明了轻量级SA注意力机制不仅不增加参数，而且对精度的提升有一定的帮助。最后D组将以上模块融合在一起，相比B组mAP@0.5提升了0.5%，相比C组mAP@0.5和mAP@0.75都提升了0.1%。从上述的实验结果可知，所改进的模块对原网络都表现出了各自优异的性能。

表4 消融实验对比结果

Table 4 Comparison results of ablation experiments

模型名称	FasterNet	MPDIoU	SA	mAP@0.5/%	mAP@0.75/%	mAP@0.5:0.95/%	参数量/($\times10^6$)	浮点数/GFLOPs
YOLOv5s	×	×	×	78	59.2	50.5	7.01	15.8
A	√	×	×	76.2	52.1	46.4	5.78	12.6
B	√	×	√	79.8	62	52.4	5.78	12.6
C	√	√	×	80.3	60.6	51.9	5.78	12.6
D	√	√	√	**80.4**	**60.7**	**52.1**	**5.78**	**12.6**

4.5 与其他算法对比

为了验证本文的改进算法的优势，与其他算法进行轻量化和精度比较。本文选用SSD、YOLOv3-tiny和YOLOv4-tiny进行对比实验，实验结果如表5所示。评价指标为参数大小、模型大小、mAP、浮点数和帧率。

表5 与其他算法对比实验

Table 5 Compared with other algorithms

算法名称	参数量/($\times10^6$)	模型大小/MB	mAP@0.5/%	mAP@0.5:0.95/%	帧率/fps	浮点数/GFLOPs
SSD(VGG)	26.28	91.6	29.81	11.7	28	62.74
YOLOv3-tiny	8.67	16.6	70.2	39.9	120	12.9
YOLOv4-tiny	5.88	22.4	53.87	24.5	131	6.8
YOLOv5s	7.01	13.7	78	50.5	99	15.8
本文算法	**5.78**	**11.3**	**80.4**	**52.1**	**102**	**12.6**

经典的SSD网络mAP@0.5达到了29.81%，在精度和轻量化方面都远不如本文的改进算法。较为轻量的YOLOv3-tiny网络mAP@0.5达到了70.2%，而本文的改进算法与其相比，参数量不仅减少2.89×10^6，mAP@0.5高出10.1%，浮点数减少了0.3 GFLOPs。YOLOv4-tiny网络mAP@0.5达到了53.87%，本文的改进算法与其相比，虽在帧率和浮点数有所劣势，但mAP@0.5高出26.43%，参数量还减少了0.1×10^6。最后与原来的YOLOv5s网络相比，参数量减少了1.23×10^6，模型大小减少了2.4×10^6，浮点数降低了3.2 GFLOPs，mAP@0.5提升了2.3%，帧率达到了102 fps。本文所改进的算法不仅保证了高精度，还兼顾了模型的轻量化，在二者之间保持了良好的平衡。

4.6 图像检测结果分析

从数据集的验证集中抽取3张含有交通标志的图像进行可视化分析，检测结果如图8所示，分别为SSD算法、YOLOv3-tiny算法、YOLOv4-tiny算法、YOLOv5s算法和本文FMS-YOLOv5s算法的检测结果。从检测的结

果可知，SSD算法只检测到了图8(a2)中一处禁止标志，漏检和误检的情况较为严重。YOLOv3-tiny和YOLOv4-tiny算法的检测结果比SSD算法要好，图8(b1)和(c1)中远处的禁止标志都能检测到，但是漏检的情况仍然存在，比如图8(b2)、(c2)中和图8(b3)、(c3)中的交通标志都未检测到。YOLOv5s算法也存在漏检的情况，每张图像都有一个交通标志未检测到。只有改进后的算法能检测到图8(e1)～(e3)中所有的交通标志，同时还保留了较高的置信度，可以看出本文改进的FMS-YOLOv5s算法在小目标方面得到了较好的提升。综上可知，FMS-YOLOv5s算法的检测效果优于原YOLOv5s、YOLOv3－tiny等算法，且更加轻便，能够满足移动端的要求。

图8　可视化检测结果

Fig. 8　Visualization of detection result

5　结　论

本文针对交通标志检测模型参数多且准确率低的问题，提出了一种改进的FMS-YOLOv5s算法。该网络通过融合FasterNet Block结构解决网络参数多的问题，再将原始的CIoU损失函数替换成MPDIoU损失函数，提升对小目标的检测能力。最后在Neck层嵌入轻量级的SA注意力机制，提高模型的泛化能力。改进后的FMS-YOLOv5s不仅保留了较高的精度，达到了80.4%，同时检测速度也得到了提升，模型大小减小到了11.3×10^{6}，能

够满足实时性的需求。接下来的工作将本文的 FMS-YOLOv5s 算法嵌入于移动端，且在降低模型实时检测中目标的漏检和误检率这一方向还有优化空间。

参考文献

[1] GUDIGAR A, CHOKKADI S, RAGHAVENDRA U, et al. Local texture patterns for traffic sign recognition using higher order spectra [J]. Pattern Recognition Letters, 2017, 94: 202-210.

[2] ELLAHYANI A, ANSARI M E, JAAFARI I E. Traffic sign detection and recognition based on random forests [J]. Applied Soft Computing, 2016, 3(46): 805-815.

[3] LIU H, LIU Y, SUN F. Traffic sign recognition using group sparse coding [J]. Information Sciences, 2014, 266(10): 75-89.

[4] LIU Z, DU J, TIAN F, et al. MR-CNN: A multi-scale region-based convolutional neural network for small traffic sign recognition [J]. IEEE Acess, 2019, 7: 57120-57128.

[5] 韩岩江，王伟，王峰萍. 融合坐标注意力和 BiFPN 的 YOLOv5s 交通标志检测方法 [J]. 国外电子测量技术，2022, 41(11): 170-179.
HAN Y J, WANG W, WANG F P. YOLOv5s traffic sign detection method based on coordinate attention and BiFPN [J]. Foreign Electronic Measurement Technology, 2022, 41(11): 170-179.

[6] 高涛，邢可，刘占文，等. 基于金字塔多尺度融合的交通标志检测算法 [J]. 交通运输工程学报，2022, 22(3): 210-224.
GAO T, XING K, LIU ZH W, et al. Traffic sign detection algorithm based on pyramid multiscale fusion [J]. Journal of Traffic and Transportation Engineering, 2022, 22(3): 210-224.

[7] 霍爱清，南思媛，胥静蓉. 改进 YOLOX 的弱光线道路交通标志检测 [J]. 电子测量技术，2023, 46(6): 62-67.
HUO A Q, NAN S Y, XU J R. Improved YOLOX's low-light road traffic sign detection [J]. Electronic Measurement Technology, 2023, 46(6): 62-67.

[8] 俞林森，陈志国. 融合前景注意力的轻量级交通标志检测网络 [J]. 电子测量与仪器学报，2023, 37(1): 21-31.
YU L S, CHEN ZH G. Lightweight traffic sign detection network with fused foreground attention [J]. Journal of Electronic Measurement and Instrumentation, 2023, 37(1): 21-31.

[9] 安胜彪，娄慧儒，白宇. 一种改进 YOLOv4 的半导体芯片表面字符识别算法 [J]. 国外电子测量技术，2022, 41(4): 77-82.
AN SH B, LOU H R, BAI Y. Character recognition of semiconductor chip surface based on improved YOLOv4 [J]. Foreign Electronic Measurement Technology, 2022, 41(4): 77-82.

[10] 郎松，曹选，张艳微，等. 融合改进 YOLOv5 算法的图像全站仪全自动测量方法 [J]. 仪器仪表学报，2022, 43(5): 120-127.
LANG S, CAO X, ZHANG Y W, et al. Fully automated measurement method of image total station based on the improved YOLOv5 algorithm [J]. Chinese Journal of Scientific Instrument, 2022, 43(5): 120-127.

[11] 吕禾丰，陆华才. 基于 YOLOv5 算法的交通标志识别技术研究 [J]. 电子测量与仪器学报，2021, 35(10): 137-144.
LV H F, LU H C. Research on traffic sign recognition technology based on YOLOv5 algorithm [J]. Journal of Electronic Measurement and Instrumentation, 2021, 35(10): 137-144.

[12] CHEN J R, KAO S H, HE H, et al. Run, don't walk: Chasing higher FLOPS for faster neural neworks [C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 12021-12031.

[13] MA S L, XU Y. MPDIoU: A Loss for efficient and accurate bounding box regression [J]. arXiv preprint arXiv: 2307. 07662, 2023.

[14] ZHANG Q L, YANG Y B. SA-net: Shuffle attention for deep convolutional neural networks [C]. ICASSP 2021 — 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2021: 2235-2239.

[15] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. YOLOv4: Optimal speed and accuracy of object detection [J]. arXiv preprint arXiv: 2004. 10934, 2020.

[16] HOWARD A, SANDLER M, Chu G, et al. Searching for mobilenetv3 [C]. Proceedings of the IEEE/CVF International Conference on Computer Vision. IEEE, 2019: 1314-1324.

[17] ZHANG X, ZHOU X, LIN M, et al. Shufflenet: An extremely efficient convolutional neural network for mobile devices [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 6848-6856.

[18] TAN M X, LE Q V. Efficientnet: Rethinking model scaling for convolutional neural neworks [C]. International Conference on Machine Learning, 2019: 6105-6114.

[19] HAN K, WANG Y, TIAN Q, et al. Ghostnet: More

features from cheap operations[C]. Proceedings of the IEEE/ CVF Conference on Computer Vision and Pattern Recognition,2020:1580-1589.

[20] REZATOFIGHI H, TSOI N, GWAK J Y, et al. Generalized intersection over union: A metric and a loss for bounding box regression[C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition,2019:658-666.

[21] ZHENG Z, WANG P, LIU W, et al. Distance-IoU loss: Faster and better learning for bounding box regression [C]. AAAI Conference on Artificial Intelligence (AAAI),2020,34(7):12993-13000.

[22] ZHANG Y F, REN W, ZHANG Z, et al. Focal and efficient IOU loss for accurate bounding box regression[J]. Neuro-Computing,2022,506:146-157.

[23] ZHANG J M, ZOU X, KUANG L D, et al. CCTSDB 2021: A more comprehensive traffic sign detection benchmark[J]. Human-Centric Computing and Information Sciences, 2022, DOI: 10.22967/HCIS.2022.12.023.

[24] CUI C, GAO T, WEI S, et al. PP-LCNet: A lightweight CPU convolutional neural network[J]. ArXiv Preprint ArXiv. 2109.15099, 2021.

[25] OUYANG D, HE S, ZHANG G, et al. Efficient multi-scale attention module with cross-spatial learning[C]. ICASSP 2023 — 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2023: 1-5.

作 者 简 介

曹立,硕士,主要研究方向为深度学习和目标检测。
E-mail:1521169257@qq.com

康少波(通信作者),博士,副教授,主要研究方向为电力变换技术、系统仿真和稳定控制算法。
E-mail:kang_shaobo@126.com