

# 基于深度学习的交通标志检测与识别方法研究

王云航

(同济大学汽车学院, 上海 201804)

【欢迎引用】王云航. 基于深度学习的交通标志检测与识别方法研究[J]. 汽车文摘, 2024(6): 30-38.

【Cite this paper】WANG Y H. Research on Detection and Recognition Methods of Traffic Signs Based on Deep Learning[J]. Automotive Digest(Chinese), 2024(6): 30-38.

**【摘要】**为克服交通标志检测与识别中交通标志目标小、尺寸多样、获取相关特征信息困难且易受到复杂背景干扰等挑战,基于YOLOv5s网络提出一种基于深度学习的交通标志检测与识别方法。针对当前交通标志识别算法在背景复杂的小目标识别方面存在的问题,将置换注意力机制嵌入YOLOv5s模型主干网络末端的C3层,提出了一种基于注意力机制的交通标志检测与识别算法,以提高对关键区域的聚焦能力,有效消除背景噪声干扰。针对目前目标检测算法在处理尺寸多变的交通标志图像时存在特征融合局限性的问题,提出了加权特征融合网络算法。该算法使主干网络中包含丰富语义信息的同尺寸浅层特征图,分别与深层的中等和大目标检测层进行加权融合,以增强多尺寸特征融合能力。实验结果表明,改进后算法在交通标志检测数据集CCTSDB 2021上,相较于原YOLOv5s方法,精确度和召回率分别提升了0.5个百分点和3.6个百分点,平均精度提升了2.8个百分点,检测速度达到123.46帧/s。因此,提出的算法能够有效提高交通标志检测和识别的准确性,同时保持原有的检测速度。

**关键词:**小目标检测;交通标志识别;YOLOv5s;置换注意力机制;加权特征融合

中图分类号:TP391.41;TN911.73 文献标识码:A DOI: 10.19822/j.cnki.1671-6329.20240075

## Research on Detection and Recognition Methods of Traffic Signs Based on Deep Learning

Wang Yunhang

(School of Automotive Studies, Tongji University, Shanghai 201804)

**【Abstract】**In order to overcome the challenges of traffic sign detection and recognition, such as small targets, diverse sizes, difficulty in obtaining relevant feature information, and susceptibility to complex background interference, traffic signs are detected and recognized. a deep learning method for traffic sign detection and recognition is proposed based on the YOLOv5s network. Addressing the issue of current traffic sign recognition algorithms struggling to identify small targets in complex backgrounds, we have integrated a shuffle attention mechanism into the C3 layer at the end of the YOLOv5s backbone network. This integration introduces a traffic sign detection and recognition algorithm that relies on an attention mechanism. This enhances the ability to focus on key areas and effectively eliminates background noise interference. To address the limitations of feature fusion in current object detection algorithms when dealing with traffic signs of varying sizes in images, we propose a weighted feature fusion network algorithm. This algorithm performs weighted fusion of shallow feature maps containing rich semantic information in the backbone network with medium and large target detection layers, enhancing the fusion ability of multi-size features. Experimental results on the traffic sign detection dataset CCTSDB 2021 show that the enhanced algorithm achieved a 0.5 percentage points increase in precision, a 3.6 percentage points increase in recall, and an average precision improvement of 2.8 percentage points compared to the original YOLOv5s method. Additionally, the detection speed reached 123.46 frame/s. Therefore, the proposed algorithm effectively enhances the accuracy of traffic sign detection and recognition while maintaining a original detection speed.

**Key words:** Small object detection, Traffic sign recognition, YOLOv5s, Shuffle attention mechanism, Weighted feature fusion

## 0 引言

交通标志检测任务旨在准确识别图像或视频中的交通标志类别及其位置。这项技术被广泛应用于自动驾驶和辅助驾驶等领域,是智能交通系统重要的组成部分,对于交通标志的准确检测具有重要的研究意义和实用价值。然而,由于实际环境中的交通标志受采集条件的影响,往往呈现目标小、背景复杂、尺寸变化多样等特点,导致识别困难甚至误判,可能对车辆的行车安全造成严重影响。因此,实时、快速地识别交通标志成为交通任务中的重点和难点之一。

国内外已有许多关于交通标志检测识别的研究,并取得了一定成果。交通标志检测可分为传统方法和基于深度学习的方法2大类。传统检测方法主要利用颜色、形状特征或二者融合进行交通标志的检测。Akatsuka等<sup>[1]</sup>提出一种阈值分割算法,通过在RGB颜色空间中设定不同颜色的阈值范围,标记交通标志的红色、黄色和蓝色等特定颜色,从而实现了对交通标志的检测。Barnes等<sup>[2]</sup>引入径向对称检测法,该方法利用点与点的对称性,将圆形映射为点的集中聚集,突出图像中的圆对称区域,以实现交通标志的粗定位。汤凯等<sup>[3]</sup>提出了一种多特征协同方法,结合颜色特征、形状特征和尺度特征,并采用支持向量机(Support Vector Machine, SVM)对融合特征进行分类,从而获得交通标志的检测结果。然而,该方法在处理小尺度曲率直方图时容易受到边缘噪声的干扰,导致尺度较小的交通标志难以被正确检测到。Qian等<sup>[4]</sup>将基于颜色的分割方法和模板匹配方法的优点结合起来,提出了一种新的几何形状特征表达方法,即多级链码直方图(Multi-level Chain Code Histogram, MCCH),用于交通标志的识别。He等<sup>[5]</sup>利用改进的局部二值模式(Local Binary Patterns, LBP)进行局部特征提取,并采用离散小波变换的低频系数作为全局信息,然后将这2种特征级联起来进行交通标志识别。然而,传统的交通标志检测方法对特征提取依赖手工操作,鲁棒性较差,无法满足交通标志检测对实时性和精确性的要求。

2014年,首次将卷积神经网络应用于目标检测<sup>[7]</sup>,自此以后,基于深度学习的方法开始广泛应用于交通标志检测。深度学习的目标检测算法主要分为2个方向:二阶段检测和一阶段检测。二阶段检测算法包括R-CNN(Regions with CNN features)及其改进版本Fast R-CNN<sup>[8]</sup>、Faster R-CNN<sup>[9]</sup>等。一阶段检测算法包括YOLO(You Only Look Once)系列<sup>[10-13]</sup>和SSD(Single

Shot multibox Detetor)<sup>[14-15]</sup>等,目标检测发展状况如图1所示。

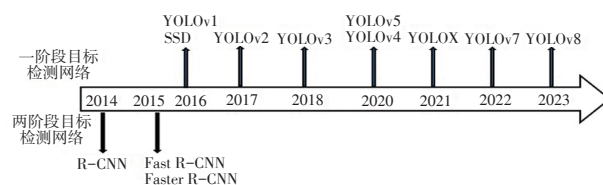


图1 目标检测发展简史

Yang等<sup>[17]</sup>基于Faster R-CNN框架引入了注意力网络,以便快速搜索所有可能的感兴趣区域,并根据颜色特征将其分为3类。然后,另外一个区域提取网络从一组锚框中生成最终的候选区域。Jin等<sup>[18]</sup>在SSD框架的基础上,引入了特征融合层,并在融合后将SE(Squeeze and Excitation)模块添加到特征提取层,极大地提高了对小型交通标志的识别率。YOLO算法于2016年首次提出,其最大的贡献在于采用S×S网格代替传统的目标区域,通过一次回归同时输出目标的坐标和类别概率。Zhang等<sup>[19]</sup>提出了一种改进的YOLOv2算法,用于交通标志识别。为了减少计算量,在网络的中间层引入了多个1×1卷积层,并减少了顶部的卷积层。同时,为了检测小型交通标志,采用更密集的网格划分图像,以获得更精细的特征图。Sichkar等<sup>[20]</sup>首先使用YOLOv3将交通标志根据形状分为4类进行定位,然后使用另一个卷积神经网络对定位到的交通标志进行准确分类,在GTSRB数据集上进行试验,其平均精确度(mean Average Precision, mAP)达到了97.22%。Liu等<sup>[21]</sup>在YOLOv5网络中引入了MobileNetV2作为主干网络,使整个模型的参数数量减少了60%,同时mAP提升了0.13个百分点。张上等<sup>[22]</sup>通过对YOLOv5网络进行模型裁剪和运算参数压缩,并在模型中嵌入了坐标注意力机制(Coordinate Attention, CA)和混合域注意力模块(Convolutional Block Attention Module, CBAM),从而使网络的平均精度提升了2.8个百分点。

由于交通标志检测中小目标占比较大,同时受到背景复杂和尺寸多样性因素的影响,本文旨在提高交通标志检测的精度和处理速度,从而提升整体的检测性能,更好地将交通标志检测用于实际工程应用。本文中的算法以YOLOv5s框架为基础进行改进,改进思路如下:针对当前交通标志识别算法对于背景复杂的小目标识别效果不佳的问题,提出了一种基于注意力机制的交通标志检测与识别算法。在特征提取主干的C3结构嵌入置换注意力机制(Shuffle Attention,



中不同位置的信息重要性。尽管CBAM<sup>[16]</sup>成功地将这2种注意力融合,以此获得了较好的性能,但是存在计算量大和收敛困难等问题。

SA注意力模块通过采用洗牌方式同时结合了通道和空间注意力(图4)。首先,在每个注意力单元起始端,对输入的特征图在通道维度上进行分组,形成多个子特征组,即 $[X_1, \dots, X_c]$ ,其中 $X_k \in R^{C/G \times H \times W}$  ( $C, H, W, G$ 分别表示通道数、空间高度、空间、宽度和分组数)。每个分支再进一步沿通道维度划分为2个部分,即 $X_{k1}, X_{k2} \in R^{C/2G \times H \times W}$ ,分别用于生成通道注意力图和空间注意力图,然后将2条分支按通道数相加,实现分支的融合,最后,聚合所有的分组子特征,利用“通道置换”(Channel Shuffle)算子,随机置换不同通道之间的信息,从而实现不同子特征之间的通信,有效结合通道和空间注意力。

通道注意力机制为了捕获更丰富的通道信息并平衡计算速度的影响,采用全局平均池化(Global Averaging Pooling, GAP)的方式,计算通道统计量 $s$ :

$$s = F_{gp}(X_{k1}) = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W X_{k1}(i, j) \quad (1)$$

式中: $s \in R^{C/2G \times 1 \times 1}$ ;  $F_{gp}$ 为全局平均池化函数;  $X_{k1}$ 为子特征 $X_k$ 的通道注意力分支,  $X_{k1} \in R^{C/2G \times H \times W}$ ;  $C, H, W, G$ 分别为空间高度、空间、宽度和分组数。

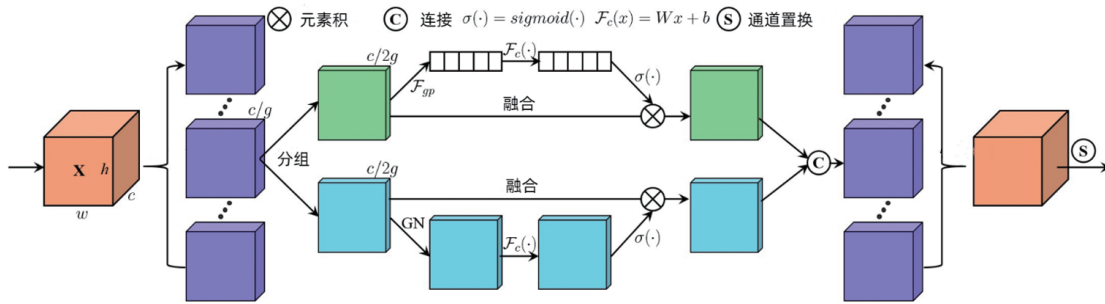


图4 SA注意力机制结构

## 2.2 加权特征融合模块

在目标检测中,深度网络的进一步发展可以增强模型的语义特征提取能力。然而,随着网络深度增加,输出特征图尺寸缩小且图像细节特征有限,对于小目标检测可能存在瓶颈。相反,浅层网络能更有效地捕获丰富的细节信息,但缺乏足够的语义信息,影响模型在物体分类上的准确性。

为了克服这些限制,学术界开始探索结合深层和浅层特征的方法,以提取更全面且丰富的特征表示。目前流行的多尺度特征融合架构包括特征金字塔网络(Feature Pyramid Networks, FPN)<sup>[28]</sup>、路径聚合网络

此外,通道注意力为了选择的精确性及自适应性,利用门控机制对 $X_{k1}$ 进行处理,该机制采用sigmoid激活函数,最终输出的计算公式如下:

$$X'_{k1} = \sigma(F_c(s)) \cdot X_{k1} = \sigma(W_1 s + b_1) \cdot X_{k1} \quad (2)$$

式中: $\sigma$ 是激活函数, $F_c(\cdot)$ 为线性函数, $W_1 \in R^{C/2G \times 1 \times 1}$ 和 $b_1 \in R^{C/2G \times 1 \times 1}$ 分别用于缩放和移位统计量 $s$ 。

在空间注意力分支中,采用群范数(Group Norm, GN)获取空间特征,然后通过全连接层 $F_c(\cdot)$ 进行线性变换和激活函数处理以增强这些特征。最终,空间注意力的输出计算公式如下:

$$X'_{k2} = \sigma(W_2 \cdot GN(X_{k2}) + b_2) \cdot X_{k2} \quad (3)$$

式中: $W_2 \in R^{C/2G \times 1 \times 1}$ 和 $b_2 \in R^{C/2G \times 1 \times 1}$ 是形状为 $R^{C/2G \times 1 \times 1}$ 的参数, $GN$ 为群范数, $X_{k2}$ 为子特征 $X_k$ 的空间注意力分支, $X_{k2} \in R^{C/2G \times H \times W}$ 。

最后,将空间和通道2个注意力分支进行连接,使得通道数与输入相同,如式(4)所示。利用通道置换算子,使得每个分支的特征信息能够沿通道跨分支流动。

$$X'_k = [X'_{k1}, X'_{k2}] \in R^{C/G \times H \times W} \quad (4)$$

本文将置换注意力SA模块嵌入到C3模块中,进行先降维、后升维卷积操作后,形成特征增强模块C3SA,如图5所示。由于关键特征提取在主干部分进行,因此特征增强模块被放置在主干的末端位置。

(Path Aggregation Network, PANet)<sup>[26]</sup>、加权双向特征金字塔网络(Bi-directional Feature Pyramid Network, BiFPN)<sup>[29]</sup>,这些架构设计在提升检测性能方面展现了极大的潜力。

如图6a所示,FRN通过自顶向下传递强语义信息,将深层特征与浅层特征进行融合。然而,FPN结构也存在一些缺陷,它更倾向于关注相邻层的特征,而深层特征经过多次下采样传递后可能丧失语义信息。为了解决这个问题,PANet同时采用自上而下和自下而上2条路径进行特征融合(见图6b)。这种结构缩短信息传递的路径,有效地避免FPN中可能发生

的语义信息丧失情况,但增加了计算开销。

在2020年,Tan<sup>[29]</sup>基于PANet提出BiFPN(见图6c),并带来以下改进:首先,删除只有单输入节点,因为这些节点对特征融合没有贡献,只会增加计算量;其次,添加跳跃连接,即在同一水平上从原始输入到输出节点添加额外的边,以便在不增加太多成本的情况下融合更多特征;再次,将每个由自顶向下和自底向上形成的双向路径视为一个网络层,对这些网络层进行多次重复,实现更高级的特征融合;最后,根据不同的输入对最终特征的贡献度,为每个输入分配不同的权重,以调整不同特征输入在输出中的比例,并在网络训练中实时地调整这些权重。

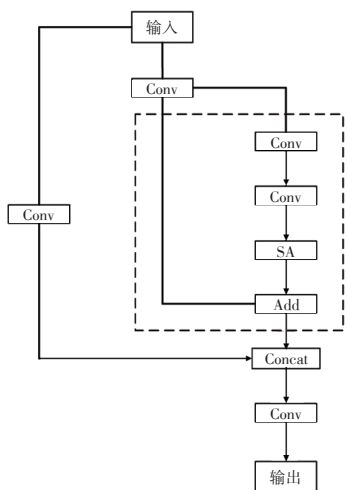


图5 C3SA 特征增强模块结构

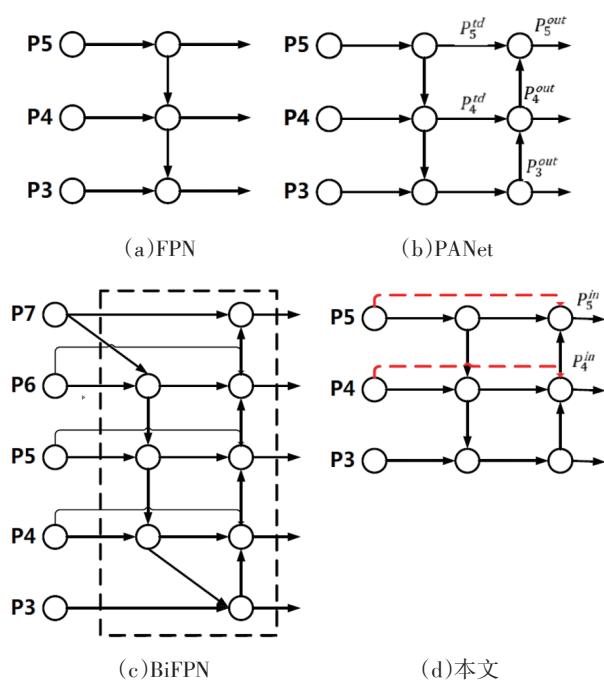


图6 多尺度特征融合网络结构

原始的YOLOv5s模型使用FPN+PAN结构进行颈

部网络特征融合,形成自顶向下和自底向上的双向路径。然而,这种处理方式相对较简单,特征的融合还有进一步优化的空间。受到加权特征金字塔网络的启发,本节提出将BiFPN的设计思想应用到YOLOv5s的多尺度特征融合部分,提出加权特征融合网络(Weighted Path Aggregation Network, WPANet)。

首先,添加跳跃连接。当特征融合层存在多个融合节点时,在同一层不同的节点之间添加多条特征传输边。这样,网络能够在增加少量参数和计算量的情况下融合更多的特征信息,如图6d所示的P<sub>4</sub>标注的间断线。

其次,增加特征融合。在YOLOv5s的设计中,检测头部署在网络的深层位置,使其能够接触到丰富的语义信息。然而,随着特征层级的深入,浅层的关键位置信息容易丢失,而这些信息对于提升检测的准确性至关重要。为应对这一挑战,本文创新性地将浅层特征与深层的检测输出进行紧密融合。特别地,在网络的早期阶段,主干特征层会产生逐步缩小尺寸的特征图,其中一些特征图的大小正好与检测层的输出匹配。因此,本文选择20×20的主干特征图,与负责大尺寸目标的检测层P<sub>5</sub>融合,以此增强对小型目标边缘信息的捕获。这种融合策略与BiFPN中消除单一节点的方法存在显著的差异(P<sub>5</sub>标注的虚线),见图6d。

最后,加权分参。在合并不同分辨率的多尺度特征方面,以往的一些目标检测算法常常将每个层级输出的特征信息等同对待,并通过上采样或下采样的方式将它们尺寸调整为一致后再进行相加。然而,近年来的研究发现,不同分辨率的输入特征对于输出特征的影响并不相同。为了提高算法的鲁棒性和增强网络的特征融合能力,本文引入了注意力的思想,为每个层级输出的特征信息引入了一个超参数。这种方法使网络能够学习到不同层级特征的重要程度,从而更好地优化特征的融合过程。对于传统的PANet,它以简单的求和方式聚合多尺度特征。

$$P_4^{out} = Conv(P_4^{id} + Resize(P_3^{out})) \quad (5)$$

$$P_5^{out} = Conv(P_5^{id} + Resize(P_4^{out})) \quad (6)$$

式中: P<sub>3</sub><sup>out</sup>、P<sub>4</sub><sup>out</sup>、P<sub>5</sub><sup>out</sup> 分别代表输出特征信息, P<sub>4</sub><sup>id</sup>、P<sub>5</sub><sup>id</sup> 分别代表输入特征信息,如图6b所示;Resize使用上采样或者下采样来调整特征图的分辨率,使得不同输入尺度的特征图分辨率相匹配;Conv是对求和后的特征图使用卷积操作来提取新的特征。

针对WPANet,采用加权特征融合的方法来整合不同分辨率的输入特征图层。不同分辨率的输入层对应

的权重也不同。通过网络自动学习每个输入层的权重参数,可以更好地表示整体特征信息。对于 $P_4$ 和 $P_5$ 输出部分,分别进行快速归一化融合,具体计算如下:

$$P_4^{out} = Conv \left( \frac{w_1 \cdot P_4^{in} + w_2 \cdot P_4^{id} + w_3 \cdot Resize(P_3^{out})}{w_1 + w_2 + w_3 + \epsilon} \right) \quad (7)$$

$$P_5^{out} = Conv \left( \frac{w_1 \cdot P_5^{in} + w_2 \cdot P_5^{id} + w_3 \cdot Resize(P_4^{out})}{w_1 + w_2 + w_3 + \epsilon} \right) \quad (8)$$

式中: $P_4^{in}$ 、 $P_5^{in}$ 分别为新增输入特征信息,如图6d虚线所示; $w_1$ 为当前层输入的权重; $w_2$ 为当前层中过渡单元输出的权重; $w_3$ 为前一层输出的权重; $\epsilon$ 为一个超参数,用于防止梯度消失; $Conv$ 表示对整体计算结果的卷积运算。

### 3 实验结果与分析

#### 3.1 数据集

为了验证算法效果,本文选用了中国交通标志检测数据集2021(CCTSDB)<sup>[30]</sup>。数据集包含16356张图片,分辨率在1000×350至1024×768之间。图片覆盖了各种尺度、光照和噪声等影响因素的变化,标注了强制(Mandatory)、禁止(Prohibitory)和警告(Warning)3大类标志(图7)。在实验中,按照随机8:2的比例将数据集划分为训练集和验证集。



图7 中国交通标志分类

#### 3.2 实验环境

本文实验使用CCTSDB 2021交通标志数据集,其中图像尺寸为640×640。网络参数设置如下:批量大小(每次输入神经网络进行训练的样本数量)设为32,权重衰减参数(momentum)设为0.98,衰减系数(decay)设为0.001,训练迭代轮数设为200,学习速率设为0.001。计算机设备采用Linux服务器,详细试验训练环境配置参数如表1所示。

#### 3.3 评价指标

常用的交通标志检测与识别方法通常采用总体样本平均精度(mAP)和每秒的检测帧数(Frames Per Second, FPS)为主要评价指标。其中,mAP用于衡量

平均检测精度,而FPS则代表模型对图像数据的处理速度,FPS值越大表示模型处理图像的速度越快。此外,本文还使用精确度(Precision, P)和召回率(Recall, R)等指标作为实验的辅助评价标准,其具体计算公式如下:

$$P = \frac{TP}{TP + FP} \quad (9)$$

$$R = \frac{TP}{TP + FN} \quad (10)$$

$$AP = \int_0^1 P dR \quad (11)$$

$$mAP = \frac{1}{n} \sum_{i=1}^n AP_i \quad (12)$$

式中: $P$ 为精度, $R$ 为召回率, $TP$ 、 $FP$ 、 $FN$ 分别为预测为正样本且实际为正样本、预测为正样本实际为负样本、预测为负样本实际为正样本数量, $n$ 为样本类别数量,本文中取值为3。

表1 实验训练环境配置

操作系统	Ubuntu 18.04(Linux)
深度学习框架	PyTorch1.7.0
CPU	15 vCPU AMD EPYC 7543 32-Core Processor
GPU	Nvidia GeForce RTX 3090(24 GB)
内存/G	80
CUDA	11.5
语言	Python 3.8
编译器	Pycharm

#### 3.4 消融实验

本实验在基于3.2章节相同环境配置下,通过在CCTSDB 2021交通标志数据集上进行200轮训练,对4种网络模型(YOLOv5s、YOLOv5s\_C3SA、YOLOv5s\_WPA Net、YOLOv5s\_C3SA+WPA Net)进行了评估,训练损失函数的示意如图8所示。由图8可知,在训练的初期阶段,损失函数曲线迅速下降。随着训练次数达到约40轮,各网络模型的损失值逐渐趋于稳定,并最终稳定在大约0.024的水平。其中,代表YOLOv5s\_C3SA+WPA Net的橙色损失函数曲线在整个训练阶段基本都位于其他曲线的下方,表现出最优的性能。

消融实验结果对比如表2所示。根据表中的不同网络模型参数和测试结果,可以得出以下结论:相较于YOLOv5s、YOLOv5s\_C3SA和YOLOv5s\_WPA Net模型,本文提出的YOLOv5s\_C3SA+WPA Net(即改进后算法)在mAP方面分别提升了2.8、1、0.9百分点。在精确度方面,分别提升了0.5、0.7、0.5百分点;召回率方面分别提升了3.6、0.6、1.1百分点。尤其重要的是,

改进后的模型在检测速度方面基本保持不变,达到了123.46帧/s。这一速度已足够满足实际情况下对模型实时性的要求,从而证明了改进后模型的有效性。

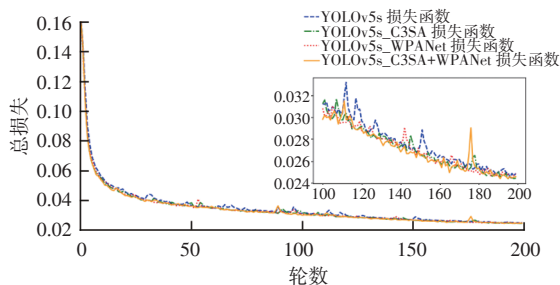


图8 网络模型损失函数

表2 消融实验测试结果对比表

网络模型	P/%	R/%	mAP/%	FPS/帧·s <sup>-1</sup>
改进后算法	89.30	70.40	77.20	123.46
YOLOv5s_WPANet	88.80	69.30	76.30	125.00
YOLOv5s_C3SA	88.60	69.80	76.20	121.95
YOLOv5s	88.80	66.80	74.40	126.58

为了验证置换注意力机制的有效性,本文还与其他注意力机制进行了横向比较。由于置换注意力机制兼具同时处理通道和空间信息的特点,其检测精度优于其他2种机制。前文已经提到,轻量化的置换注意力机制在检测速度方面优于混合域注意力机制<sup>[24]</sup>。本文将置换注意力机制与坐标注意力机制(CA)<sup>[27]</sup>,在相同位置以相同形式分别嵌入YOLOv5s模型,进行对比实验,结果见表3。嵌入注意力机制的2种算法相比于原YOLOv5s算法,检测精度均有明显提升。实验YOLOv5s\_C3SA网络模型相对于YOLOv5s\_C3CA网络模型在检测精度和检测速度方面表现出一定优势,而在精确度和召回率方面,分别提升了1.1、0.4个百分点,因此,本文采用置换注意力机制。

表3 嵌入不同注意力机制的模型测试结果对比

网络模型	P/%	R/%	mAP/%	FPS/帧·s <sup>-1</sup>
YOLOv5s_C3SA	88.60	69.80	76.20	121.95
YOLOv5s_C3CA	87.50	69.40	76.00	120.48
YOLOv5s	88.80	66.80	74.40	126.58

为深入研究改进后的网络模型在所有类别上的检测效果,本文整理了模型在不同类别上的对比结果,见表4。相对于原始的YOLOv5s网络模型,改进后的网络模型mAP提升显著:在强制、禁止和警告标志类别上,分别提高了1.9、3.3、3个百分点。在精确度方面,与原始模型相比,改进后的网络模型在强制标志上略微下降了1.5个百分点,在禁止标志上提升了3百分

点,而在警告标志类别上保持稳定。在召回率指标上,改进后的网络模型在强制、禁止和警告标志类别分别提升了3.7、2.9、4.1个百分点。通过这些数据可以看出,改进后的网络模型明显提升了在各类交通标志上的检测效果。

表4 模型检测不同类别测试结果对比

网络模型	类别	P/%	R/%	mAP/%
改进后算法	强制	85.40	64.10	69.50
	禁止	92.10	68.40	77.20
	警告	90.50	78.70	84.80
原算法 (YOLOv5s)	强制	86.90	60.40	67.60
	禁止	89.10	65.50	73.90
	警告	90.50	74.60	81.80

### 3.5 与主流算法对比实验

将经过优化的网络模型Ours与现有主流模型如SSD、Faster R-CNN,以及早期版本的YOLO进行对比。根据表5中展示的实验结果,优化后的网络模型在检测速度方面虽比原始的YOLOv5s略微下降,但在平均精度、准确度和召回率指标上均表现出显著优势。综合考虑,本文提出的算法能够更准确地识别CCTSDb 2021数据集中的交通标志,进而有助于提升交通的安全水平。此外,改进后的模型能够满足实时应用场景对实时性的关键要求。

### 3.6 检测效果对比

本节选取了CCTSDb数据集中的几类特殊图片,包括大雾、雪天、夜晚逆光和雨天夜晚条件下的交通标志,进行了识别效果对比,结果如图9所示。

表5 与主流算法测试结果对比

网络模型	P/%	R/%	mAP/%	FPS/帧·s <sup>-1</sup>
Faster R-CNN <sup>[30]</sup>	84.43	54.98	56.58	4.87
SSD <sup>[30]</sup>	86.47	27.74	49.2	22.33
YOLOv3 <sup>[30]</sup>	84.63	42.71	50.48	20.34
YOLOv4 <sup>[30]</sup>	76.16	52.50	51.69	16.55
YOLOv5s	88.80	66.80	74.40	126.58
改进后算法	89.30	70.40	77.20	123.46

在大雾条件下(见图9a、图9b),改进后算法相较于原算法(YOLOv5s)在识别禁止标志时置信度提高了21%。在雪天环境下(见图9c、图9d),改进后算法相较于原算法在识别警告标志时置信度提高了7%。在夜晚逆光的情况下(见图9e、图9f),改进后算法相较于原算法在识别强制标志时置信度提高了5%。在雨

天夜晚的情况下(见图9g、图9h),改进后的算法相较于原算法在识别禁止标志时置信度提高了5%。这些结果表明,即使在特殊复杂的环境下,改进后算法仍能够保持较高的精度,具有一定的鲁棒性。



(a)大雾原算法



(b)大雾改进算法



(c)雪天原算法



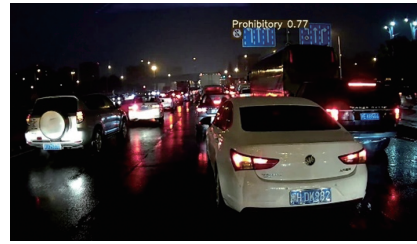
(d)雪天改进算法



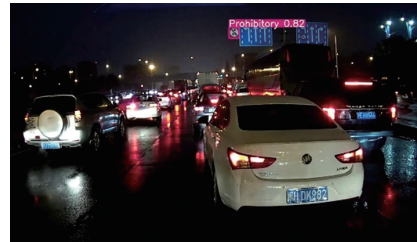
(e)夜晚逆光原算法



(f)夜晚逆光改进算法



(g)雨天夜晚原算法



(h)雨天夜晚改进算法

图9 特殊环境下检测效果对比

#### 4 结束语

针对交通标志检测中存在的目标小、尺寸多样等难题,本文在YOLOv5s的基础上进行了改进。首先,利用特征增强模块C3SA提高对关键区域的聚焦能力;其次,采用加权融合网络增强多尺寸特征的融合能力。改进后算法改善了交通标志检测精度。通过消融试验和对比试验分析,证明了本文改进后算法在检测精度上有一定的提升,并且检测速度上能够达到实时性要求。未来的研究可以围绕将算法优化应用于嵌入式工程应用中的交通标志识别。

#### 参 考 文 献

- [1] AKATSUKA H, IMAI S. Road Signposts Recognition System[J]. SAE Transactions, 1987, 96(1): 936-943.
- [2] BARNES N, ZELINSKY A, FLETCHER L S. Real-Time Speed Sign Detection Using the Radial Symmetry Detector[J]. IEEE Transactions on Intelligent Transportation Systems, 2008, 9(2): 322-332.
- [3] 汤凯, 李实英, 刘娟, 等. 基于多特征协同的交通标志检测[J]. 计算机工程, 2015, 41(3): 211-217.
- [4] QIAN R, ZHANG B, YUE Y, et al. Traffic Sign Detection By Template Matching Based on Multi-level Chain Code Histogram[C]// 2015 12th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD). Zhang Jia jie, China: IEEE, 2015.
- [5] HE X, DAI B. A New Traffic Signs Classification Approach Based on Local and Global Features Extraction[C]// 2016 6th International Conference on Information Communication and Management (ICICM). Hatfield, UK: IEEE, 2016.

- [6] ZOU Z, CHEN K, SHI Z, GUO Y, et al. Object Detection in 20 Years: A Survey[J]. Proceedings of the IEEE, 2023, 111(3): 257–276.
- [7] 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. Columbus, OH, USA: IEEE, 2014.
- [8] GIRSHICK R. Fast R-CNN[C]//2015 IEEE International Conference on Computer Vision (ICCV). Santiago, Chile, USA: IEEE, 2015.
- [9] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2017, 39(6): 1137–1149.
- [10] REDMON J, DIVVALA S, GIRSHICK R, et al. You Only Look Once: Unified, Real-Time Object Detection[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, NV, USA: IEEE, 2016.
- [11] REDMON J, FARHADI A. YOLO9000: Better, Faster, Stronger[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, HI, USA: IEEE, 2017.
- [12] REDMON J, FARHADI A. YOLOv3: An Incremental Improvement[J/OL]. arXiv e-prints, 2018[2024-04-26]. <https://arxiv.org/pdf/1804.02767>.
- [13] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. YOLOv4: Optimal Speed and Accuracy of Object Detection[J/OL]. arXiv, 2020[2024-04-25]. <https://arxiv.org/abs/2004.10934>.
- [14] WEI L, DRAGOMIR A, DUMITRU E, et al. SSD: Single Shot MultiBox Detector[C]//14th European Conference Amsterdam, The Netherlands: ECCV, 2016.
- [15] SIMONYAN K, ZISSERMAN A. Very Deep Convolutional Networks for Large-Scale Image Recognition[J/OL]. arXiv [2015-04-10](2024-05-21) <https://doi.org/10.48550/arXiv.1409.1556>.
- [16] WOO S, PARK J, LEE J Y, et al. CBAM: Convolutional Block Attention Module[J/OL]. arXiv[2018-01-18](2024-05-21) <https://doi.org/10.48550/arXiv.1807.06521>.
- [17] YANG T, LONG X, SANGAIAH A K, et al. Deep Detection Network for Real-Life Traffic Sign in Vehicular Networks[J]. Computer Networks, 2018, 13(6): 95–104.
- [18] JIN Y, FU Y, WANG W, et al. Multi-Feature Fusion and Enhancement Single Shot Detector for Traffic Sign Recognition[J]. IEEE Access, 2020, 8: 38931–38940.
- [19] ZHANG Z, WANG H, ZHANG J, et al. A Vehicle Real-time Detection Algorithm Based on YOLOv2 Framework[C]//Real-Time Image and Video Processing, 2018.
- [20] SICHKAR V N, KOLYUBIN S A. Real Time Detection and Classification of Traffic Signs Based on YOLO Version 3 Algorithm[J]. Scientific and Technical Journal of Information Technologies Mechanics and Optics, 2020, 20(3): 418–424.
- [21] LIU X, JIANG X K, HU H C, et al. Traffic Sign Recognition Algorithm Based on Improved YOLOv5s[C]//Proceedings of the 2021 International Conference on Control, Automation and Information Sciences (ICCAIS). Xi'an: IEEE, 2021: 980–985.
- [22] 张上, 王恒涛, 冉秀康. 基于YOLOv5的轻量化交通标志检测方法[J]. 电子测量技术, 2022, 45(8): 129–135.
- [23] 胡昭华, 王莹. 改进YOLOv5的交通标志检测算法[J]. 计算机工程与应用, 2023, 59(1): 82–91.
- [24] ZHANG Q L, YANG Y B. SA-Net: Shuffle Attention for Deep Convolutional Neural Networks[C]//Proceedings of 2021 IEEE International Conference on Acoustics. Toronto, Canada: IEEE, 2021: 2235–2239.
- [25] TIAN Z, SHEN C, CHEN H, et al. Focus: Fully Convolutional One-Stage Object Detection[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision, 2019: 9627–9636.
- [26] LIU S, QI L, QIN H, et al. Path Aggregation Network for Instance Segmentation[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, UT, USA: IEEE, 2018.
- [27] HOU Q, ZHOU D, FENG J. Coordinate Attention for Efficient Mobile Network Design[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Nashville, TN, USA: IEEE, 2021.
- [28] LIN T Y, DOLLÁR P, GIRSHICK R, et al. Feature Pyramid Networks for Object Detection[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, HI, USA: 2017.
- [29] TAN M, LE Q V. EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks[C]//Proceedings of the IEEE International Conference on Intelligent Transportation Systems, 2015, 17(7): 2022–2031.
- [30] ZHANG J, ZOU X, LI D K, et al. CCTSDB 2021: A More Comprehensive Traffic Sign Detection Benchmark[J]. Human-centric Computing and Information Sciences, 2022, 12: 23.

(责任编辑 明慧)