



计算机工程
Computer Engineering
ISSN 1000-3428, CN 31-1289/TP

《计算机工程》网络首发论文

题目: 基于 DMC-YOLO 的交通标志实时检测算法
作者: 栾孟娜, 郑秋梅, 王风华
DOI: 10.19678/j.issn.1000-3428.0069409
网络首发日期: 2024-06-28
引用格式: 栾孟娜, 郑秋梅, 王风华. 基于 DMC-YOLO 的交通标志实时检测算法[J/OL]. 计算机工程. <https://doi.org/10.19678/j.issn.1000-3428.0069409>



网络首发: 在编辑部工作流程中, 稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定, 且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式 (包括网络呈现版式) 排版后的稿件, 可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定; 学术研究成果具有创新性、科学性和先进性, 符合编辑部对刊文的录用要求, 不存在学术不端行为及其他侵权行为; 稿件内容应基本符合国家有关书刊编辑、出版的技术标准, 正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性, 录用定稿一经发布, 不得修改论文题目、作者、机构名称和学术内容, 只可基于编辑规范进行少量文字的修改。

出版确认: 纸质期刊编辑部通过与《中国学术期刊 (光盘版)》电子杂志社有限公司签约, 在《中国学术期刊 (网络版)》出版传播平台上创办与纸质期刊内容一致的网络版, 以单篇或整期出版形式, 在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊 (网络版)》是国家新闻出版广电总局批准的网络连续型出版物 (ISSN 2096-4188, CN 11-6037/Z), 所以签约期刊的网络版上网络首发论文视为正式出版。

基于 DMC-YOLO 的交通标志实时检测算法

栾孟娜, 郑秋梅, 王风华

(中国石油大学(华东), 计算机科学与技术学院, 山东 青岛 266580)

摘 要: 在交通标志检测中由于外界环境的干扰以及驾驶场景中交通标志目标较小的特点, 导致提升交通标志检测性能一直是一项具有挑战性的任务。本文提出了一种新的交通标志检测算法, 其能够在保证实时检测的情况下明显提高模型的检测精度。该算法首先设计一种新的多尺度特征提取网络, 引入大尺度特征来增加小目标定位信息, 同时设计多尺度特征注意增强模块进一步获得目标的上下文信息。其次, 为了降低模型的计算量和复杂度, 本文对原始模型的多尺度检测头进行改进, 选取两个大尺度检测头对小目标进行检测。最后该算法对 CIOU 损失函数进行改进, 增强算法对小目标的感知能力, 同时提高网络的训练效率。将改进后的模型在两个开源的公共数据集上进行实验, 实验结果表明, 该方法在 TT100K 和 CCTSDB2021 交通标志数据集上对交通标志小目标的检测精度均有提高, 在两个数据集的测试集上 mAP 分别达到 84.8%、83.6% 的检测精度, 较基准算法分别提升了 3%、3.6%, 具有更高的检测性能和特征提取能力, 且满足实时检测的需求。

关键词: 交通标志检测; 多尺度特征融合; 注意力机制; 膨胀卷积; 小目标检测

Real-time traffic sign detection algorithm based on DMC-YOLO

Luan Mengna, Zheng Qiumei, Wang Fenghua

(College of Computer Science and Technology, China University of Petroleum (East China), Qingdao, Shandong 266580, China)

[Abstract] In traffic sign detection, the challenge of enhancing detection performance persists due to external environmental interference and the characteristic small size of traffic sign targets in driving scenarios. This paper introduces a novel traffic sign detection algorithm that significantly improves model detection accuracy while ensuring real-time detection capabilities. The algorithm initially designs a new multi-scale feature extraction network, incorporating large-scale features to augment small target localization information, and simultaneously designs a multi-scale feature attention enhancement module to further enhance the model's feature extraction capability. Secondly, to reduce the model's computational load and complexity, this paper improves upon the original model's multi-scale detection heads, selecting two large-scale detection heads for detecting small targets. Finally, the algorithm modifies the CIOU loss function to enhance the algorithm's perception of small targets and improve the network's training efficiency. Experiments conducted on two open-source public datasets demonstrate that the improved model achieves enhanced detection accuracy for small traffic sign targets on the TT100K and CCTSDB2021 traffic sign datasets, with mAP reaching 84.8% and 83.6% on the test sets of the two datasets, respectively. This represents an improvement of 3% and 3.6% over the baseline algorithms, showcasing higher detection performance and feature extraction capabilities, while meeting the requirements for real-time detection.

[Key word] traffic sign detection; multi-scale feature fusion; attention mechanism; dilated convolution; small object detection

DOI:10.19678/j.issn.1000-3428.0069409

0 概述

随着深度学习技术的火热发展, 目标检测技术发展的也越来越成熟。目前目标检测技术已应用于自动驾驶、视频监控、医学影像分析、工业自动化以及农业等领域中。在交通标志检测中, 目标检测技术通过分析道路图像, 识别并定位行人、车辆、交通标志等目标, 并通过分析这些信息辅助驾驶员做出正确的决策。但是由于交通标志目标较小、车辆行驶过程中存在遮挡模糊等问题、且交通标志目标易受周围复杂环境影响导致系统对于交通标志

的检测性能不佳。小的交通标志检测是目标检测领域中一项具有挑战性的问题。

传统的交通标志检测算法^[1-4]通常是通过手工特征提取得到目标的形状、颜色以及纹理等信息。然后再利用机器学习方法对检测到的交通标志进行分类和识别。但是由于传统交通标志检测算法依赖手工提取特征, 实时性较差, 相对于深度学习方法来说, 在复杂场景下表现不佳。

随着深度学习的发展, 基于深度学习的交通标志检测算法逐渐成为主流算法。越来越多的学者将焦点集中于如何提升交通标志检测的性能上。

基金项目: 国家自然科学基金 (52074341、51874340); 中央高校基本科研业务费专项资金 (19CX02030A)。

E-mail: s21070040@s.upc.edu.cn

赵宏等人^[5]提出一种基于坐标注意力的轻量级交通标志检测模型 CGS-GhostYOLO, 通过使用 StemBlock 模块进行采样降低模型的参数量, 同时通过引入坐标注意力机制提高模型的特征提取能力。为了解决交通标志检测缺乏底层特征的问题, Zhang 等人^[6]提出一种多尺度注意特征模块来融合不同层次的特征信息, 并引入空间信息聚合器, 使得底层特征映射的空间信息能够融合到高级特征映射中。为了能更好地检测小型交通标志, Gao 等人^[7]通过学习空间注意力图来融合不同尺度的特征图, 并在不同区域强调或抑制特征, 缓解特征之间的不一致性, 增强小目标的特征表示。李嘉豪等人^[8]针对交通标志检测方法在复杂场景下对相似物体存在误检的问题, 设计一种高精度检测模型 LFRD, 设计分层聚类的锚框机制来降低漏检的问题, 同时提出分组损失来指导模型学习相似类别之间的差异, 有效提高了模型的准确率。

由于现实场景中交通标志检测的特殊性, 交通标志检测对于算法检测精度和速度都有着较高的要求。为了能在保证实时检测的同时提高多尺度交通标志的检测精度, Wang 等人^[9]提出了一种改进的特征金字塔模型 AF-FPN, 该模型利用自适应注意模块 (AAM) 和特征增强模块 (FEM) 来减少特征生成过程中的信息丢失。同时提出一种新的自动学习数据增强方法来丰富数据集, 提高模型的鲁棒性。为了解决小目标检测不准确、移动部署困难、工作环境复杂等问题, Wang 等人^[10]提出了一种适用于多种工况小型交通标志的车载自适应交通标志检测器, 该模型设计 DCSP 特征提取模块, 并提出一种具有自适应联合滤波策略的图像增强网络 IENet 显著提高了对小目标的检测精度和速度。Han 等人^[11]针对复杂场景下小交通标志的误检和漏检问题, 提出了一种基于 YOLOv5 的 YOLO-SG 模型, 该模型采用 SPD-Conv 作为下采样结构来减少下采样过程中特征信息的损失。采用 GhostNet 重新设计特征提取架构, 有效减少了模型的参数数量。这些深度学习方法虽然可以有效检测小目标, 但现实场景中道路交通标志的检测精度和检测速度仍受到环境、天气

等因素的影响, 交通标志检测性能仍有待提升。

为了能更好地提升模型对小型交通标志的检测性能, 本文基于 YOLOv5s 7.0 模型进行改进提出了 DMC-YOLO 模型, 能在保证实时检测的情况下明显提高对交通标志的检测精度。本文的主要贡献如下:

- 1) 设计了一种新的特征提取网络, 引入大尺度特征进行特征融合, 并通过多尺度注意增强模块来获得目标的上下文信息, 进一步增强模型对小目标的特征提取能力。
- 2) 为提高对小尺寸交通标志的检测能力, 本文对网络的 Head 检测层进行优化, 增加了包含较多小目标特征信息的大尺度检测层, 同时去除了较小尺度的检测层, 增强模型对小目标的检测性能。
- 3) 改进原网络中的回归损失函数 CIOU, 提出 FEA-CIOU 损失对训练模型进行优化, 加快模型的收敛速度。

1 DMC-YOLO 交通标志检测算法

1.1 基于改进 YOLOv5s7.0 的 DMC-YOLO 模型

YOLOv5 是一个实时高效的目标检测模型, 已经实现了很多场景的落地应用。为了满足边缘设备计算能力和实时性的需求, 本文选用 YOLOv5s7.0 模型作为基准进行实验。通过改进 YOLOv5s7.0 的颈部网络、预测头和损失函数部分, 提高算法对小交通标志的检测性能。

在 YOLOv5s7.0 颈部网络中, 不同尺寸特征之间存在差异与冲突, 直接进行 Concat 融合会造成特征表达不充分。此外, 浅层特征图中包含较多的小目标位置信息, 有助于提高模型对小目标的检测性能。因此本文重新设计模型颈部网络中不同尺度特征的融合方法并引入浅层特征来有意识地引导不同尺度特征的有效融合。改进后的网络称为 DMC-YOLO, 其整体网络结构如图 1 所示。

在模型的颈部网络部分, 在自上而下路径中设计 TD-DAE 模块来融合不同尺度的特征信息, 接着在自下而上路径中设计 DT-AE 模块来融合不同尺度的特征图。TD-DAE 和 DT-AE 模块的详细结构如图 2 所示。

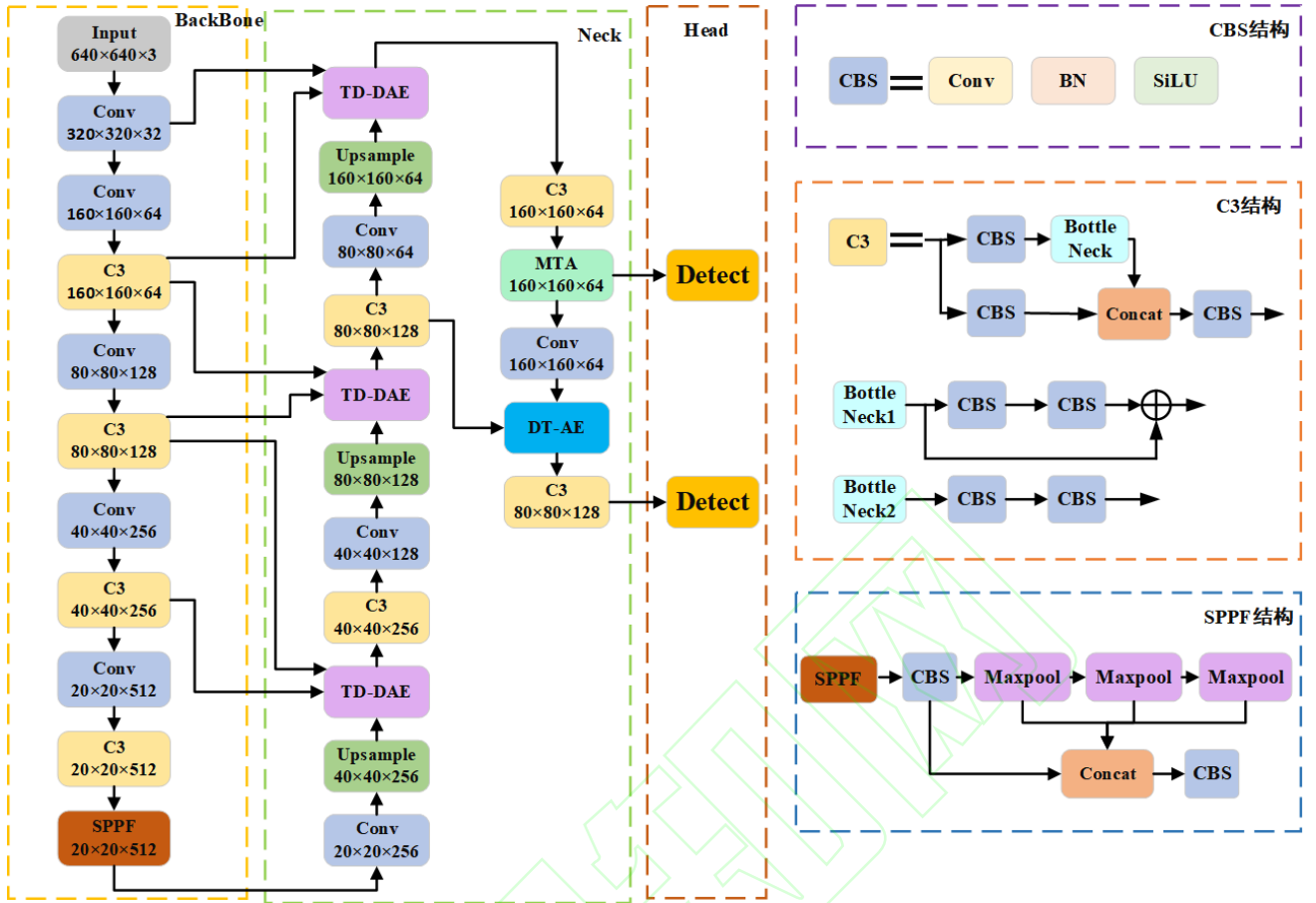


图 1 DMC-YOLO 整体网络结构图

Fig.1 The overall network architecture of DMC-YOLO

为了进一步充分利用浅层特征信息, 本文设计多尺度注意增强模块 MTA 来提高浅层特征图中包含的小目标信息。MTA 的详细结构如图 4 所示。由于在交通标志数据集中目标尺寸较小, 属于小目标的范畴, 因此本文对预测头进行筛选, 选取两个尺寸较大的特征图对目标进行检测, 可以降低网络的计算量和复杂性。在检测头进行回归预测时, 本文对边界框损失函数进行改进提出 FEA-CIOU 损失, 进而提高模型对小目标的敏感度。经过上述改进, 本文提出一种新的交通标志检测算法 DMC-YOLO 模型。

1.2 注意增强的颈部特征提取网络 DMC-PAFPN

通过对模型原始颈部网络进行改进, 提出了一种密集注意增强的颈部特征提取网络。该网络分别在自上而下和自下而上路径中分别设计 TD-DAE 和 DT-AE 模块来融合不同尺度特征的有效融合, 增强网络对小目标的特征提取能力。

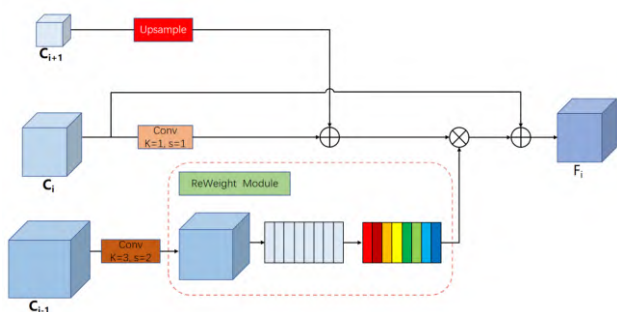
在 TD-DAE 模块中如图 2 (a) 所示, C_i 代表当前层特征图、 C_{i-1} 代表其前一层的浅层特征图、 C_{i+1} 代表其后一层的深层特征图。

在不同尺度特征融合过程中, 首先调整 C_i 的通道数并通过卷积和上采样操作调整 C_{i+1} , 使其与 C_i 尺度一致。这两个特征图相加后生成 D_i 特征, 此操作强化了特征图中的语义信息。接下来, C_{i-1} 经过下采样以匹配 C_i 的尺度, 并通过 ReWeight 通道注意力模块进行处理, 该模块通过全局平均池化、卷积层、ReLU 激活函数、以及 Sigmoid 函数的组合, 学习每个通道的重要性, 生成一组特征权重。这些权重与 D_i 相乘, 赋予不同通道以不同的重要性, 优化了信息的融合过程。最终, 通过将处理后的特征与 C_i 相加, 得到了新的特征图 F_i , 它不仅包含深层特征中的高级语义信息, 也融合了浅层特征中的目标定位信息。

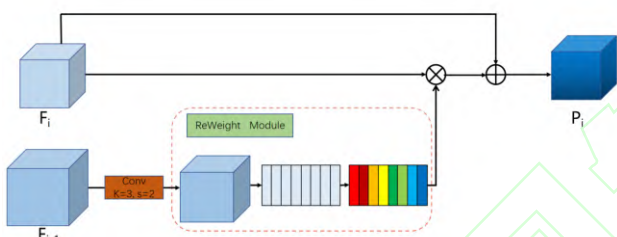
这种自上而下的特征融合策略, 使得 TD-DAE 能够在网络的不同层次间实现信息的有效流动和整合, 生成信息丰富的 F_i 特征集。

DT-AE 的融合方式如图 2 (b) 所示, F_i 表示当前层的特征图, 而 F_{i-1} 是它前一层的特征图。通过使用 3×3 大小、步长为 2 的卷积核对 F_{i-1} 进行下采样, 可以生成一个与 F_i 尺寸相匹配的特征图。接下来,

这个特征图经过 ReWeight 操作(与 TD-DAE 模块的注意力机制相同)处理,产生一组表示不同通道重要性的权重参数。这些权重参数被用来与 F_i 进行相乘,以此提升网络对各通道细节的敏感度。通过残差连接将乘法结果与 F_i 相加,得到增强的特征图 P_i ,此步骤有助于减少在特征融合过程中的信息丢失。通过这种自上而下的注意力增强特征融合方法最终生成了一系列新的信息丰富的特征 P_i 。



(a) 自上而下路径中 TD-DAE 特征融合方式



(b) 自下而上路径中 DT-AE 特征融合方式

图 2 DMC-PAFPN 网络的不同路径的融合方法
Fig.2 Convergence methods for different paths in DMC-PAFPN networks

1.3 MTA 多尺度注意增强模块

经过密集注意增强颈部网络处理之后,得到的特征映射包含更多的底层特征信息。为了进一步提取交通标志的有效信息,本文提出一种多尺度注意增强模块 MTA 如图 4 所示。该模块能够捕捉目标周围的上下文信息,对检测小目标起到辅助增强的作用。

在 MTA 模块中首先对颈部网络输出的最大尺寸特征图采用不同膨胀率的膨胀卷积进行处理,得到具有不同感受野的特征映射如图 3 所示。图 (a)、(b)、(c) 分别代表膨胀率为 1、2、3 的膨胀卷积。假设有一个 $K \times K$ 大小的卷积核,那么卷积核参数之间就有 $K-1$ 个间隔,而膨胀率 r 是指当进行卷积运算时,在每个间隔上增加 $r-1$ 个零值,再进行运算。膨胀卷积可以在不增加网络参数的情况下,增大感受野。

如图 3 所示,当 $r=1$ 时,卷积核参数之间不添加零值,此时为普通卷积,卷积核大小为 3×3 ,感

受野大小为 3×3 。当 $r=2$ 时,在卷积核每个参数间隔中增加了一个零值,此时膨胀卷积就等效于 5×5 的卷积核,等效卷积核大小的计算公式如 (1) 所示,相应的感受野为 7×7 ,感受野计算公式如 (2) 所示。当 $r=3$ 时,就是在卷积核每个参数间隔增加两个零值,此时膨胀卷积就等效于 7×7 卷积,相应的感受野大小为 13×13 。

膨胀卷积等效卷积核大小计算公式如下:

$$K_{DC} = K + (K - 1) \times (r - 1) \quad (1)$$

其中 K_{DC} 代表膨胀卷积等效卷积核大小, K 代表卷积核的大小, r 代表膨胀率的大小。

当前层感受野的计算公式如下所示:

$$RF_i = RF_{i-1} + (K - 1) \times \prod_{i=1}^i \text{Stride} \quad (2)$$

其中 RF_i 代表当前层的感受野, RF_{i-1} 代表前一层的感受野, K 代表卷积核的大小, Stride 代表步长,

$\prod_{i=1}^i \text{Stride}$ 代表之前所有层步长的乘积。

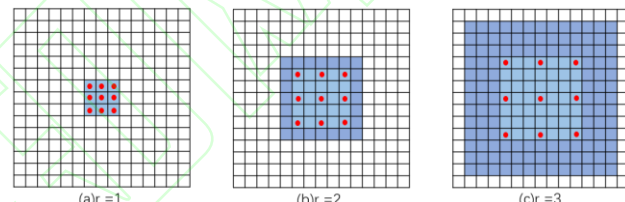


图 3 不同膨胀率的膨胀卷积

Fig.3 Dilated convolutions with different dilated rates

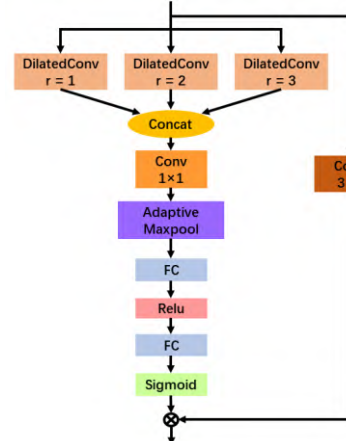


图 4 MTA 多尺度注意增强结构

Fig.4 MTA Multiscale Attention Enhanced Structure

在 MTA 模块中通过不同膨胀率的膨胀卷积处理后,得到了具有不同感受野且输出尺度不变的特征映射。接着将三个特征映射进行 Concat 拼接,然后再通过 1×1 卷积进行融合和降维操作。最后对得到的特征图通过 SE 注意力操作^[12]来锁定候选焦点区域,并过滤掉一些背景和冗余特征。将得到的一组特征权重与最大尺度特征相乘来增强特征中的多尺度上下文信息。

1.4 改进小目标检测层

在 YOLOv5s 原始 Head 检测层中, 选用三种不同尺度检测头对不同大小的交通目标进行检测。每个检测头匹配三个锚点。在交通道路场景中, 交通标志距离驾驶员较远, 可获得的有效信息少, 导致目标检测困难。因此本文对模型的检测头进行重筛选, 引入浅层大尺度检测头对交通标志进行检测。

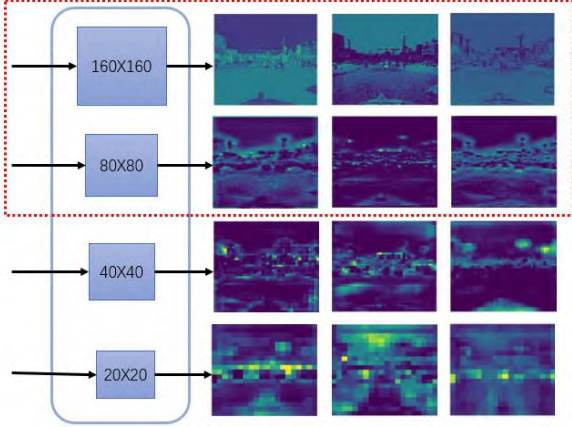


图 5 不同尺度检测层的可视化特征图

Fig.5 Visualized feature maps for different scales of detection layers

在原始 Head 检测层中, 模型通过 80×80 、 40×40 和 20×20 三种不同尺寸的检测层对不同大小的目标进行检测。本文通过统计 TT100K 等交通标志数据集中不同目标的尺寸, 发现在交通标志数据集中绝大部分目标都属于小目标。图 5 展示了在不同层中特征图的可视化结果。从图中可以看出, 浅层大尺度特征层中包含着较多的底层信息, 而随着网络的加深, 深层特征图中小目标信息逐渐退化消失。因此本文引入浅层大尺度特征层, 选用 160×160 、 80×80 两种特征层对目标进行检测。该方法可以有效提高对小目标的检测精度, 并降低模型的参数量和计算量。为了使模型更准确地进行检测, 本文采用 K-means 聚类方法重新定义锚点 (anchor) 的大小, 使得 anchors 编码更具代表性的先验信息, 并在回归时进行更准确地预测。

1.5 FEA-CIOU 损失函数

YOLOv5s7.0 模型的损失函数由三个部分组成, 分别为回归损失 (localization loss)、目标置信度损失 (confidence loss) 以及分类损失 (classification loss)。置信度损失和分类损失都用二值交叉熵损失 (BCE Loss) 进行计算, 而边界框回归损失采用 CIOU 损失进行计算。CIOU 损失的公式如公式 (3) 所示。CIOU 损失考虑了目标真实框和预测框的重叠面积、中心距离和长宽比。但是损失中仅反映长宽

比的差异而不是 w 与 w_{gt} 和 h 与 h_{gt} 之间的实际关系, 减缓了 CIOU 的收敛速度。当预测框与真实框的长宽比非常接近或相等时, CIOU 中长宽比的惩罚项 ϑ 不起作用。且对惩罚项求偏导数可以看出 w 、 h 不能同增同减, 这也会对模型的检测效果造成影响。

CIOU 损失函数的公式如下:

$$CIOU = IOU - \frac{\rho^2}{c^2} - \alpha \vartheta \quad (3)$$

$$\vartheta = \frac{4}{\pi^2} \left(\arctan \frac{w_{gt}}{h_{gt}} - \arctan \frac{w_{pre}}{h_{pre}} \right)^2 \quad (4)$$

$$\alpha = \frac{\vartheta}{1 - IOU + \vartheta} \quad (5)$$

其中 IOU 为交并比损失, ρ 为矩形框的中心点距离, c 为外接矩形框的对角线, ϑ 为长宽比的相似度, α 为 ϑ 的影响因子。

为了解决上述问题, 本文对 CIOU 损失进行改进提出 FEA-CIOU 损失函数, 如公式 (7) 所示。该损失改进了 CIOU 的宽高比惩罚项, 分别计算真实框和预测框的宽和高的差距, 有助于加快模型的收敛速度。另一方面, 当目标越大时, 预测框和真实框之间的中心点距离通常也会越大, 那么在计算损失时大目标损失就会明显高于小目标, 不利于小目标的损失计算。而且在交通标志数据集中存在大量的小目标, 因此针对小目标定位的问题, 本文进一步改进了损失中的惩罚项, 将惩罚项中的 \arctan 函数替换成 \tanh 函数对宽高距离进行计算如式 (8) 所示, 可以提高模型对小目标框的敏感度。最后为了解决数据集中高低质量样本类别不平衡的问题, 采用 Focal Loss 进一步对损失函数进行处理, 从而把高质量锚框和低质量锚框分开。

Focal Loss 函数公式如下:

$$FocalLoss = -\alpha_t (1 - P_t)^\gamma \log(P_t) \quad (6)$$

其中 α_t 用于解决正负样本不平衡的问题, γ 用于解决难易样本不平衡的问题。

改进后的 FEA-CIOU 的公式如式 (7) 所示:

$$FEA-CIOU = -(1 - IOU)^\gamma \log(IOU) (IOU - \frac{\rho^2}{c^2} - \alpha_1 \vartheta_1) \quad (7)$$

$$\vartheta_1 = \frac{4}{\pi^2} \left(\tanh \frac{\rho^2 (w_{gt} - w_{pre})}{c_w^2} - \tanh \frac{\rho^2 (h_{gt} - h_{pre})}{c_h^2} \right) \quad (8)$$

$$\alpha_1 = \frac{\vartheta_1}{1 - IOU + \vartheta_1} \quad (9)$$

其中 ϑ_1 和 α_1 代表改进后的值, 其他参数的含义同上。

2 实验结果及分析

2.1 交通标志数据集

本文在两个开源的中国交通标志数据集 TT100K^[13]和 CCTSDB2021^[14]数据集上进行实验。

下面是对两个数据集的具体介绍。

TT100K 数据集是由清华和腾讯的联合实验室进行整理和公布的。该数据集中共有 10 万张图像,其中有一万多张包含交通标志,总共有 30000 个交通标志实例。TT100K 数据集中的图片拍摄于中国多个城市,而且在不同拍摄地点的光照条件和天气条件有所不同。由于原始数据集中包含较多无目标图片以及实例数较少的类别。因此为了避免数据类别不均衡,本文对图片进行筛选,选择类别中实例数大于 100 的 45 个类别进行实验。与其他交通标志数据集相比,TT100K 数据集的分类更精细,目标更小,因此检测难度较大。

CCTSDB2021 数据集是在 CCTSDB2017 数据集的基础上增加了 4000 多张真实交通场景图,并且将原数据集中的易检测样本换成困难样本整理而成。CCTSDB2021 数据集中一共包含 17856 张图片,其中训练集中有 16356 张图片,测试集中有 1500 张图片。数据集中不仅包括高速公路、城镇等多种路况,还包含雨、雪、雾、夜晚等多种复杂天气和远程拍摄画面。因此对目标的检测难度增加。

2.2 实验环境与评价指标

本章所有实验均在 Windows10 操作系统下进行,采用 Pytorch 深度学习框架搭建 DMC-YOLO 模型,在 NVIDIA RTX2080Ti GPU 上进行训练和测试,网络初始参数设置如表 1 所示。

在训练中,本文使用 YOLOv5s7.0 的预训练权重来进行训练,以便减少训练时间。本文采用随机梯度下降算法作为模型的优化器,当数据集为 TT100K 数据集时,BatchSize 的大小设置为 16, Momentum 设置为 0.937,衰减系数设置为 0.0005,训练轮数 Epoch 设为 300 轮,学习率初始大小设置为 0.001, NMS 的阈值设为 0.5,最终完成 300 轮训练后就得到训练好的模型。当数据集为 CCTSDB2021 交通标志数据集时,网络的参数设置除了 BatchSize 和 Epoch 之外都一样。由于 CCTSDB2021 数据量较少,因此模型训练的批次 Batchsize 大小设置为 32,训练轮数设置为 200。

表 1 网络初始参数设置

Tab.1 Initial network parameterization

配置参数	设定值
Batchsize (批次大小)	16
Momentum (动量参数)	0.937
Weight_decay (学习率衰减系数)	0.0005
Epoch (总训练轮数)	300/200
Pre_learningrate (初始学习率)	0.001

衡量一个交通标志检测模型的好坏主要看检测精度和速度两个方面。它们对应着不同的评价指标。常用的评价指标有准确率 P、召回率 R、单个类别准确率 AP、所有类别 AP 的平均值 mAP、参数量 Params、每秒传输帧数 FPS。P、R 的定义如式 (10)、(11) 所示。其中 TP 表示正确预测的样本数量, FP 表示错误预测的样本数量, FN 表示预测错为负样本的样本数量。AP、mAP 的定义如式 (12)、(13) 所示。mAP 的值越大,表示算法的精度越高。而参数代表着模型的复杂度,参数越大表明模型所占用的计算机资源越多。FPS 越大表明模型检测速度越快,实时性越高。

$$P = \frac{TP}{TP + FP} \quad (10)$$

$$R = \frac{TP}{TP + FN} \quad (11)$$

$$AP = \int_0^1 P(R) \quad (12)$$

$$mAP = \frac{\sum_{i=1}^{N_{class}} AP_i}{N_{class}} \quad (13)$$

2.3 在 TT100K 上的实验结果

为了验证本文模型 DMC-YOLO 检测交通标志的优越性,本文在开源数据集 TT100K 上进行实验,与不同模型检测结果进行对比,如表 2 所示。

通过分析表 2 的实验结果可知,在 TT100K 数据集上本文模型 DMC-YOLO 相比于基准模型 YOLOv5s7.0 在精度上 mAP@0.5 提升了 3%,且本文模型的参数量为 5.2M,比基准降低了 27%,且低于绝大部分模型,有利于模型进行部署。本文模型的 FPS 为 30.58 帧/秒,能够满足实时检测的效果。与两阶段模型 Faster R-CNN、Sparse R-CNN 模型相比,本文模型对交通标志小目标具有较好的检测精度,且模型的参数量明显降低,提高了模型的检测

性能。其次,与其他主流的单阶段检测模型相比,DMC-YOLO 模型对小目标的检测精度具有较好的检测效果,在 TT100K 数据集上取得了最佳表现。

表 2 不同方法在 TT100K 数据集上的结果比较
Tab.2 Comparison of results on the TT100K dataset using different methods

Model	mAP@0.5	Params(M)	FPS
Faster R-CNN ^[15]	73.4	143.7	9.3
SSD ^[16]	76.3	26.28	14.1
QueryDet ^[17]	77.1	—	—
Sparse R-CNN ^[18]	82	—	—
YOLOv5s7.0 ^[19]	81.8	7.13	28.25
YOLOv7tiny ^[20]	80.58	6.13	50.51
VATSD ^[10]	82.8	—	—
YOLO_SG ^[11]	75.8	4	131.6
Improved YOLOv5 ^[21]	82	—	30.42
Ours	84.8	5.2	30.58

为了进一步验证 DMC-YOLO 对于交通标志检测的有效性,本文选取 YOLOv5s 模型、YOLOv7-tiny 模型和 DMC-YOLO 模型分别对 TT100K 测试集进行检测,检测的可视化结果对比图如图 6 所示。

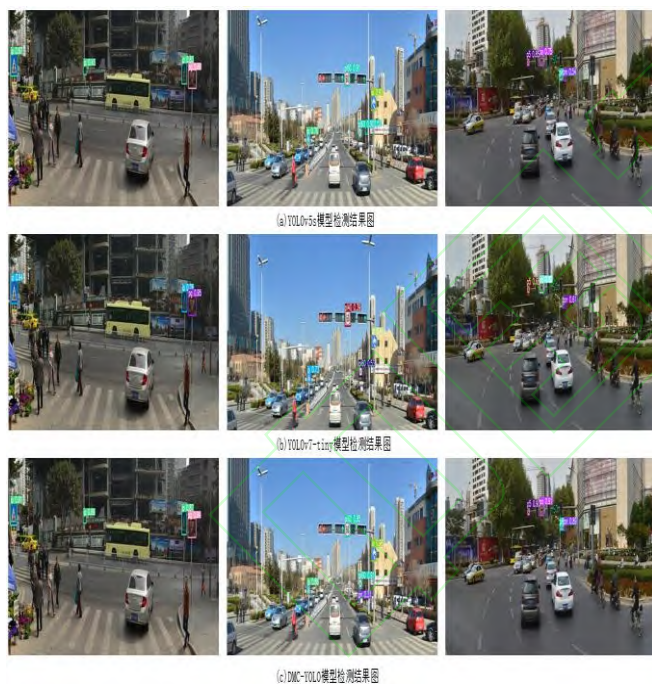


图 6 不同模型的检测结果图

Fig.6 Detection result images from different models

通过对比第一列图片可以发现, YOLOv7-tiny 模型未能将黄色公交车前类别为“ip”的交通标志检测出来,出现了漏检的情况,而本章提出的模型不仅能够检测出图片中所有类别的交通标志且获得了较高的置信度得分,具有更好的检测性能。第二列图片是在光线较强的环境下进行拍摄的,通过分析第二列图片可知, YOLOv5s7.0 模型将黄色楼房上的广告牌误检为交通标志, YOLOv7-tiny 模型将

“pl60”标志误检为“pl40”标志,而本章模型整体表现较好。通过分析第三列图片可以发现,三种模型都能够将道路场景中的交通标志检测出来,但是本章模型对于不同的类别都具有最高的置信度分数。

综合而言,本章模型对于现实道路场景中的交通标志检测性能较好,更适用于交通标志检测任务。

2.4 在 CCTSDB2021 上的实验结果

为了验证本文模型的有效性,本文又在另一个开源交通数据集 CCTSDB2021 上进行实验。在该数据集下与不同主流模型进行比较,实验对比如表 3 所示。

表 3 不同方法在 CCTSDB2021 数据集上的结果
Tab.3 Comparison of results on the CCTSDB2021 dataset using different methods

Model	P	R	mAP@0.5	FPS
Faster R-CNN ^[15]	84.4	54.9	56.58	4.87
SSD ^[16]	86.5	27.7	49.20	22.33
RetinaNet ^[22]	86.7	52.8	57.78	8.88
Sparse R-CNN ^[18]	94.1	52.8	59.65	8.45
YOLOv4 ^[23]	76.2	52.5	51.69	16.55
YOLOv5s7.0 ^[19]	90.1	73.9	80	93.46
YOLOv7-tiny ^[20]	84.9	69.7	76.3	68.49
PSG-YOLOv5 ^[24]	91.6	80.9	80.2	114.35
Improved YOLOv5s ^[25]	—	—	82.8	115.37
Ours	91.9	77.4	83.6	86.96

通过分析表 3 结果可知,本文提出模型 DMC-YOLO 对交通标志的检测精度 mAP 为 83.6%,在所有模型中占有较大优势。在实时性方面,模型每秒检测帧数(FPS)为 86.96 帧/秒,超过了大部分模型的检测速度,能够实现实时检测效果。DMC-YOLO 模型与两阶段模型 Faster R-CNN、Sparse R-CNN 模型相比具有较好的模型检测精度,同时具有较高的检测速度。本文模型与其他主流的目标检测算法相比,具有较好的检测精度,其比基准模型 YOLOv5s7.0 在 mAP 上提升了 3.6%,明显的提升了模型对于交通标志的检测性能,在 CCTSDB2021 数据集上取得了较好的检效果,更适用于交通标志检测任务。

为了能更好的看出模型的检测效果,本文选取 YOLOv5s7.0、YOLOv7-tiny 和 DMC-YOLO 模型对 CCTSDB2021 的测试集进行检测,检测的部分结果图如图 7 所示。

通过图 7 可以看出,通过比较第一列中的三个图片可以发现,这三种模型都能将图片中现有的交通标志检测出来,但是存在部分类别检测置信度分数较低的问题,而本文提出的模型检测置信度分数

都在 80% 以上,对交通标志小目标检测的表现更好。通过分析第二列图片可以看出,YOLOv5s7.0 模型误将右下方物体以极低的置信度分数检测为“prohibitory”标志,而 YOLOv7-tiny 模型虽然将道路场景中的交通标志都检测出来,但是在检测置信度上仍低于本章模型。通过对比这三种模型的检测结果图可知,本章模型相比于其他模型具有较好的检测性能,YOLOv5s7.0 模型将一个背面的交通标志牌误检为一个“mandatory”标志,而 YOLOv7-tiny 模型则漏检了右侧的一个“mandatory”标志。从整体上来说,本章提出的模型具有较强的鲁棒性,能够将复杂环境下的交通标志更准确地检测出来,有助于更好的辅助驾驶,保障道路行车安全。



图 7 不同模型的检测结果图

Fig.7 Detection result images from different models

2.5 消融实验

为了进一步验证本文提出的不同改进方法对于提升模型对交通标志检测的有效性,本文在 TT100K 数据集上对 DMC-YOLO 模型进行消融实验,实验结果如表 4 所示。

表 4 消融实验
Tab.4 Ablation experiment

模型	A	B	C	D	mAP@0.5	GFLOPs
模型 1					81.8	16
模型 2	√				82.9	24.1
模型 3	√	√			83.9	32.3
模型 4	√	√	√		83.9	26.3
模型 5	√	√	√	√	84.8	26.3

表中 A、B、C、D 分别表示在 YOLOv5s7.0 模型的基础上依次改进模型的颈部特征提取网络 DMC-PAFPN、增加多尺度特征增强模块 MTA、改进小目标检测头 two-head 以及优化回归损失函数等操作。

从表 4 中可以看到,YOLOv5s7.0 作为基准模型,其检测精度达到了 81.8%。为了提高模型颈部的特征提取能力,从而提高模型检测精度,我们首先引入了 DMC-PAFPN 网络结构,该结构通过充分利用各阶段生成的不同尺度特征,加强特征间的互联,使得检测精度有了 1.1% 的提高。随后,将 DMC-PAFPN 输出的最大尺度特征作为 MTA 模块的输入,利用 MTA 模块中不同感受野的膨胀卷积模块进行处理,进一步提高模型的特征表达能力,使检测精度再次提升 1.0%。

虽然 DMC-PAFPN 和 MTA 模块的引入有效提升了模型的检测精度,但这也导致了计算成本的增加和检测速度的降低。为了弥补这一缺陷,本文在模型的预测部分采用了 two-head 改进策略,在不降低准确率的情况下,显著的将模型的计算量从 32.3 降低至 26.3。除上述模型内部改进外,本文还针对整体训练过程提出了一种新的 FEA-CIOU 损失函数,在该损失函数的约束下训练,提高了模型对交通标志目标的敏感度,进一步的将检测精度提高了 0.9%。

总的看来,DMC-PAFPN、MTA 模块的加入增强了模型内部的特征提取与表达能力,提高了检测精度,但同时带来了计算量的增加,为了弥补这一不足,在输出层使用 two-head 策略进行改进,显著降低了模型计算量。模型内部的三部分改进在提高模型精度的同时,有效控制了计算成本,模型外部引入的 FEA-CIOU 损失函数对训练过程进行约束,在加快模型收敛的同时进一步提高检测精度。内部与外部的改进策略相互配合,不可或缺的组成了一个协同作用的整体,对于提高模型的综合性能起到了关键作用。

为了能够更加直观地看出本文对模型进行各种改进后的效果,本文在 YOLOv5s7.0 模型的基准上通过改进小目标检测层、改进颈部网络结构并增加 MTA 模块以及优化损失函数后模型的训练精度对比图如图 8 所示。从图 8 中可以看出,通过改进模型的颈部网络、增加 MTA 模块能够有效提升模型对交通标志小目标的检测精度,通过改进模型的小目标检测头以及优化模型的回归损失能够明显地加快模型的收敛速度。

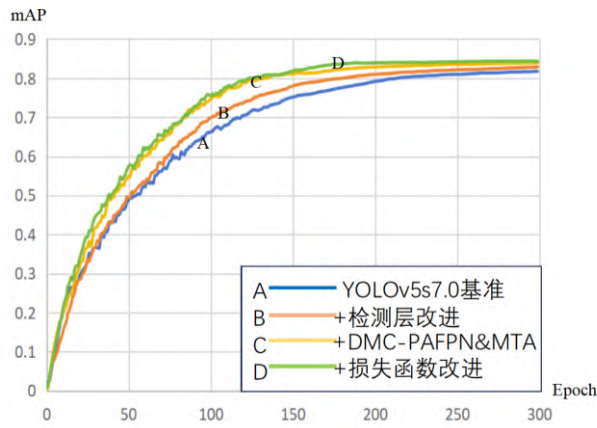


图8 不同改进方法的训练精度对比图

Fig.8 Comparison of training accuracy of different improved methods

综上所述,通过分析不同改进后模型对交通标志的检测结果可以看出,本文提出的几种改进策略能够有效提升网络对于交通标志的检测精度,且模型能够满足实时检测的要求,适用于实时的交通标志检测任务。

3 结束语

为了保证模型能够在实时检测的情况下提高精度,本文提出了一种新的交通标志检测算法DMC-YOLO。该模型采用新的特征提取网络DMC-PAFPN来增强不同尺度特征的有效融合,并利用MTA模块对浅层特征进行增强特征提取。针对检测头进行改进减少了模型的计算量和复杂度。同时提出一种新的损失函数FEA-CIOU提高模型对于小目标的检测效果。在两个开源的数据集上证明了本文模型的优越性。但是由于现实场景中恶劣的天气环境会对图像采集以及目标检测识别产生影响,因此后续工作会在本文改进模型的基础上,结合图像预处理等方法提高模型的鲁棒性。

4 参考文献

- [1] 黄志勇,孙光民,李芳.基于RGB视觉模型的交通标志分割[J].微电子学与计算机,2004,(10):147-148+152.DOI:10.19304/j.cnki.issn1000-7180.2004.10.040.
- [2] Benallal M, Meunier J. Real-time color segmentation of road signs[C]//CCECE 2003-Canadian Conference on Electrical and Computer Engineering. Toward a Caring and Humane

Technology (Cat. No. 03CH37436). IEEE, 2003, 3: 1823-1826.

[3] Barnes N, Loy G, Shaw D. The regular polygon detector[J]. Pattern Recognition, 2010, 43(3):592-602.

[4] Chen Y, Xie Y, Wang Y. Detection and Recognition of Traffic Signs Based on HSV Vision Model and Shape features[J]. J. Comput., 2013, 8(5): 1366-1370.

[5] 赵宏,冯宇博.基于CGS-Ghost YOLO的交通标志检测研究[J].计算机工程,2023,49(12):194-204.DOI:10.19678/j.issn.1000-3428.0066520.

ZHAO H, FENG Y B. Research on Traffic Sign Detection Based on CGS-Ghost YOLO[J]. computer engineering, 2023, 49(12): 194-204. DOI: 10.19678/j.issn.1000-3428.0066520.

[6] Zhang J, Ye Z I, ** X, et al. Real-time traffic sign detection based on multiscale attention and spatial information aggregator[J]. Journal of Real-Time Image Processing, 2022, 19(6): 1155-1167.

[7] Gao E, Huang W, Shi J, et al. Long-tailed traffic sign detection using attentive fusion and hierarchical group softmax[J]. IEEE Transactions on Intelligent Transportation Systems, 2022, 23(12): 24105-24115.

[8] 李嘉豪,闵卫东,陈炯缙等.一种复杂场景下高精度交通标志检测模型[J].计算机工程,2023,49(11):311-320.DOI:10.19678/j.issn.1000-3428.0066372.

LI J H, MIN W D, CHEN J J, et al. A high precision traffic sign detection model in complex scenes [J]. Computer Engineering, 2023, 49 (11): 311-320.

[9] Wang J, Chen Y, Gao M, et al. Improved YOLOv5 network for real-time multi-scale traffic sign detection. ar**v 2021[J]. ar**v preprint ar**v:2112.08782.

[10] Wang J, Chen Y, Ji X, et al. Vehicle-mounted adaptive traffic sign detector for small-sized signs in multiple working conditions[J]. IEEE Transactions on Intelligent Transportation Systems, 2023.

[11] Han Y, Wang F, Wang W, et al. YOLO-SG: Small traffic signs detection method in complex scene[J]. The Journal of Supercomputing, 2023: 1-22.

[12] Hu J, Shen L, Sun G. Squeeze-and-excitation networks[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2018: 7132-7141.

[13] Zhu Z, Liang D, Zhang S, et al. Traffic-sign detection and classification in the wild[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 2110-2118.

[14] Zhang J, Zou X, Kuang L D, et al. CCTSDB 2021: a more comprehensive traffic sign detection benchmark[J]. Human-centric Computing and Information Sciences, 2022, 12.

[15] Ren S, He K, Girshick R, et al. Faster r-cnn: Towards real-time object detection with region

proposal networks[J]. Advances in neural information processing systems, 2015, 28.

[16] Liu W, Anguelov D, Erhan D, et al. Ssd: Single shot multibox detector[C]//Computer Vision – ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11 – 14, 2016, Proceedings, Part I 14. Springer International Publishing, 2016: 21-37.

[17] Yang C, Huang Z, Wang N. QueryDet: Cascaded sparse query for accelerating high-resolution small object detection[C]//Proceedings of the IEEE/CVF Conference on computer vision and pattern recognition. 2022: 13668-13677.

[18] Sun P, Zhang R, Jiang Y, et al. Sparse r-cnn: End-to-end object detection with learnable proposals[C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2021: 14454-14463.

[19] Ultralytics (2020) YOLOv5 2020 Available from: <https://github.com/ultralytics/yolov5>

[20] Wang C Y, Bochkovskiy A, Liao H Y M. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2023: 7464-7475.

[21] Zhang R, Zheng K, Shi P, et al. Traffic Sign Detection Based on the Improved YOLOv5[J]. Applied Sciences, 2023, 13(17): 9748.

[22] Lin T Y, Goyal P, Girshick R, et al. Focal loss for dense object detection[C]//Proceedings of the IEEE international conference on computer vision. 2017: 2980-2988.

[23] Bochkovskiy A, Wang C Y, Liao H Y M. Yolov4: Optimal speed and accuracy of object detection[J]. arXiv preprint arXiv:2004.10934, 2020.

[24] Hu J, Wang Z, Chang M, et al. PSG-Yolov5: A Paradigm for Traffic Sign Detection and Recognition Algorithm Based on Deep Learning[J]. Symmetry, 2022, 14(11): 2262.

[25] Qu S, Yang X, Zhou H, et al. Improved YOLOv5-based for small traffic sign detection under complex weather[J]. Scientific reports, 2023, 13(1): 16219.