

光电工程

Opto-Electronic Engineering

中文核心期刊 中国科技核心期刊
Scopus CSCD

基于改进YOLOv5的交通标志识别方法

曲立国, 张鑫, 卢自宝, 刘玉玲, 陈国豪

引用本文:

曲立国, 张鑫, 卢自宝, 等. 基于改进YOLOv5的交通标志识别方法[J]. 光电工程, 2024, 51(6): 240055.

Qu L G, Zhang X, Lu Z B, et al. A traffic sign recognition method based on improved YOLOv5[J]. *Opto-Electron Eng*, 2024, 51(6): 240055.

<https://doi.org/10.12086/oe.2024.240055>

收稿日期: 2024-03-07; 修改日期: 2024-06-03; 录用日期: 2024-06-04

相关论文

基于改进YOLOv5s的无人机图像实时目标检测

陈旭, 彭冬亮, 谷雨

光电工程 2022, 49(3): 210372 doi: 10.12086/oe.2022.210372

基于ZYNQ的轻量化YOLOv5声呐图像目标检测算法及实现

赵冬冬, 谢墩翰, 陈朋, 梁荣华, 沈伊, 郭新新

光电工程 2024, 51(1): 230284 doi: 10.12086/oe.2024.230284

PAW-YOLOv7: 河道微小漂浮物检测算法

栾庆磊, 常昕昱, 吴叶, 邓从龙, 史艳琼, 陈梓华

光电工程 2024, 51(4): 240025 doi: 10.12086/oe.2024.240025

改进YOLOv7的无人机视角下复杂环境目标检测算法

张润梅, 肖钰霏, 贾振楠, 陈中, 陈梓华, 袁彬, 曹炜威, 宋妮妮

光电工程 2024, 51(5): 240051 doi: 10.12086/oe.2024.240051

更多相关论文见光电期刊集群网站 



<http://cn.ojournal.org/oe>



OE_Journal



Website

DOI: 10.12086/oe.2024.240055

基于改进 YOLOv5 的交通标志识别方法

曲立国^{1,2*}, 张鑫¹, 卢自宝¹, 刘玉玲¹, 陈国豪³¹安徽师范大学物理与电子信息学院, 安徽 芜湖 241002;²安徽省智能机器人信息融合与控制工程研究中心, 安徽 芜湖 241002;³湖北省武汉市武汉铭科轨道交通装备有限公司, 湖北 武汉 430074

摘要: 交通标志检测是自动驾驶领域重要的环节, 针对当前交通标志的识别存在漏检、误检、模型参数多, 以及常见的代表性真实环境情况, 如雾天鲁棒性差的问题, 提出一种改进 YOLOv5 的小目标交通标志识别算法。首先对数据集进行雾化操作以适应在雾天情况下的准确识别, 使用更加轻量的部分卷积 (partial convolution, PConv) 构建 PC3 特征提取模块; 随后在颈部网络中提出延伸的特征金字塔 (extended feature pyramid network, EFPN), 为小目标添加一个小目标检测头, 同时删去原始颈部网络中针对大目标的检测头, 提高小目标识别准确率的同时降低网络参数; 最后引入 Focal-EIOU 替换 CIOU 作为损失函数, 以此来解决小目标的误检和漏检问题, 嵌入 CBAM 注意力机制, 提升网络模型的特征提取能力。改进的模型性能在 TT100K 数据集上得到验证, 与原 YOLOv5 算法相比, 改进模型在精确率 (P)、mAP0.5 上分别提高了 8.9%、4.4%, 参数量降低了 44.4%, 在 NVIDIA 3080 设备上 FPS 值为 151.5, 可满足真实场景中交通标志的实时检测。

关键词: 小目标检测; 深度学习; YOLOv5; 轻量化卷积; EFPN; 损失函数**中图分类号:** TP391**文献标志码:** A

曲立国, 张鑫, 卢自宝, 等. 基于改进 YOLOv5 的交通标志识别方法 [J]. 光电工程, 2024, 51(6): 240055

Qu L G, Zhang X, Lu Z B, et al. A traffic sign recognition method based on improved YOLOv5[J]. *Opto-Electron Eng*, 2024, 51(6): 240055

A traffic sign recognition method based on improved YOLOv5

Qu Ligu^{1,2*}, Zhang Xin¹, Lu Zibao¹, Liu Yuling¹, Chen Guohao³¹School of Physics and Electronic Information, Anhui Normal University, Wuhu, Anhui 241002, China;²Anhui Provincial Engineering Research Center for Information Fusion and Control of Intelligent Robots, Wuhu, Anhui 241002, China;³Wuhan Mingke Rail Transit Equipment Co., Ltd., Wuhan, Hubei 430074, China

Abstract: Traffic sign detection is an important link in the field of autonomous driving, and given the problems of missed detections, false detections, many model parameters, and common and complex representative real environment conditions, such as poor robustness in foggy days, an improved YOLOv5 micro-target traffic sign recognition algorithm was proposed. Firstly, the dataset was atomized to adapt to the accurate identification in the

收稿日期: 2024-03-07; 修回日期: 2024-06-03; 录用日期: 2024-06-04

基金项目: 芜湖市科技计划资助项目 (2022jc07); 安徽省教育厅研究生创新创业实践资助项目 (2022cxycysj058, 2023cxycysj043)

*通信作者: 曲立国, lgqu@ahnu.edu.cn

版权所有©2024 中国科学院光电技术研究所

foggy weather, and the PC3 feature extraction module was constructed by using a lighter partial convolution (PConv), and then the Extended Feature Pyramid Network (EFPN) was proposed in the neck network. Finally, Focal-EIOU is introduced to replace CIOU as the loss function to solve the problem of false detection and missed detection of micro targets, and the CBAM attention mechanism is embedded to realize the lightweight model and significantly improves the feature extraction ability of the network model. Compared with the original YOLOv5 algorithm, the improved model is increased by 8.9% and 4.4% respectively on P and mAP0.5, the number of parameters is reduced by 44.4%, and the FPS value on NVIDIA 3080 device is 151.5, which can meet the real-time detection of traffic signs in the real scenes.

Keywords: small object detection; deep learning; YOLOv5; lightweight convolution; EFPN; loss function

1 引言

交通标志识别 (Traffic sign recognition, TSR) 是汽车自动驾驶系统^[1]和驾驶辅助系统领域的一个重要研究内容。目前交通路况复杂, 警力有限, 交通事故频发, 智能交通管理系统 (intelligent traffic management system, ITMS) 有望借助人工智能技术来解决这一问题, 辅助驾驶员实时对路况信息做出正确响应, 提高驾驶安全性。交通标志识别作为 ITMS 系统的重要组成部分, 众多学者开展了大量研究工作, 传统算法主要依靠人工特征提取, 如局部二值模式^[2](local binary pattern, LBP), Gabor^[3]、方向梯度直方图^[4](histogram of oriented gradient, HOG) 等, 并使用支持向量机^[5](support vector machines, SVM)、AdaBoost^[6]等分类器完成交通标志识别。由于车载相机远距离拍摄图像, 目标图像占比较小, 且要满足实时检测, 人工特征提取并不能满足实际需要, 目前还没有一种全面的方法来识别小目标交通标志的最佳特征集和最佳分类器。

近年来, 基于卷积神经网络的深度学习模型在目标检测领域取得了显著成果。深度学习中涵盖两种类型的目标检测模型: 一种是区域建议的两阶段检测算法, 如区域卷积神经网络 R-CNN^[7]、Fast R-CNN^[8]、Faster R-CNN^[9]等; 另一种是单阶段检测算法, 如 YOLO、SSD^[10], 它们可以直接从网络中提取特征来预测目标的分类和位置。YOLO 的最早版本是由先驱 Redmon 提出, 随后的改进中, 许多研究人员在物体检测方面进行广泛研究, 并将其集成到 YOLO 算法中, 产生了很多高效的检测模型, 如 YOLOv4^[11]、YOLOv5^[12]、YOLOX^[13]、YOLOv6^[14]、YOLOv7^[15]和 YOLOv8^[16]。相比于其它深度学习算法, 单阶段检测算法 YOLO 最大的特点是运行速度快, 其中 YOLOv5 算法在识别精度和速度方面实现最佳均衡, 可以很好

地满足实时检测需求。

针对小目标交通标志检测中的识别精度低、漏检率高的问题, 很多深度学习学者进行了有效的改进并取得了一定的成效^[17]。为了提高骨干网络对交通标志的提取能力, Yang 等人^[18]在网络结构中增加跨层连接, 缩短特征金字塔 (feature pyramid network, FPN) 网络中的信息传递路径, 融合更多的特征, 提高网络特征识别精度。由于小目标尺寸较小, 在传感器中成像清晰度不高^[19], 为了提高复杂背景下小目标的检测和识别性能, Zhang 等人^[20]对小目标交通标志识别, 提出级联 R-CNN 获得金字塔的多尺度特征, 使用点卷积和 Softmax 加权多尺度特征, 此方法提高了交通标志检测的准确性。Zhang 等人^[21]在其文章中提出 MSA_YOLOv3, 为使小目标交通标志得到准确定位和分类, 使用混合图像增强, 并引入多尺度空间金字塔模块来学习更丰富的特征, 该算法在检测小目标交通标志的精度优于 YOLOv3。Guo 等人^[22]提出了一种基于卷积神经网络 (CNN) 图像增强的跨域自适应雾天物体检测方法, 提高物体检测算法的准确性。Lin 等人^[23]基于 YOLOv4 的模型结构, 增加了一个 160×160 的小目标检测层, 同时删去了 20×20 的大目标检测层, 并将 ECA (efficient channel attention) 机制无缝插入到主干结构中, 提升了模型性能。Wang 等人^[24]在检测网络中增加了雾浓度判别模块, 以提高其适应性和鲁棒性。Lang 等人^[25]将 CA (coordinate attention) 机制嵌入到 YOLOv5 模型中, 提升了针对小目标对象的检测性能。Zhu 等人^[26]改进了颈部结构, 使用 BiFPN (bidirectional feature pyramid network) 来丰富特征的提取, 降低了模型漏检率。

通过上述分析, 针对小目标识别精度低、漏检率高, 以及雾天环境由于能见度低导致鲁棒性差等问题, 本文提出一种改进 YOLOv5 的小目标交通标志识别算法方法, 其主要工作如下: 1) 对数据样本进行数据

增强处理; 2) 使用更加轻量的 PConv 构建 PC3 特征提取模块; 3) 采用延伸的特征金字塔 (extended feature pyramid network, EFPN) 为小目标添加一个小目标检测头; 4) 删去原始针对大目标的检测头; 5) 引入 Focal-IoU 替换 CIoU 作为损失函数; 6) 嵌入 CBAM 注意力机制。经实验验证, 本文算法针对小目标漏检、误检以及雾天环境鲁棒性差的问题有了一定的改善。

2 相关工作

2.1 YOLOv5 目标检测模型

作为一种目标检测算法, YOLOv5 根据其不同的网络模型深度和特征宽度分为多个版本, 包括 n、s、m、l 和 x。网络的深度和宽度可以根据任务需求自由更改。YOLOv5 网络结构如图 1 所示, 由四个部分组成: 输入、骨干、颈部和头部。

输入端的图片首先被缩放到模型训练或推理时所需的大小, YOLOv5 提供了自动聚类^[27]合适锚框的方法, 通过聚类, 模型可以选择最适合数据集中目标大小的锚框, 提高预测的准确性; 骨干网络负责从输入端的图片中提取特征图片; 经骨干网络特征提取之后, 特征图进入到由特征金字塔网络 (feature pyramid network, FPN) 和多路径信息融合网络 (path aggregation network, PANet) 组成的颈部网络, 颈部网络位于骨干网络和头部之间, 负责进一步处理骨干网络提取的特征, 用于后续的预测, FPN 将骨干网络提取的特征进

行多尺度融合, 使得模型能够同时检测不同大小的目标。它通过自顶向下的路径将高层语义信息传递给低层, 增强低层特征图的语义信息, PANet 在 FPN 的基础上增加了一个自底向上的路径, 将低层特征图中的细节信息传递给高层, 从而进一步提高了特征图的融合能力; 经颈部网络特征融合之后, 特征图最后被送入到头部网络进行识别分类, YOLOv5 设有 3 个检测头, 分别针对小目标、中目标、大目标进行检测。

2.2 YOLOv5 网络模型优化

2.2.1 数据增强

本文所使用的 TT100K 数据集, 其样本多是在理想情况拍摄的, 存在样本数据不均衡问题, 采用数据增强丰富样本, 提高训练后模型的泛化能力。本文采用 ImageNet 分类, 通过对原始图像进行变换, 可以生成更多的训练样本, 从而使模型更好地学习图像的各种变化。另外在真实环境中天气是多变的, 包括晴天、阴天、雨天、雪天、雾天等, 但在驾驶过程中, 雾的影响相较于其它天气更容易影响驾驶员的视线, 所以本文主要针对雾天情况增加数据集。

采用合成雾算法对数据集进行增强, 从而提高模型在雾场景下的鲁棒性。暗通道先验可以描述为式 (1):

$$J^{\text{dark}}(x) = \left| \min_{y \in \Omega(x)} \left(\min_{c \in \{r, g, b\}} J^{c(y)} \right) \right| \rightarrow 0, \quad (1)$$

其中: J^{dark} 表示经过暗通道处理之后输出的灰度图像,

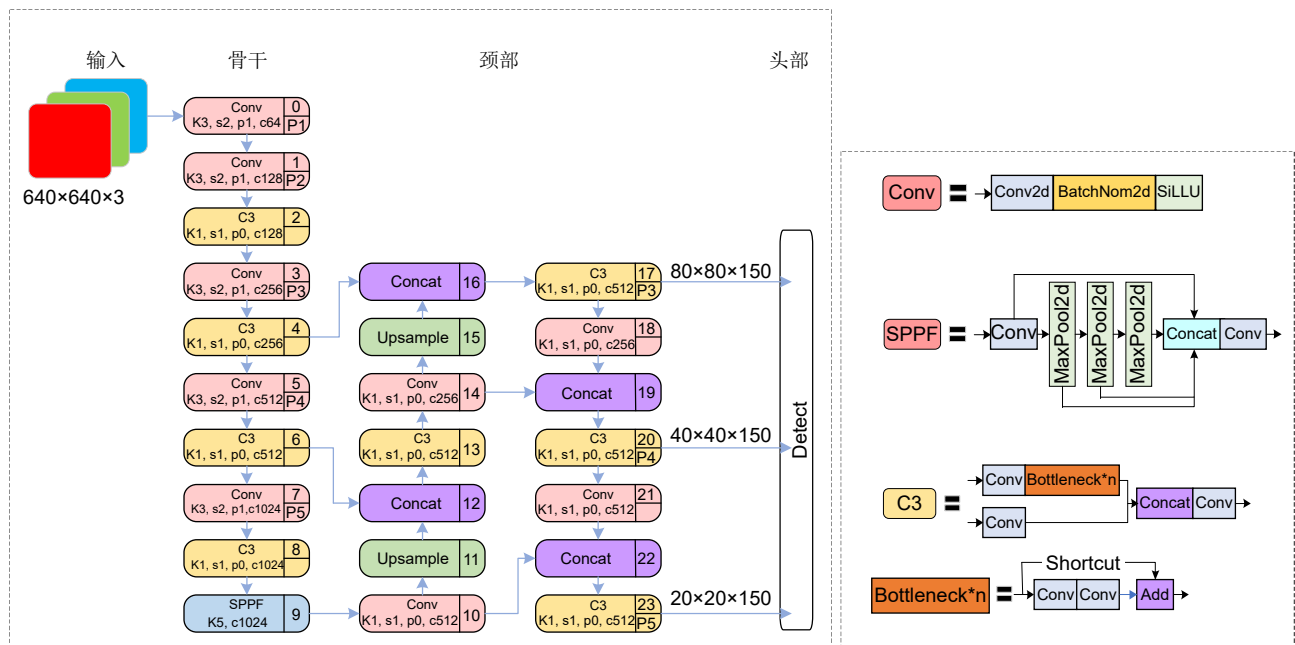


图 1 YOLOv5 网络结构

Fig. 1 YOLOv5 network structure

J 表示 RGB 输入的三通道图像。暗通道先验显示, 在彩色图像的非天空区域总是至少有一个通道具有非常低的像素值。基于暗通道先验, 利用雾图像形成合成雾模型, 如式 (2) 所示^[28]:

$$I(x) = J(x)t(x) + A(1 - t(x)), \quad (2)$$

式中: $J(x)$ 为原始无雾图像, A 为大气光, $t(x)$ 为透射率。 A 的范围为 0~1, 表示所添加的雾的灰度。当 $A=1$ 时加入纯白雾, 当 $A=0$ 时加入纯黑雾。 $t(x)$ 的范围为 0~1, 表示原始图像和雾化输出图像中的比例。当 $t(x)=1$ 时, 所有的场景都可以透过雾, 即没有雾; 当 $t(x)=0$ 时, 没有任何场景可以通过纯白雾, 即输出为纯白雾。不同透光率 $t(x)$ 的雾化图像如图 2 所示。



图 2 雾化图像
Fig. 2 Fog image

2.2.2 轻量级卷积神经网络

在 YOLOv5 的骨干部位, C3 部分包含 3 个 Conv 和 1 个 Bottleneck, 两个结果进行叠加之后再进行一次 Conv, 虽然这种连接方式可以最大化传播梯度差异, 但普通卷积的稠密卷积操作会带来速度的延迟。为了解决此问题, 本文借助轻量化部分卷积 PConv 替换掉原 C3 中的普通卷积, 替换后的模块被称为 PC3, 两种结构图分别如图 3 所示。

PC3 模块用于特征提取的模块, 通过堆叠多个卷积层和改变其通道数来提取不同尺度和复杂性的特征。有两路分支, 左边分支卷积核大小为 1×1 , 用于降低输入特征图的通道数, 以减少计算量, 右侧分支是进行特征提取的核心, 卷积核大小为 3×3 , 通过增大卷积核的尺寸来增加感受野, 最后再使用 1×1 卷积来进

一步提取高级特征。此设计在提高网络的检测能力和准确性的同时使得结构轻量化。

PConv 的计算如式 (3) 所示:

$$x' = \begin{cases} W^T(X \cdot M) \frac{\text{sum}(1)}{\text{sum}(M)} + b, & \text{if } \text{sum}(M) > 0 \\ 0, & \text{otherwise} \end{cases}, \quad (3)$$

其中: X 为输入的特征层, W 为卷积核, b 为卷积运算的偏置, 输出的 x' 是单个像素。 M 为输入的掩膜, 与输入特征层 X 具有相同的维度的二值 (0, 1) 矩阵。 $\text{sum}(1)$ 中 1 是指与卷积核 W 维度相同元素全为 1 的矩阵。 $\text{sum}()$ 函数代表对二值 (0, 1) 矩阵元素求和运算。

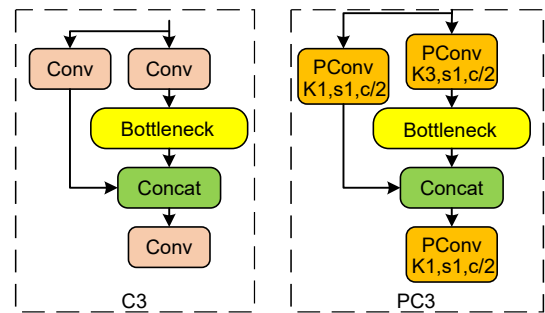


图 3 C3 和 PC3 结构图
Fig. 3 Structure diagram of C3 and PC3

如图 4 所示, PConv 在不影响模型性能的同时克服了对计算资源的需求。因此, 使用 PC3 模块替换原 C3 结构, 使整个模型更加轻量。

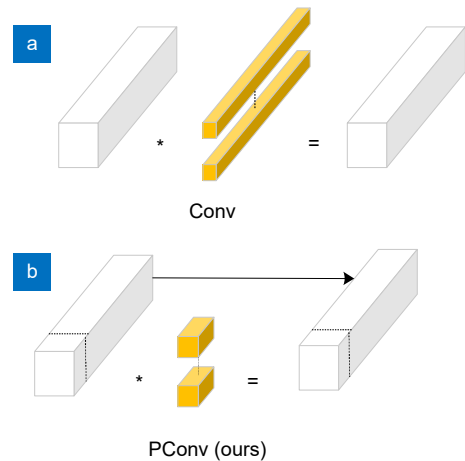


图 4 Conv 和 PConv
Fig. 4 Conv and PConv

2.2.3 延伸的特征金字塔 EFPN

本文根据文献 [29] 小目标的定义: 目标锚框的宽高与原图宽高比小于 0.1。根据 TT100K 数据集的相

对大小绘制了训练集标签宽高比分布图, 如图 5 所示, 横纵坐标值代表标签宽高占整张图像比例, 其中, 训练集中目标的高宽比主要集中在 0.1 以内, 比值在 0.05 以内的占比也很大, 由此得知交通标志符合小目标特征。

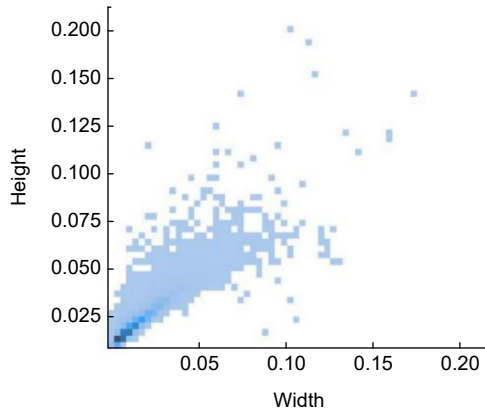


图 5 标签宽高比分布图
Fig. 5 Distribution of label aspect ratio

虽然不同金字塔层次的特征负责不同大小的物体, 但在 TT100K 数据集中交通标志基本都是小目标, 如果使用 YOLOv5 附带的三个检测头, 会使得小目标和中目标检测耦合在同一底层 P3'上, 如图 6(a) 中 YOLOv5 默认路径中黄色圆圈所示。为了缓解这一问题, 建议将普通的特征金字塔延伸到一个新的层次形成延伸的特征金字塔 EFPN, EFPN 的结构如图 6(b) 所示, 其中红色圆圈 P2'层即为新增加的小目标检测层, 该检测层支持小目标交通标志检测, 此改进能更多识别遗漏的小目标交通标志。

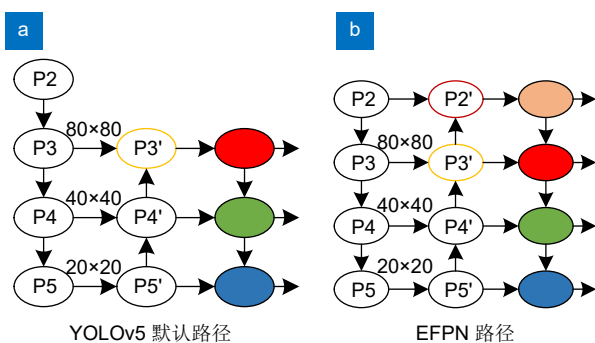


图 6 YOLOv5 默认路径和 EFPN 路径
Fig. 6 YOLOv5 default path and EFPN path

2.2.4 删去大目标检测层 EFPN

经图 5 分析, TT100K 数据集中的标签大小占比较小, 所以 YOLOv5 设置的大目标检测功能基本闲

置, 基于此分析, 为了减少模型参数量和计算量, 删去 EFPN 的 P5'层 20×20 大目标检测层, 定义为 EFPN' 路径, 如图 7 所示, P2'、P3'、P4'将被输入右边检测器, 以实现进一步的对象定位和分类, 实现对小尺寸目标的更优检测。

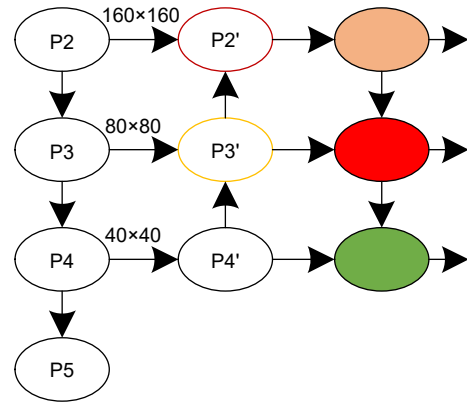


图 7 删去大目标检测层路径 EFPN'
Fig. 7 Removing the path EFPN' from the large object detection layer

针对图 6(b) 结构 (EFPN) 和图 7 结构 (EFPN') 进行了实验对比。由表 1 可知, 删除大目标检测头后, 实验结果证明模型参数量下降 23.4%, 特别针对于数据集中相对微小的目标, 例如 pn (禁止停车)、w32 (施工)、ph4 (限高 4 m) 的检测精确率都有进一步的提升, 而且检测过程中无检测丢失情况。

表 1 EFPN 与 EFPN' 结构对比
Table 1 Comparison of EFPN and EFPN' structures

改进方法	精确率 P/%			Params/M
	pn	w32	ph4	
YOLOv5+PC3+EFPN	0.93	0.70	0.76	4.96
YOLOv5+PC3+ EFPN'	0.96	0.96	0.98	3.80

模型说明: YOLOv5+PC3+EFPN 表示 YOLOv5 中 C3 中的普通卷积用轻量化部分卷积 PConv 替换, 构成 PC3 替换掉原 C3 结构, 检测头用 EFPN 替换; YOLOv5+PC3+EFPN' 表示 YOLOv5 中 C3 中的普通卷积用轻量化部分卷积 PConv 替换, 构成 PC3 替换掉原 C3 结构, 检测头用 EFPN' 替换。

当特征融合网络发生变化时, 原 YOLOv5 默认锚框大小如表 2 所示, 不适用于新的特征空间, 因此本文采用基于 TT100K 数据集的 K 均值聚类算法来重新计算和调整锚框的尺寸和比例, 聚类结果如表 3 所示。

对比表 2、表 3 锚框大小, 表 3 中通过 K 均值聚类算法获得的新锚框能更好地适应本文所研究的小目

标尺寸, 更符合小目标交通标志识别的实际情况, 从而提高小目标检测的准确率。

表 2 YOLOv5 默认锚框大小
Table 2 YOLOv5 default anchor box size

检测尺度	Anchor1	Anchor2	Anchor3
小尺寸	[10,13]	[16,30]	[33,23]
中尺寸	[30,61]	[62,45]	[59,119]
大尺寸	[116,90]	[156,198]	[373,326]

表 3 K 均值聚类算法结果
Table 3 Results of K-means clustering algorithm

检测尺度	Anchor1	Anchor2	Anchor3
小尺寸	[5,5]	[6,7]	[8,9]
中尺寸	[9,14]	[9,14]	[14,15]
大尺寸	[19,20]	[19,20]	[25,26]

2.2.5 损失函数的改进

虽然完全交并比 (complete intersection over union, CIoU) 考虑了边界框的中心点距离和宽高比一致性, 没有特别设计来解决类别不平衡问题, 也没有特别针对小目标检测进行优化。交通标志中包含大量小尺寸目标, 更难检测, 需要特别关注。CIoU 通常作为边界框回归的损失函数使用, 而类别不平衡问题则需要额外的损失函数 (如交叉熵损失) 来解决。损失函数 Focal-EIoU^[30] 通过结合 Focal-Loss, 旨在用单一损失函数同时解决边界框回归和小目标检测问题。基于此, 本文采用 Focal-EIoU 取代 YOLOv5 默认函数 CIoU 提升算法对多分类的识别精确度和鲁棒性。Focal-EIoU 的公式如式 (4)、式 (5) 所示:

$$L_{EIoU} = L_{IoU} + L_{dis} + L_{asp} = 1 - IoU + \frac{\rho^2(b, b^{gt})}{(w^c)^2 + (h^c)^2} + \frac{\rho^2(w, w^{gt})}{(w^c)^2} + \frac{\rho^2(h, h^{gt})}{(h^c)^2}, \quad (4)$$

$$L_{Focal-EIoU} = IoU^\gamma L_{EIoU}, \quad (5)$$

式中: γ 是超参数, 作用是控制所抑制的程度, 两个边界框的最小闭合区的宽度和高度在公式中用 w^c 和 h^c 表示; IoU 表示交并比, b 、 w 、 h 分别为预测框的中心点、宽度和高度; b^{gt} 、 w^{gt} 、 h^{gt} 分别为实际框的中心点、宽度和高度; $\rho^2(b, b^{gt})$ 、 $\rho^2(w, w^{gt})$ 、 $\rho^2(h, h^{gt})$ 分别为预测框与真实框之间的中心点、宽度和高度之间的欧式距离。

2.2.6 注意力机制

CBAM^[31] 由两个独立的子模块组成, 即通道注意力模块 (channel attention module, CAM) 和空间注意力模块 (spatial attention module, SAM), 它们在通道和空间维度上执行注意力操作。CBAM 使模型能够关注图像中更重要的区域, 进一步提升特征的空间表达能力。本部分在模型主干网络最后一个 C3 模块之后嵌入 CBAM 空间和通道注意力模块。CBAM 设计为即插即用模块集成到现有的 CNN 架构中, 作为一个轻量级的模块不会显著增加计算负担, 其结构如图 8 所示。

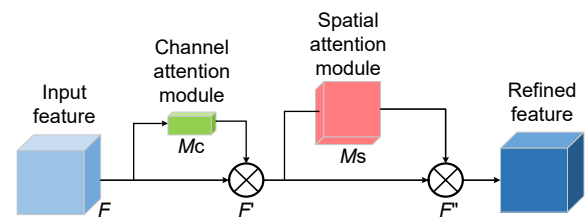


图 8 CBAM 注意力机制
Fig. 8 CBAM attention mechanism

CAM 对输入特征图 F 中的各个通道进行自适应的缩放, 提炼出更加有区分性的特征表示 F' ; SAM 进一步对特征图 F' 的空间维度进行聚焦, 强化模型对图像中关键区域的感知能力, 输出特征图 F'' 。算法过程如式 (6) 所示:

$$F'' = M_s(F') \otimes F' = M_s(M_c(F) \otimes F) \otimes (M_c(F) \otimes F), \quad (6)$$

式中: $F \in \mathbb{R}^{C \times H \times W}$ 为输入特征图, $M_c \in \mathbb{R}^{C \times 1 \times 1}$ 和 $M_s \in \mathbb{R}^{1 \times H \times W}$ 分别表示通道注意力特征和空间注意力特征, \otimes 表示矩阵元素依次相乘, F'' 为输出特征图。通道注意力模块的计算过程如式 (7) 所示:

$$M_c = \sigma(MLP(AvgPool(F)) + MLP(MaxPool(F))) = \sigma(W_1(W_0(F_{avg}^c)) + W_1(W_0(F_{max}^c))), \quad (7)$$

其中: σ 表示 Sigmoid 函数, MLP 是一个双层共享的人工神经网络, $F_{max}^c \in \mathbb{R}^{C \times 1 \times 1}$ 和 $F_{avg}^c \in \mathbb{R}^{C \times 1 \times 1}$ 分别表示在最大池化和平均池化操作之后获得的特征。空间注意力模块的计算过程如式 (8) 所示:

$$M_s(F) = \sigma(f^{7 \times 7}([AvgPool(F); MaxPool(F)])) = \sigma(f^{7 \times 7}([F_{avg}^s; F_{max}^s])), \quad (8)$$

式中: $F_{max}^s \in \mathbb{R}^{1 \times H \times W}$ 和 $F_{avg}^s \in \mathbb{R}^{1 \times H \times W}$ 表示使用两种池化操作得到的二维空间映射, $f^{7 \times 7}$ 表示卷积核为 7×7 的卷积操作。

改进后的网络整体结构如图 9 所示。

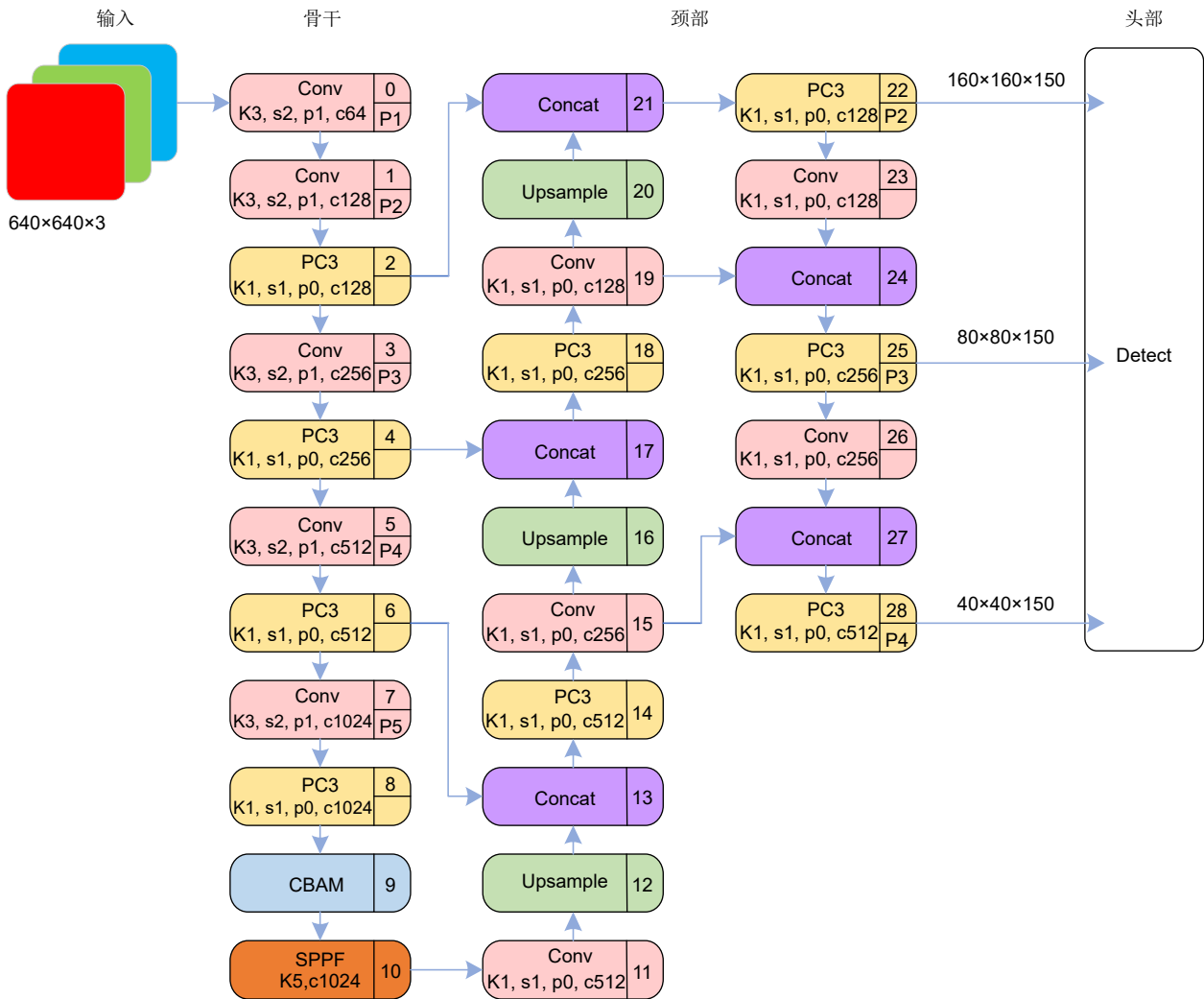


图 9 改进后的网络结构

Fig. 9 Improved network structure

3 实验配置

3.1 数据集筛选

TT100K 数据集由清华大学与腾讯联合制作，是一个专注于中国交通标志的数据集。此数据集已被广泛用于交通标志识别、目标检测等计算机视觉任务的研究和实验中。其丰富的语义信息和高质量的图像使得研究者能够开发出更加准确和鲁棒性强的模型。虽然数据集包含了丰富的交通标志类别，但各类交通标志之间的数量差距大，小目标交通标志较多，可能导致模型在训练时对某些类别的学习不够充分，本文筛选出其中 45 个具有重要语义信息的交通标志。

3.2 数据集预处理

由于不同种类的数量不均匀，不利于模型的训练，同时拓展雾天应用场景，本文利用雾化增强技术对 TT100K 数据集 (共 7962 张) 做扩充处理，得到 TT100K-

FOG 数据集 (共 55020 张)。扩充后的交通标志数量分布如图 10 所示，各类交通标志的数量均有一定的增加，其中 w32 (施工) 交通标志的数量由 104 个增加到 624 个，虽然部分类别的数量仍存在一定差距，但足够满足于本文的模型训练，通过雾化进行数据集的扩充，也适应了雾天情况下的准确识别。

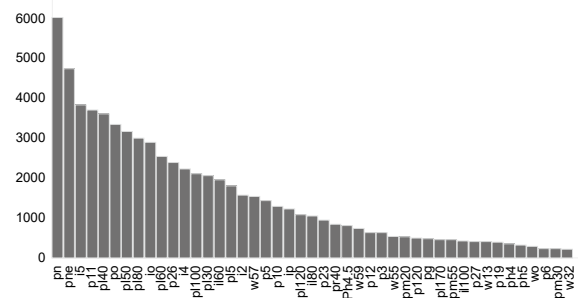


图 10 扩充后交通标志数量分布

Fig. 10 Distribution of traffic signs after expansion

3.3 硬件配置与训练参数细节

进行模型训练和测试的硬件配置为 Intel core(TM) i7-11700K 处理器和显存 10 GB 的 NVIDIA RTX 3080 GPU, 计算机操作系统为 Ubuntu 18.04, 软件环境配置为 Python 3.8、Pytorch 1.8、Opencv 4.1.2 和 CUDA v11.1。

本实验设置训练迭代期为 300, 初始学习速率为 $1e-4$, 动量系数为 0.9, 权重衰减正则项为 0.0005。考虑到 GPU 的容量, 批量大小设置为 16。训练完成后, 选择保存最好的模型进行验证。

3.4 评价指标

本实验选取一些经典指标来评价检测算法的性能, 包括精确率 (precision, P)、召回率 (recall, R)、参数量、推理时间 (latency, L)、FPS 以及 mAP0.5 来评估模型的检测能力。 P 和 R 的计算如式 (9) 和 (10) 所示:

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

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

其中: TP 表示正确检测到的交通标志数量; FP 表示已检测到交通标志, 但检测不正确的数量; FN 表示未检测到交通标志的数量。

模型的参数量通常被用作衡量模型复杂性和大小的指标, 其计算过程如式 (11) 所示:

$$Params = C_{in} \times C_{out} \times K_h \times K_w, \quad (11)$$

其中: C_{in} 代表输入通道数, 即卷积操作前输入的特征图数量, C_{out} 代表输出通道数, 即卷积操作后输出的特征图数量, K_h 、 K_w 分别代表卷积核的高度和宽度。

推理时间是衡量网络是否快速的重要标准, 其计算如式 (12) 所示:

$$Latency = \frac{1}{FPS}, \quad (12)$$

其中: FPS 为每秒传输的图片帧数, FPS 的值越高, 意味着网络的检测速度越快。

mAP0.5 表示 IoU 阈值为 0.5 时所有类别的平均精度。

4 实验结果分析

4.1 消融实验

本文采用了多种方法来改进 YOLOv5 算法, 本节设计几组消融实验来对比分析改进的效果, 实验结果汇总如表 4 所示。

经表中分析, 对比模型 0、1 表明数据集扩充效果明显, 扩充数据集后, 在识别准确率上提升了 6.1%, 达到 89.3%; 模型 2 使用更加轻量的 PConv 构建 PC3 特征提取模块来取代 YOLOv5 骨干和颈部网络中的 C3 模块, 实验结果显示整合 PC3 之后模型精确率和召回率虽然出现一定的略低波动, 但 FPS 相比模型 1 提升 14.1%, 达到 166.7, 推理时间降低 0.9 ms, 参数量降低 33.1%, 达到 4.87 M, 说明改进后的 PC3 模块实现了模型的轻量化; 模型 3 在模型 2 基础上采用 EFPN 结构替代 YOLOv5 中检测头, 相较于模型 2 在精确率上提升了 1.1%, 同时 mAP0.5 提升了 0.2%, 说明添加微小尺寸检测层能够改善对小目标交通标志的检测效果, 降低漏检率; 模型 4 是在模型 3 基础上删去 EFPN 结构中的大目标检测层, 相比模型 3 模型参数量下降 23.4%, 模型 4 主要是进一步减少模型参数量; 模型 5 是在模型 4 基础上采用 Focal-EIoU 作为边界框回归的损失函数取代 YOLOv5 默认函数 CIoU, Focal-EIoU 损失函数通过平衡不同

表 4 消融实验结果

Table 4 Results of ablation experiment

编号	模型	P	R	mAP0.5	FPS	L/ms	Params/M
0	YOLOv5	0.842	0.824	0.861	145.7	6.9	7.10
1	YOLOv5+FOG	0.893	0.838	0.870	145.7	6.9	7.10
2	YOLOv5+FOG+PC3	0.853	0.768	0.840	166.7	6.0	4.87
3	YOLOv5+FOG+PC3+EFPN	0.862	0.759	0.842	135.1	7.4	4.96
4	YOLOv5+FOG+PC3+EFPN'	0.876	0.782	0.854	161.3	6.2	3.80
5	YOLOv5+FOG+PC3+EFPN'+Focal-EIoU	0.906	0.790	0.860	161.3	6.2	3.80
6	YOLOv5+FOG+PC3+EFPN'+Focal-EIoU+CBAM	0.917	0.853	0.899	151.5	6.7	3.95

模型说明: FOG代表扩充雾化数据集TT100K-FOG; PC3代表使用更加轻量的PConv构建PC3特征提取模块来取代YOLOv5骨干和颈部网络中的C3模块; EFPN代表采用延伸的特征金字塔结构, 替代YOLOv5中检测头; EFPN'代表在EFPN结构中删除大目标检测层后, 替代YOLOv5中检测头; Focal-EIoU代表采用Focal-EIoU取代YOLOv5默认函数CIoU; CBAM代表在YOLOv5主干网络中嵌入空间和通道注意力模块。

类别和不同质量的样本, 提升了模型的泛化能力和鲁棒性, EIoU 损失函数克服了 CIoU 的一些缺陷, 通过综合考虑重叠面积、中心点和边长等几何因素, 更准确地衡量预测框与真实框之间的差异。通过分析实验结果, 在引入 Focal-EIoU 损失函数之后模型的参数量维持在 3.80 M, 而精确率和召回率继续提升, 达到了 90.6% 和 79.0%。模型 6 是在模型 5 基础上通过嵌入 CBAM 注意力机制, 提高了模型对小目标交通标志的识别精确率, 进一步提升了模型的性能, 对比模型 5 精确率提高 1.2%, 召回率提升 8.0%, mAP0.5 提升 4.5%。模型 6 虽然在识别精确率上有所提高, 但是 CBAM 的嵌入也导致模型参数量增加 3.9%, 相比精确率的提升这也是值得的, 这些改进不仅增强了模型的泛化能力和鲁棒性, 还使得模型能够更准确地捕获小目标交通标志的信息, 从而提高了检测的精确率和召回率。综合来看, 这些改进方法为交通标志检测任务带来了实质性的进步。

为了进一步验证改进前后 YOLOv5 模型检测效果, 对原 YOLOv5 算法 (表 3 中模型 0) 和改进后的本文算法 (表 3 中模型 6) 进行验证。选择一些不同场景下的图片, 进一步进行验证优化算法的有效性, 尤其针对小目标交通标志的漏检、误检问题以可视化方式验证改进算法的有效性。

由图 11 可以看出, 图 11(a, b) 都可以正确检测

出 p150 (限速 50 km/h) 交通标志, 图 11(c, d) 都正确检测出 i4 (机动车行驶)、i2 (非机动车行驶) 交通标志, 但是通过改进前后实验结果对比分析, 改进后的模型对目标的识别准确率更高。如图中的两组结果展示, 改进前图 11(a) 中的 p150 准确率仅为 79%, 改进之后的模型可高达 92%, 大大提升了检测性能; 改进前图 11(c) 中 i4、i2 的准确率分别为 79%、73%, 改进之后的模型可高达 90%、82%, 分别提高了 14.0%、12.3%。

图 12(a, b) 包含 po (其他禁止标志)、w32 (前方施工) 类交通标志的样本, 因为待检测交通标志处于背光环境, 原 YOLOv5 算法未能检测出 w32 交通标志, 只检测出 po 类交通标志, 本文算法能够完全检测出这两个交通标志。图 12(c, d) 为雾天天气情况下对包含 pne (禁止驶入)、i5 (靠右侧道路行驶)、p150 (限速 50 km/h)、pn (禁止停车) 等交通标志样本, 原 YOLOv5 算法未能检测出 pne 交通标志, 本文算法可以检测出来。由此可以看出, 针对前景背景相似的情况, YOLOv5 算法存在漏检情况, 检测效果不够理想, 本文提出的改进方法显著提升了 YOLOv5 算法对小目标交通标志的漏检情况。

图 13(a, b) 包含 ph4 (限高 4 m)、pne (禁止驶入)、i5 (靠右侧道路行驶) 等交通标志样本, 由于交通标志中包含的数字尺寸较小, 极易容易与其它数字混淆,



图 11 改进前后交通标志精度效果对比图

Fig. 11 Comparison of accuracy effects of traffic signs before and after improvement



图 12 改进前后交通标志漏检效果对比图

Fig. 12 Comparison of missed detection effects of traffic signs before and after improvement

原 YOLOv5 算法把 ph4 (限高 4 m) 误检成 ph4.5 (限高 4.5 m), 本文算法能够正确检测出此标志。图 13(c,d) 为雾天天气情况下对包含 pl40 (限速 40 km/h)、p11 (禁止鸣笛)、pn (禁止停车) 等交通标志样本, YOLOv5 算法把 pl40 误检成 pl80 (限速 80 km/h), 本文算法可以正确检测出来。由此可以看出, 原 YOLOv5 算法存在误检情况, 检测效果不够理想, 本文改进后的模型 6 对良好以及雾天天气下小目标交通标志的误检情况检测效果更好。

4.2 与其它算法的对比实验结果与分析

本节在保证实验环境和训练参数一定的情况下, 训练多种当前比较流行的目标检测算法, 包括 YOLO 系列以及 Faster R-CNN, 所有对比模型在各自的开源平台上实现, 实验数据汇总如表 5 所示。其中 YOLOv6、YOLOv7 以及 YOLOv8 是在 YOLOv5 发表不久随之新发表的流行算法, 每个网络模型在创新点、训练过程等方面各有千秋, 因此在训练的时候, 输入大小均采用 640×640。

本文改进方法与经典两阶段模型 Faster R-CNN 以及一阶段模型 YOLOv4、YOLOX 相比, 无论是精

度还是速度方面都远超前于它们; 与原 YOLOv5 相比, 在 P 、 $mAP_{0.5}$ 上分别提高了 8.9%、4.4%, FPS 也提高了 4.0%, 可满足真实场景中交通标志的实时检测; 在较新的 YOLOv6、YOLOv7 和目前最新的 YOLOv8 模型中, 它们虽然在各自的论文中所验证的结果均优于 YOLOv5, 但在做小目标交通标志检测方面, 表 5 的结果证明改进之后的模型在 P 和 $mAP_{0.5}$ 方面均优于这三种模型, 在检测速度上, 本文模型在 NVIDIA RTX 3080 GPU 设备上的 FPS 值保持在 151.5, 虽然略低于 YOLOv6 和 YOLOv8, 但比 YOLOv7 快 16.4 个百分点。

综上所述, 改进后的模型在 P 和 $mAP_{0.5}$ 都是最优的, 分别达到 91.7% 和 89.9%, 检测精度领先其它模型, 并且检测速度达到 151.5 FPS, 超过原模型 4 个百分点, 实现模型轻量化, 同时保持对小目标交通标志的高检测性能, 因此本文改进模型更加适合于小目标交通标志的检测识别。

5 结语

为了提升在雾天环境下的小目标交通标志识别效

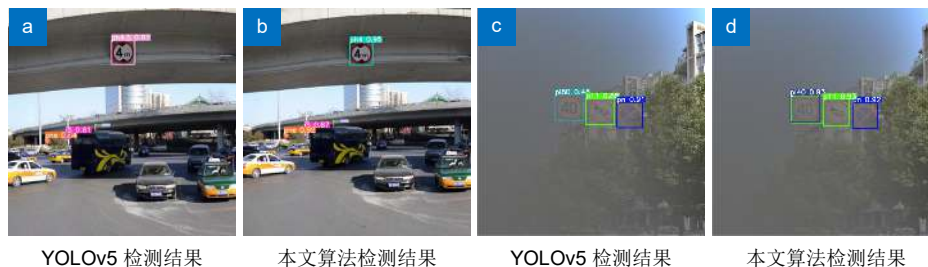


图 13 改进前后交通标志误检效果对比图

Fig. 13 Comparison of false detection effects of traffic signs before and after improvement

表 5 与其它算法的性能对比

Table 5 Performance comparison with other algorithms

模型	平台	主干网	类型	$P/\%$	$mAP_{0.5}/\%$	FPS
Faster R-CNN	MMDetection	ResNet50	Anchor-based	71.9	79.9	57.7
YOLOv4	Darknet	Darknet	Anchor-based	58.7	82.2	80.9
YOLOv5	YOLOv5	Darknet	Anchor-based	84.2	86.1	145.7
YOLOX	MMDetection	Darknet	Anchor-free	72.6	79.7	93.6
YOLOv6	YOLOv6	EfficientRep	Anchor-free	77.7	81.1	162.8
YOLOv7	YOLOv7	E-ELAN	Anchor-based	72.0	77.4	130.2
YOLOv8	YOLOv8	Darknet	Anchor-free	87.7	83.7	171.4
Ours	YOLOv5	Darknet	Anchor-based	91.7	89.9	151.5

率, 本文基于 YOLOv5 模型进行改进, 首先对数据集进行雾化操作以适应在雾天情况下的准确识别, 使用更加轻量的 PConv 构建 PC3 特征提取模块; 随后在颈部网络中采用 EFPN 结构, 为小目标添加一个小目标检测头, 同时删去原始颈部网络中针对大目标的检测头, 在提高小目标交通标志识别精度的同时降低网络参数; 最后引入 Focal-EIOU 替换 CIOU 作为损失函数, 以此来解决小目标交通标志的误检和漏检问题; 嵌入 CBAM 注意力机制, 实现模型轻量化的同时显著提升网络模型的特征提取能力, 在预处理后的 TT100K 数据集上训练模型并进行验证。实验结果表明, 改进的模型对小目标交通标志的误检、漏检问题以及雾天情况下的交通标志识别效果有所提升, 优于大部分最新的交通标志识别模型。在 NVIDIA 3080 设备上 FPS 值达到 151.5, 可满足真实场景中交通标志的实时检测需求。该模型通过雾化增强技术提高了在雾天情况下的识别精度, 但雨雪天气未能充分考虑, 同时该模型由 TT100K 中的 45 类交通标志训练而来, 后期有望继续扩展数据集, 识别更多种类的交通标志。

参考文献

- [1] Wang R X, Wu J P, Xu H. Overview of research and application on autonomous vehicle oriented perception system simulation[J]. *J Syst Simul*, 2022, **34**(12): 2507–2521.
王若萱, 吴建平, 徐辉. 自动驾驶汽车感知系统仿真的研究及应用综述[J]. *系统仿真学报*, 2022, **34**(12): 2507–2521.
- [2] Acharya S, Nanda P K. Adjacent LBP and LTP based background modeling with mixed-mode learning for foreground detection[J]. *Pattern Anal Appl*, 2021, **24**(3): 1047–1074.
- [3] Shao F M, Wang X Q, Meng F J, et al. Real-time traffic sign detection and recognition method based on simplified Gabor wavelets and CNNs[J]. *Sensors*, 2018, **18**(10): 3192.
- [4] Maria Dominic Savio M, Deepa T, Bonasu A, et al. Image processing for face recognition using HAAR, HOG, and SVM algorithms[J]. *J Phys Conf Ser*, 2021, **1964**(6): 062023.
- [5] Burges C J C. A tutorial on support vector machines for pattern recognition[J]. *Data Min Knowl Discovery*, 1998, **2**(2): 121–167.
- [6] Thamilselvan P. Lung cancer prediction and classification using adaboost data mining algorithm[J]. *Int J Comput Theory Eng*, 2022, **14**(4): 149–154.
- [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, USA, 2014: 580–587.
<https://doi.org/10.1109/CVPR.2014.81>.
- [8] Girshick R. Fast R-CNN[C]//2015 *IEEE International Conference on Computer Vision (ICCV)*, Santiago, Chile, 2015: 1440–1448.
<https://doi.org/10.1109/ICCV.2015.169>.
- [9] Ren S Q, He K M, Girshick R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[J]. *IEEE Trans Pattern Anal Mach Intell*, 2017, **39**(6): 1137–1149.
- [10] Liu W, Anguelov D, Erhan D, et al. SSD: single shot MultiBox detector[C]//*The 14th European Conference on Computer Vision*, Amsterdam, The Netherlands, 2016: 21–37.
https://doi.org/10.1007/978-3-319-46448-0_2.
- [11] Bochkovskiy A, Wang C Y, Liao H Y M. YOLOv4: optimal speed and accuracy of object detection[Z]. arXiv: 2004.10934, 2020.
<https://arxiv.org/abs/2004.10934v1>.
- [12] Ultralytics. YOLOv5[EB/OL].
<https://github.com/ultralytics/YOLOv5>.
- [13] Ge Z, Liu S T, Wang F, et al. YOLOX: exceeding YOLO series in 2021[Z]. arXiv: 2107.08430, 2021.
<https://arxiv.org/abs/2107.08430>.
- [14] Li C Y, Li L L, Jiang H L, et al. YOLOv6: a single-stage object detection framework for industrial applications[Z]. arXiv: 2209.02976, 2022.
<https://arxiv.org/abs/2209.02976>.
- [15] 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]//2023 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Vancouver, Canada, 2023: 7464–7475.
<https://doi.org/10.1109/CVPR52729.2023.00721>.
- [16] Reis D, Kupec J, Hong J, et al. Real-time flying object detection with YOLOv8[Z]. arXiv: 2305.09972, 2024.
<https://arxiv.org/abs/2305.09972>.
- [17] Chen X, Peng D L, Gu Y. Real-time object detection for UAV images based on improved YOLOv5s[J]. *Opto-Electron Eng*, 2022, **49**(3): 210372.
陈旭, 彭冬亮, 谷雨. 基于改进 YOLOv5s 的无人机图像实时目标检测[J]. *光电工程*, 2022, **49**(3): 210372.
- [18] Yang J, Sun T, Zhu W C, et al. A lightweight traffic sign recognition model based on improved YOLOv5[J]. *IEEE Access*, 2023, **11**: 115998–116010.
- [19] Chen L, Zhang J L, Peng H, et al. Few-shot image classification via multi-scale attention and domain adaptation[J]. *Opto-Electron Eng*, 2023, **50**(4): 220232.
陈龙, 张建新, 彭昊, 等. 多尺度注意力与领域自适应的小样本图像识别[J]. *光电工程*, 2023, **50**(4): 220232.
- [20] Zhang J M, Xie Z P, Sun J, et al. A cascaded R-CNN with multiscale attention and imbalanced samples for traffic sign detection[J]. *IEEE Access*, 2020, **8**: 29742–29754.
- [21] Zhang H B, Qin L F, Li J et al. Real-time detection method for small traffic signs based on Yolov3[J]. *IEEE Access*, 2020, **8**: 64145–64156.
- [22] Guo Y, Liang R L, Wang R M. Cross-domain adaptive object detection based on CNN image enhancement in foggy conditions[J]. *Comput Eng Appl*, 2023, **59**(16): 187–195.
郭迎, 梁睿琳, 王润民. 基于 CNN 图像增强的雾天跨域自适应目标检测[J]. *计算机工程与应用*, 2023, **59**(16): 187–195.
- [23] Lin H B, Zhou J L, Chen M Z. Traffic sign detection algorithm based on improved YOLOv4[C]//2022 *IEEE 10th Joint International Information Technology and Artificial Intelligence Conference (ITAIC)*, Chongqing, China, 2022: 2156–2160.
<https://doi.org/10.1109/ITAIC54216.2022.9836923>.
- [24] Wang Y D, Guo J C, Wang T B. Algorithm for foggy-image pedestrian and vehicle detection[J]. *J Xidian Univ*, 2020, **47**(4): 70–77.
汪昱东, 郭继昌, 王天保. 一种改进的雾天图像行人和车辆检测算法[J]. *西安电子科技大学学报*, 2020, **47**(4): 70–77.
- [25] Lang B K, Lü B, Wu J Q, et al. A traffic sign detection model based on coordinate attention-bidirectional feature pyramid

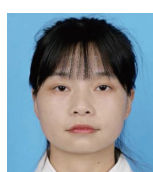
- network[J]. *J Shenzhen Univ (Sci Eng)*, 2023, **40**(3): 335–343.
郎斌柯, 吕斌, 吴建清, 等. 基于 CA-BIFPN 的交通标志检测模型[J]. *深圳大学学报 (理工版)*, 2023, **40**(3): 335–343.
- [26] Zhu H Y, Han J N, Xu Y. Printed circuit board blemishes detection based on the improved YOLOv5s[J]. *Foreign Electron Meas Technol*, 2023, **42**(3): 152–159.
朱宏禹, 韩建宁, 徐勇. 基于改进型 YOLOv5s 的印刷电路板瑕疵检测[J]. *国外电子测量技术*, 2023, **42**(3): 152–159.
- [27] Wang Y W, Lu Y, Dou Y H, et al. Synchronous GPS spoofing Identification based on K-means clustering[J]. *J Electron Inf Technol*, 2023, **45**(11): 4137–4149.
王屹伟, 路寅, 寇艳红, 等. 基于 K-means 聚类的 GPS 同步式欺骗识别方法[J]. *电子与信息学报*, 2023, **45**(11): 4137–4149.
- [28] Zhang Z D, Tan M L, Lan Z C, et al. CDNet: a real-time and robust crosswalk detection network on Jetson nano based on YOLOv5[J]. *Neural Comput Appl*, 2022, **34**(13): 10719–10730.
- [29] Chen C Y, Liu M Y, Tuzel O, et al. R-CNN for small object detection[C]//*Proceeding of the 13th Asian Conference on Computer Vision*, Taipei, China, 2016: 214–230.
https://doi.org/10.1007/978-3-319-54193-8_14.
- [30] Lin T Y, Goyal P, Girshick R, et al. Focal loss for dense object detection[J]. *IEEE Trans Pattern Anal Mach Intell*, 2020, **42**(2): 318–327.
- [31] Woo S, Park J, Lee J T, et al. CBAM: convolutional block attention module[C]//*15th European Conference on Computer Vision*, Munich, Germany, 2018: 3–19.
https://doi.org/10.1007/978-3-030-01234-2_1.

作者简介



【通信作者】曲立国(1979-), 男, 博士, 副教授, 硕士生导师, 主要研究方向为智能信息处理、检测技术与自动化装置、FPGA 及嵌入式系统开发与设计。

E-mail: lqgu@ahnu.edu.cn



刘玉玲(1998-), 女, 硕士研究生, 主要研究方向为图像处理、嵌入式系统设计。

E-mail: 1316895061@qq.com



张鑫(1998-), 女, 硕士研究生, 主要研究方向为深度学习、图像处理。

E-mail: 2897069764@qq.com



陈国豪(1998-), 男, 硕士, 主要研究方向为图像处理、嵌入式系统设计、深度学习及应用。

E-mail: cgh591614156@163.com



卢自宝(1985-), 男, 博士, 副教授, 硕士生导师, 主要研究方向为网络化控制系统、微电网控制技术、移动机器人控制。

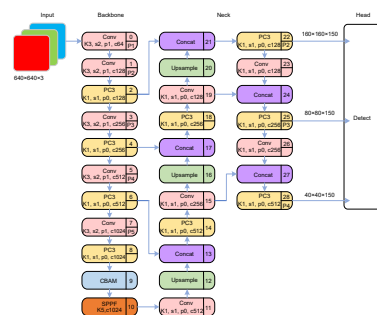
E-mail: luzibao@163.com



扫描二维码, 获取PDF全文

A traffic sign recognition method based on improved YOLOv5

Qu Ligu^{1,2*}, Zhang Xin¹, Lu Zibao¹, Liu Yuling¹, Chen Guohao³



Improved network structure

Overview: Traffic sign detection is a crucial component in the field of autonomous driving. Current traffic sign recognition methods face challenges such as missed detections, false positives, high model complexity, and poor robustness in these representative and complex real-world conditions, particularly under the foggy weather. To address these issues, this paper proposes an improved method for small traffic sign detection based on YOLOv5. Firstly, to tackle the problem of imbalanced sample data, the traffic sign dataset is augmented with foggy conditions to enhance the model's generalization capability. This augmentation helps the model to better handle diverse environmental conditions, improving its robustness. Secondly, a lightweight Partial Convolution (PConv) is introduced to construct the PC3 feature extraction module, replacing the C3 module in the original YOLOv5 model. This modification reduces the number of model parameters and enhances processing speed without compromising detection performance. Subsequently, an Extended Feature Pyramid Network (EFPN) is employed in the neck network, adding detection heads specifically for small objects while removing the heads for large object detection. This specialization optimizes the model's performance for small object detection. Additionally, the K-means clustering algorithm is used on the TT100K dataset to recalculate and adjust the size and ratio of anchor boxes, better accommodating small object detection. The Focal-EIOU loss function replaces the original CIOU loss function to address class imbalance and false positive issues in small object detection. Finally, a Convolutional Block Attention Module (CBAM) is embedded in the backbone network. This module performs attention operations on both channel and spatial dimensions, further enhancing the model's feature extraction capability. Compared with the original YOLOv5 algorithm, ablation experiments on the TT100K dataset show that the improved model achieves an 8.9% increase in precision (*P*) and a 4.4% increase in mean Average Precision (mAP0.5), with a 44.4% reduction in parameter count, reaching a frame rate (FPS) of 151.5 on NVIDIA 3080 devices. Furthermore, comparative experiments with mainstream object detection algorithms such as Faster RCNN, YOLOv7, and YOLOv8 demonstrate that the improved model achieves optimal *P* and mAP0.5, reaching 91.7% and 89.9% respectively, leading other models in detection accuracy. The model parameter count is 3.95M, realizing a lightweight design while maintaining high detection performance for small traffic signs. This improved model is suitable for real-time traffic sign detection in real-world scenarios, ensuring the reliable operation in autonomous driving systems.

Qu L G, Zhang X, Lu Z B, et al. A traffic sign recognition method based on improved YOLOv5[J]. *Opto-Electron Eng*, 2024, 51(6): 240055; DOI: [10.12086/oe.2024.240055](https://doi.org/10.12086/oe.2024.240055)

Foundation item: Project supported by Wuhu Science and Technology Plan Funding Project (2022jc07), and Anhui Provincial Department of Education Graduate Innovation and Entrepreneurship Practice Funding Project (2022cxcysj058, 2023cxcysj043)

¹School of Physics and Electronic Information, Anhui Normal University, Wuhu, Anhui 241002, China; ²Anhui Provincial Engineering Research Center for Information Fusion and Control of Intelligent Robots, Wuhu, Anhui 241002, China; ³Wuhan Mingke Rail Transit Equipment Co., Ltd., Wuhan, Hubei 430074, China

* E-mail: lgqu@ahnu.edu.cn