

基于改进Yolov8的交通标志检测算法

苟 皓, 唐 岚, 孟忠伟

西华大学汽车与交通学院, 四川 成都

收稿日期: 2024年6月15日; 录用日期: 2024年7月5日; 发布日期: 2024年7月17日

摘 要

针对在自动驾驶行驶中交通标志背景复杂和检测目标小等问题, 设计了一种基于yolov8改进的交通标志识别算法。首先在yolov8的DarkNet53中使用动态蛇形卷积, 通过引入动态蛇形卷积自适应调整卷积核形状和大小, 可以更好地适应小目标尺寸, 从而提高准确性, 而且可通过目标形状的和轮廓, 局部调整卷积核, 使在卷积过程中更加关注目标特征, 计算复杂度降低, 提高检测精度; 其次引入局部敏感卷积(LSKA)对SPPF进行改进, 通过金字塔池化和频率金字塔来获得多尺度和多频率的特征表示, 再利用LSKA来提取局部位置不变性, 让其具有更好的鲁棒性和精度, 提高网络对交通标志识别的准确性。实验结果表明, 改进后的算法在精确率、召回率、平均精确率上相比原算法分别提升2.8%、2.7%、3.3%, 检测速度满足实时性要求。

关键词

Yolov8算法, 交通标志, 动态蛇形卷积, 可分离卷积

Improving Yolov8 for Traffic Sign Detection Algorithm

Hao Gou, Lan Tang, Zhongwei Meng

School of Automotive and Transportation Engineering, Xihua University, Chengdu Sichuan

Received: Jun. 15th, 2024; accepted: Jul. 5th, 2024; published: Jul. 17th, 2024

Abstract

A traffic sign recognition algorithm based on Yolov8 improvement is designed to address issues such as complex backgrounds of traffic signs and small target detection in autonomous driving. Firstly, dynamic snake convolution is introduced into DarkNet53 of Yolov8. By dynamically adjusting the shape and size of convolution kernels, adaptive adjustments are made to better accommodate small target sizes, thereby improving accuracy. Moreover, by adjusting convolution

kernels locally based on target shapes and contours, attention to target features during the convolution process is enhanced, reducing computational complexity and improving detection accuracy. Secondly, Local Sensitivity Kernel Aggregation (LSKA) is introduced to improve the Spatial Pyramid Pooling Fusion (SPPF). By utilizing pyramid pooling and frequency pyramids to obtain multi-scale and multi-frequency feature representations, and then using LSKA to extract local position invariance, the algorithm exhibits better robustness and accuracy, enhancing the network's accuracy in traffic sign recognition. Experimental results show that the improved algorithm achieves a 2.8% increase in precision, a 2.7% increase in recall, and a 3.3% increase in average precision compared to the original algorithm, while meeting the requirements for practical detection speed.

Keywords

Yolov8 Algorithm, Traffic Signs, Dynamic Snake Convolution, Separable Convolution

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

道路交通标志识别[1]是自动驾驶系统中比较重要的环节,即影响着自动驾驶的安全性和智能性等,所以实现对交通标志信号检测的准确性和快速性对自动驾驶尤其重要。为此对交通标志识别的精确度和速度都提出了更高的要求[2]。在实际道路行驶时,交通标志检测受天气、行驶速度、标志背景、视角变化的影响,使交通标志出现检测精度和速度的差异,甚至出现漏检的情况。所以设计一个高速度和准率的检测方法显得尤为重要[3]。

目前为止,关于道路检测方法有基于 Faster R-CNN [4]的交通标识检测、基于 SSD 的交通标识识别、基于 SVM 和 HOG 特征的交通标识识别和基于深度学习[5]的交通目标检测。深度学习是现在比较主要的目标检测算法,与其它检测算法相比,其具有高准确性、灵活性、处理复杂场景的能力。

对于深度学习的道路交通标志检测算法,Li 等提出了一种改进的 Faster R-CNN [6]复杂环境交通检测方法,针对复杂环境和小目标遮挡等问题,通过 VGG16 和多尺度融合获取更多的特征信息;顾峰等[7]提出了设计亮度与对比度增强模块以凸显低照度图像特征信息,然后结合空间转换单元构建轻量化网络,通过弱化图像携带的无关信息聚焦数据中感兴趣区域,有效分离了背景噪声,改善了空间不变形;Chu 等人[8]进 Yolov5 的 TRD-YOLO 网络,在主干中加入了 Transformer 机制,同时加入了轻量级解耦头,在面对复杂环境中的小型交通标识时,它具有更好的检测性能和更强的鲁棒性,但在特殊天气下检测交通标识时,准确性和精度还有待提高。石镇岳等人[9]提出了基于注意力基础的交通标志检测模型识别,通过 K-means++聚类算法对数据集进行聚类,其次在主干特征网络中引入坐标注意力机制,最后引入空洞空间金字塔池化,捕获上下文信息,使其增强了小目标识别能力,但由于行车环境复杂,准确度容易受到天气、光照以及其它不利因素影响;Gao 等人[10]提出了一种改进 Faster-R-CNN 交通标识检测算法,将特征金字塔融入到算法中,可以精确地提取交通标识,其次在骨干加入了可变形卷积(DCN),使相似的标识更容易被区分,最后采用 ROI 对齐取代 ROI 池化,避免池化导致交通标识的特征细节丢失,该算法有助于提高极端环境下的交通标识检测性能,但检测效率过小,不能满足实时检测的要求。

对此本文提出了基于 yolov8 改进算法对道路交通标志检测算法,通过在 Darknet53 中使用动态蛇形

卷积，使其自适应调整卷积核形状和大小，可以更好地适应小目标尺寸，从而提高准确性，然后再引入 LSKA 通过利用大且可分离的卷积核以及空间扩张卷积来捕捉图像的广泛上下层信息，同时 LSKA 无需额外的块，使其降低了计算复杂性，而且内存使用更低。

2. YOLOv8 模型

YOLOv8 算法简介

YOLOv8 是目标检测领域的算法，YOLOv8 的核心思想是将目标检测任务转化为单一的神经网络执行，实现了端到端的训练和目标检测，并且使用了更强大的 Backbone 网络，采用更高分辨率输入图像以提高检测速度，以及引入一些优化策略来提高推理速度[11]。

YOLOv8 算法是由 Ultralytics LCC 公司在 YOLOv5 模型上改进的单阶段检测模型[11]，其结构模型如图 1 所示，YOLOv8 网络由主干网络(backbone)、颈部网络(neck)和头部网络(head)组成，主干网络负责特征提取，颈部主要对特征图进行融合，头部进行预测。

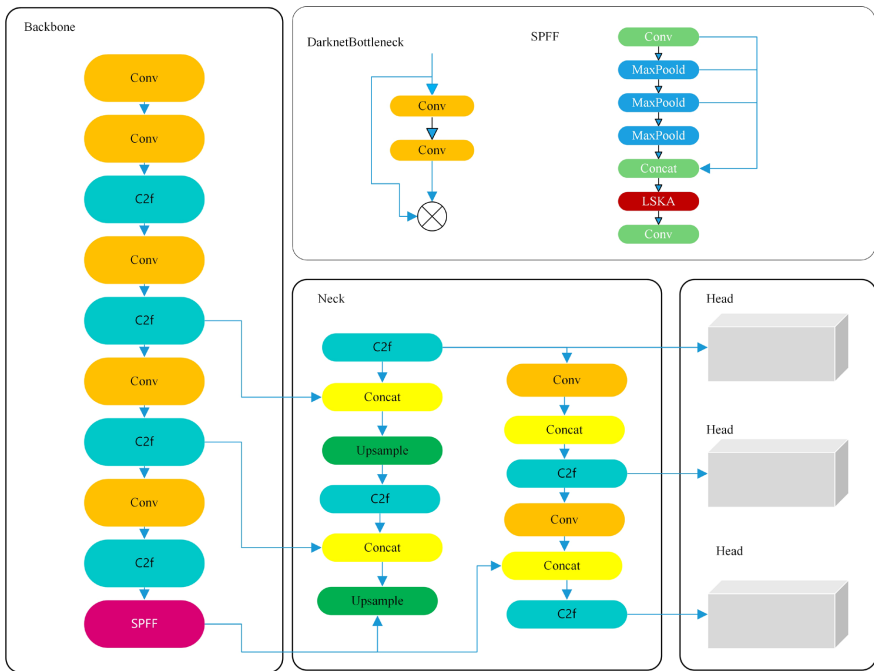


Figure 1. YOLOv8 model

图 1. YOLOv8 模型

3. YOLOv8 算法改进

在 yolo8 的 CSPDarknet53 中使用动态蛇形卷积，通过引入动态蛇形卷积自适应调整卷积核形状和大小，可以更好地适应小目标尺寸，从而提高准确性，而且可通过目标形状的和轮廓，局部调整卷积核，使在卷积过程中更加关注目标特征，提高检测精度；同时其次引入局部敏感卷积(LSKA)对 SPFF 进行改进，通过金字塔池化和频率金字塔来获得多尺度和多评率的特征表示，再利用 LSKA 来提取局部位置不变性，让其具有更好的鲁棒性和精度。

3.1. 动态蛇形卷积(Dynamic Spatial Pyramid Convolution)

在道路交通标志识别中小目标图像识别是一大难点，原因在于小目标较小，往往在特征图就丢失了

细微特征。因此，在 yolov8 中加入 DSPC 算法。

DSPC 主要原理是：首先，DSPC 通过学习或设计生成一个初始的卷积核，然后对输入特征图动态的调整卷积核形状，使之更适应当前目标的尺度和形状，调整后的卷积核再次对输入特征进行卷积操作，以及提取更准确的目标特征，最后，将经过动态蛇形卷积操作得到的特征与其它层的特征进行融合，以获取更全面的和准确的特征表示。

以标准 2D 卷积 3*3 卷积核 K 为例，其公示如下：

$$K = \{(x-1, y-1), (x-1, y), \dots, (x+1, y+1)\} \quad (1)$$

式中：\$(x, y)\$ 为卷积核的中心坐标，\$(x+m, y+n)\$ 为初始坐标 \$(x, y)\$ 的偏移量，其中 \$m\$ 和 \$n\$ 为水平和垂直方向上的移动量。

为了使卷积核更加灵活的关注目标图像的几何特征，引入了偏移量 \$\Delta\$，但是让模型自由学习变形偏移量，则感知范围容易偏移目标，所以本文采用迭代策略，对每个待处理的目标依次选择观察位置，这样既保证了注意力的连续性，又不会因为变形量偏大使感知范围过于分散。

在 DSConv 中，在 X 轴和 Y 轴上拉直卷积核。以大小为 9 的卷积核为示范，并且以坐标轴为例，\$K\$ 中每一个网格具体位置表示为：\$K_{i\pm c}=(x_{i\pm c}, y_{i\pm c})\$，其中 \$c=\{0,1,2,3,4\}\$ 表示与中心网格的水平距离。卷积核 \$K\$ 中每个网格位置 \$K_{i\pm c}\$ 的选择是一个累积过程。从中心位置 \$K_i\$ 开始，距离中心网格的位置取决于前一个网格的位置：\$K_{i+1}\$ 与 \$K_i\$ 相比增加了一个偏移量 \$\Delta=\{\delta|\delta\in[-1,1]\}\$。因此，偏移量需要求和，从而确保卷积核符合线性形态结构[12]。

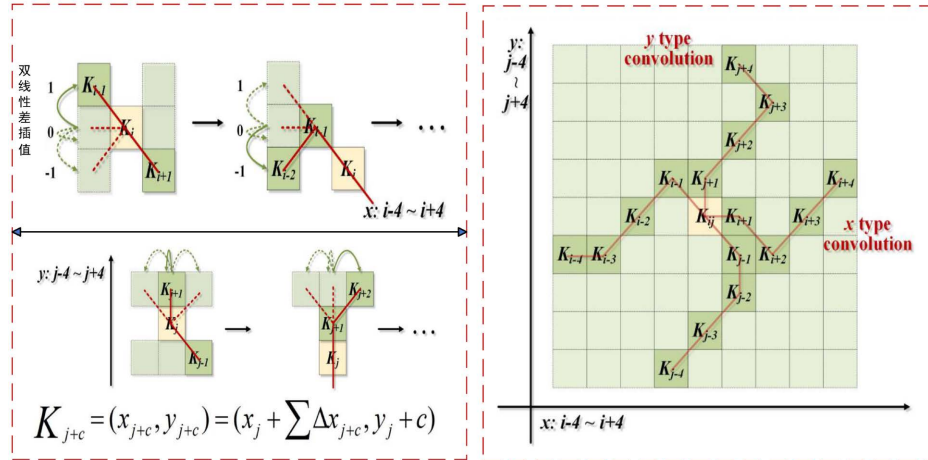


Figure 2. Left: Schematic diagram of coordinates; Right: Change in receptive field

图 2. 图左为坐标示意图、图右为感受野变化图

在 \$x\$ 轴方向上为：

$$K_{i\pm c} = \begin{cases} (x_{i+c}, y_{i+c}) = (x_i + c, y_i + \sum_{i=1}^{i+c} \Delta y), \\ (x_{i-c}, y_{i-c}) = (x_i - c, y_i + \sum_{i=1}^i \Delta y), \end{cases} \quad (2)$$

在 \$y\$ 轴方向上为：

$$K_{i\pm c} = \begin{cases} (x_{j+c}, y_{j+c}) = (x_j + c, y_j + \sum_{j=1}^{j+c} \Delta y), \\ (x_{j-c}, y_{j-c}) = (x_j - c, y_j + \sum_{j=1}^j \Delta y), \end{cases} \quad (3)$$

由于偏移量 Δ 通常为分数，双线性插值实现如下：

$$K = \sum_{K'} B(K', K) \cdot K' \tag{4}$$

由以上公式可得得到一个原理图如图 2 所示(此图是以 3*3 卷积核为例)。

采用以上方法可以在交通标志图像别中，动态蛇形卷积可以根据目标的尺寸和形状动态的调整卷积核大小。

3.2. 改进 SPPF-LSKA 模块

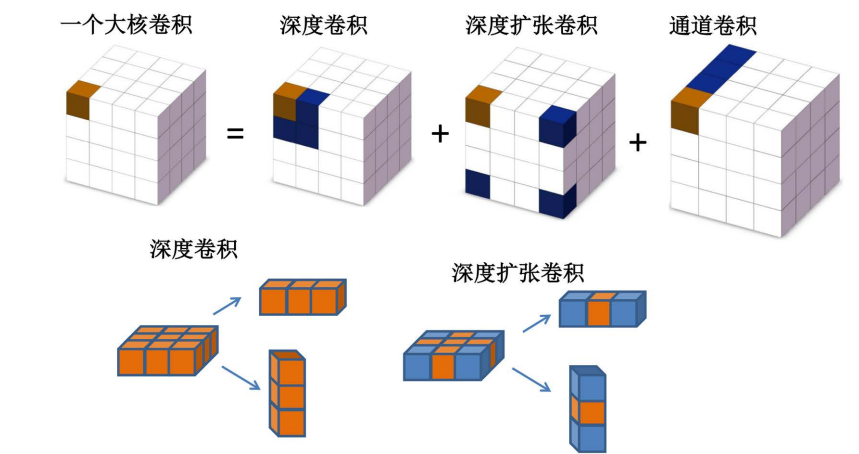


Figure 3. LSKA decomposition structure diagram
图 3. LSKA 分解结构图



Figure 4. LSKA series decomposition diagram
图 4. LSKA 串联分解示意图

由于交通标志目标检测背景复杂，为了应对模型在复杂多变的环境的影响，在 YOLOv8 骨干网络中引入了 LSKA (Large Separable Kernel Attention) 注意力机制，使模型忽略无关背景信息干扰，注意到有效的标志特征信息。以往注意力机制存在许多不足，例如 self-attention 注意力机制具有较强的长期依赖性和适应性，但忽略了图片的二次结构。LSKA 注意力机制避免了 self-attention 注意力机制存在的忽略图片二维结构问题，但其处理大尺寸卷积核时计算量过大的问题[13]。

LSKA 是一种大可分离核注意力模块，它通过大核卷积操作进行分解来获取长期适应性和依赖性，同时又将分解得到的 2D 卷积核再次分解为串联的 1D 核，降低了计算的复杂程度和内存占用需求。首先将 $K \times K$ 的卷积分解为 $(2d - 1) \times (2d - 1)$ 深度卷积、 $K/d \times K/d$ 深度扩张卷积和 1×1 卷积，其次将 2D 的深

度卷积与深度扩张卷积继续分解为 1D 的横向卷积核与垂直卷积核，最后将分解后的卷积核按照顺序串联，其分解结构图如图 3 和图 4 所示[13]。

改进后的 LSKA 结构如图 5 所示。

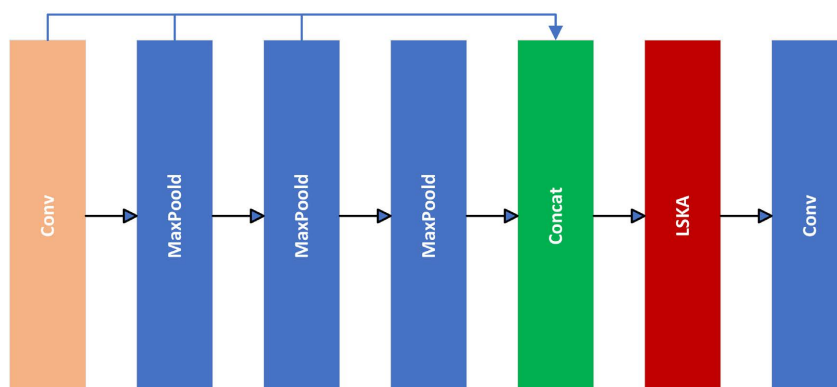


Figure 5. Improved LSKA structure diagram

图 5. 改进 LSKA 结构图

采用以上方法可以让模型在检测过程中忽略无关背景的信息干扰，提高检测精度，并提高鲁棒性。

4. 实验结果及分析

4.1. 数据集

本文使用的数据集为中国交通标志检测数据集 2021 (Chinese Traffic Sign Detection Benchmark, CCTSDB) [14]。数据集共有 17,856 张图片，其类别为指示、禁令及警告标识三类。

4.2. 实验环境

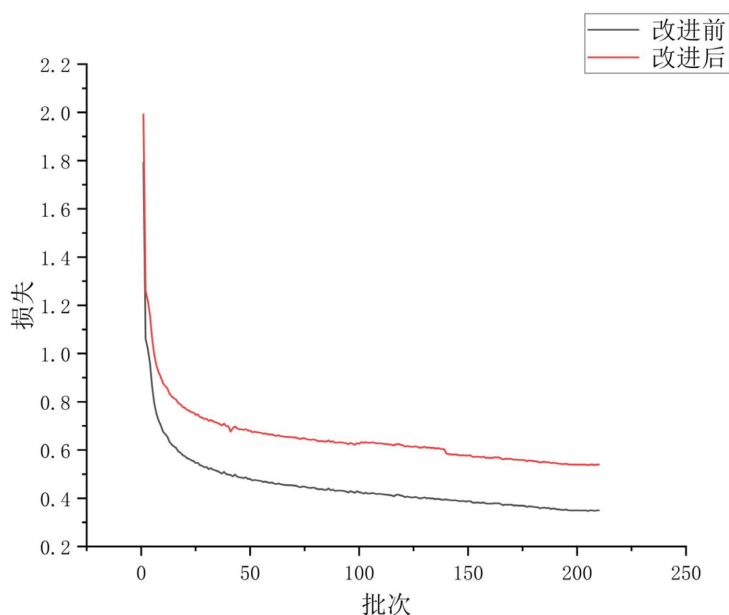


Figure 6. Training loss function curve

图 6. 训练损失函数曲线

本文实验环境系统为 Windows10 操作系统，电脑显卡配置为 NVIDIA GeForce RTX4060 显卡，显存为 8 GB，CPU 为 Intel Core i5-12400，使用 Pytorch (环境为 Python-3.8.19 torch-2.1.0 + cu121) 框架进行部署。参数如下：输入图片大小 640×640 ，批次量大小为 16，数据器加载数量为 16，初始学习率 0.01，训练批次为 250 次，优化器为 SGD。如图所示，当训练次数达到 200 次左右时，改进前后的模型逐渐收敛，并且从运行结果看出，改进后的模型比原模型收敛更快和损失更低，运行结果如图 6 所示。

4.3. 模型评价指标

模型通过训练和验证可通过精确率 AP、召回率 AR 和平均精确率 R_{mAP} 作为评价指标，其公式如 (5)~(8)。这些指标常用目标检测领域，也是 YOLO 算法结果评价的指标。这些指标可作为联合评价，对 YOLO 模型在交通标识检测中表现进行评估。

$$P = \frac{N_{TP}}{N_{TP} + N_{FP}} \quad (5)$$

$$P = \frac{N_{TP}}{N_{TP} + N_{FN}} \quad (6)$$

$$R_{AP} = \int_0^1 P \quad (7)$$

$$R_{mAP} = \frac{\sum_{i=1}^C R_{APi}}{C} \quad (8)$$

式中 N_{TP} 为真阳性，意为检测到正样本的物体； N_{FP} 为假阳性，意为将负样本检测为正样本的物体； N_{FN} 为假阴性，意为将正样本检测为负样本。 $P(r)$ 为精确率和召回率平滑曲线； P_i 中的 i 为第 i 个内别； C 为总类别。

Table 1. Comparison table of ablation experiment data

表 1. 消融实验数据对照表

算法	P/%	R/%	RmAP/%	GFLOPs(G)	参数量(M)
Yolov8n	93.3	89.7	94.2	8.2	3.2
DynamicConv	94.6	91.2	95.1	7.7	4.3
SPPF-LSKA	95.3	90.8	94.6	8.5	3.3
整合改进	96.1	92.4	97.5	8.4	3.5

验证本文提出的改进方法对交通标志的检测的性能进行评测，采用单个改进的消融实验如表 1 所示。

从表 1 中可以得到，添加 DynamicConv 时，虽然模型的参数量大幅度增加，但精确率和平均准确率有所上升，并且计算复杂度下降 0.5 GFLOPs；添加 SPPF-LSKA 模块时，虽然模型复杂度增加，但是模型的精确率、平均准确率上升。整合修改后，改进的模型进增加小部分的参数量和计算复杂度增加 0.2 GFLOPs，但是在精确率、召回率和平均精确率上提升效果显著，分别提升 2.8%、2.7%、3.3%。

4.4. 实际效果验证

原 Yolov8 算法在面对不同尺寸、小目标和复杂背景的情况下检测精度和效果有待提升，本文挑选了部分复杂场景进行原算法与改进 Yolov8 算法测试效果如下图 7~9 (左图为改进的，右图为原算法) 所示。



Figure 7. Scene one

图 7. 场景一



Figure 8. Scene two

图 8. 场景二

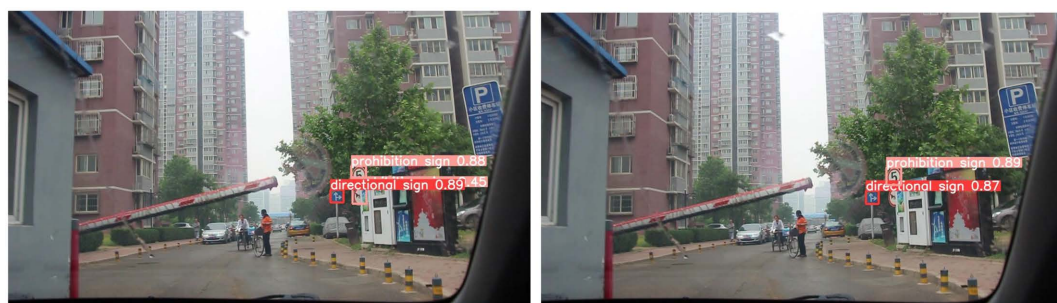


Figure 9. Scene three

图 9. 场景三

经过以上对比可知,改进后的算法检测到了更多的交通标识,并且置信度也高于原算法,验证了本文开始提出改进型意见在准确性上面的提升,以及有效改进了在复杂背景下的识别。

4.5. 目标检测算法对比

为验证本文改进算法的优越性，将本文算法与目前主流的交通标志检测算法进行对比，选取 SSD [15]、Faster R-CNN、yolov5s、yolov8n 与本文改进算法进行对比实验，实验结果如表 2。

Table 2. Algorithm comparison
表 2. 算法对比

算法	P/%	R/%	R _{mAP} /%	FLOPS(G)	FPS
SSD	23.5	33.4	39.8	15.4	81
Faster R-CNN	55.6	76.6	77.6	92.4	26
Yolov5s	94.2	87.3	93.6	24	111
Yolov8n	93.3	89.7	94.2	8.2	86
改进算法	96.1	92.4	97.5	8.5	103

5. 结束语

本文基于 yolov8 算法改进了一种交通标志检测算法，目的在于解决当前主流算法在小目标检测过程的精度低，并且在复杂背景下出现的漏检问题。改进的算法在实验数据集上平均精度达到 97.5%，相比原算法提升 3.3 个百分点，并且在复杂环境下的检测效果也得到了显著性提升，但是改进算法也存在一定的漏检情况。在后面的工作将会对算法进一步优化，提升精确度，并且将会把算法部署到硬件，使其落地，为自动驾驶进行赋能增效。

参考文献

- [1] 陈晗晗, 王俊英, 任肖月. 交通标志的检测与识别方法研究综述[J]. 信息技术与信息化, 2024(3): 77-82.
- [2] 陈红, 王相超, 陈志琳. 自然场景下的交通标志检测与识别[J]. 电子测量技术, 2021, 44(12): 102-109.
- [3] 赵坤, 刘立, 孟宇, 等. 弱光照条件下交通标志检测与识别[J]. 工程科学学报, 2020, 42(8): 1074-1084.
- [4] Hu, P., Lu, J., Cui, Y., Hu, B. and Liang, F. (2023). Insulator Defect Detection Based on Faster R-CNN and Yolov3 Algorithm. 2023 *IEEE 4th China International Youth Conference on Electrical Engineering (CIYCEE)*, Chengdu, 8-10 December 2023, 1-4. <https://doi.org/10.1109/ciycee59789.2023.10401677>
- [5] Wang, Y., Pan, F., Li, Z., Xin, X. and Li, W. (2023). CoT-YOLOv8: Improved YOLOv8 for Aerial Images Small Target Detection. 2023 *China Automation Congress (CAC)*, Chongqing, 17-19 November 2023, 4943-4948. <https://doi.org/10.1109/cac59555.2023.10451989>
- [6] Bochkovskiy, A., Wang, C.Y. and Liao, H.Y.M. (2020) Yolov4: Optimal Speed and Accuracy of Object Detection. arXiv: 2004.10934.
- [7] 顾峰, 四兵锋. 基于多源特征增强的交通标志识别方法[J]. 北京交通大学学报, 2023, 47(4): 73-80.
- [8] Chu, J., Zhang, C., Yan, M., Zhang, H. and Ge, T. (2023) TRD-YOLO: A Real-Time, High-Performance Small Traffic Sign Detection Algorithm. *Sensors*, **23**, Article 3871. <https://doi.org/10.3390/s23083871>
- [9] 石镇岳, 侯婷, 苏勇东. 改进 YOLOv7 的交通标志检测算法[J]. 计算机系统应用, 2023, 32(10): 157-165.
- [10] Gao, X., Chen, L., Wang, K., Xiong, X., Wang, H. and Li, Y. (2022) Improved Traffic Sign Detection Algorithm Based on Faster R-CNN. *Applied Sciences*, **12**, Article 8948. <https://doi.org/10.3390/app12188948>
- [11] 王琳毅, 白静, 李文静, 等. YOLO 系列目标检测算法研究进展[J]. 计算机工程与应用, 2023, 59(14): 15-29.
- [12] Qi, Y., He, Y., Qi, X., Zhang, Y. and Yang, G. (2023). Dynamic Snake Convolution Based on Topological Geometric Constraints for Tubular Structure Segmentation. 2023 *IEEE/CVF International Conference on Computer Vision (ICCV)*, Paris, 1-6 October 2023, 6047-6056. <https://doi.org/10.1109/iccv51070.2023.00558>
- [13] Lau, K.W., Po, L. and Rehman, Y.A.U. (2024) Large Separable Kernel Attention: Rethinking the Large Kernel Attention Design in CNN. *Expert Systems with Applications*, **236**, Article ID: 121352.

<https://doi.org/10.1016/j.eswa.2023.121352>

- [14] Zhang, J., Zou, X., Kuang, L.D., *et al.* (2022) CCTSDB 2021: A More Comprehensive Traffic Sign Detection Benchmark. *Human-Centric Computing and Information Sciences*, **12**, Article 23.
- [15] 史涛, 崔杰, 李松. 优化改进 YOLOv8 实现实时无人机车辆检测的算法[J]. 计算机工程与应用, 2024, 60(9): 79-89.