

基于改进 YOLOv8 的自动驾驶交通标志检测算法

郑凯东¹ 戴典珍¹

ZHENG Kaidong DAI Dianzhen

摘要

随着自动驾驶技术的迅速发展,准确且高效的交通标志检测成为确保行车安全的关键,为确保自动驾驶车辆在复杂道路环境中对交通标志有高效的识别能力,文章提出了一种基于改进 YOLOv8 的交通标志检测算法。在 Head 部分的每个 C2f 层之后引入注意力机制 CBAM,其次,利用动态非单调聚焦系数优化的 WIoU 损失函数,通过给小目标赋予更高的权重,更好评估小目标的分割结果,以此提升对交通标志的检测准确性。实验证明,相比 YOLOv8 模型,改进后的 YOLOv8-CW 模型精确度提高 1.8%,MPA50-95 提升了 2.4%。

关键词

深度学习;自动驾驶;卷积块注意模块;聚焦系数;WIoU 损失函数

doi: 10.3969/j.issn.1672-9528.2025.06.003

0 引言

随着自动驾驶技术的发展,目标检测任务对于感知模块、环境理解以及决策支持等有关键作用,交通标志检测作为其核心组件之一,对于提高道路安全性和驾驶效率具有重要意义,提供了关键的导航信息、法规指示和警告信息,准确识别这些标志对于自动驾驶系统至关重要。然而,由于交通环境的复杂性和多样性,如光照变化、天气条件、遮挡物以及标志的磨损和褪色等因素,交通标志检测成为一项具有挑战性的任务。近年来,深度学习技术在计算机视觉领域的突破性进展为交通标志检测带来了新的机遇。特别是卷积神经网络(CNN)在图像识别任务中的成功应用,为交通标志检测提供了强大的工具。

目标检测作为一种学习目标,主要任务是输入图像类别和图像中目标对象位置的信息,其结果是使用边界框来识别目标对象。在深度学习和人工智能有效发展的趋势下,计算机识别技术和图像处理在深度学习领域已经成为模型的发展。当前,深度学习领域的物体检测模型主要分单阶段(One-Stage)和两阶段(Two-Stage)^[1]。

其中,Two-Stage 是基于分类的算法,通过搜索两个方面来识别事物并划分问题:选择候选区域和分类并发展候选位置。R-CNN 模型是检测算法的基石,Girschick 等人^[2]使用 SPP-NET 作为主要模型改进 R-CNN 算法,并通过实现和上传卷积共享对象算法提供快速的 R-CNN。2016 年,Dai 等人^[3]提出了基于 FAST-R-CNN 的 R-FCN 算法,并最终利用图像位置符号来提高特征位置信息,提高检测算法的精确度。

后来,由于池化后的 RoIPooling 变形和搜索确定性的降低,He 等人^[4]在 2017 年提出了 Mask R-CNN 算法。在大多场景中,应用检测到的数据量较为庞大,为适应实时处理的要求,检测系统必须在准确率和处理效率之间取得平衡。双目标阶段的检测算法,由于引入了 RPN 机制,因此精度上可以得到保证,但是在速度上却较慢。为解决此类问题,部分研究人员开始思考在结构上做简化,后研究发现可将其视为一个两端直接进行的回归算法,这样既能够获得目标物的位置,又提取了特征信息。与双阶段模型比较而言,这种单阶段算法比双阶段的处理过程更为简化,在精度上也优于其他算法。

2016 年,Huang 等人^[5]创新性地提出了 YOLO(you only look once)目标检测算法,该算法作为首个单阶段检测方法,在检测速度和精度方面均展现出优异性能。其模型结构的关键设计参照于 MultiGrasp^[6],并且 YOLO 在此基础上添加了分类机制,改变了 LOSS 函数,这种改进机制对于多组目标的检测任务能够顺利完成。YOLO 模型因其独特的设计和高效的性能,在目标检测领域占据了重要地位。基于自动驾驶领域的目标检测要求具有实时性和精准性,本文研究是在 YOLOv8 框架基础上进行优化,提升交通标志识别系统的检测效能。通过引入 CBAM 模块来增强模型的特征表示能力,并采用 WIoU 损失函数来优化训练过程。

1 YOLOv8 算法

2023 年,YOLOv8 算法由 Ultralytics 公司开发,YOLOv8 模型相比于此前发布的模型,在 Backbone 部分进行了明显改进,由 C2f 模块代替之前的 C3 模块,卷积层的数量减少到两个,且增加了更多的跳层连接,帮助模型更好

1. 西安石油大学计算机学院 陕西西安 710000

地捕捉不同层次的特征信息，从而提升模型的轻量化和梯度信息获取能力。在 Neck 部分，采用 PANet 结构，通过双向路径融合不同层次的特征。YOLOv8 在头部设计中采用分离式架构，实现了分类任务与检测任务的独立处理。在简化模型结构的同时提高了检测精度和速度。

YOLOv8 算法通过对多项技术创新和优化，大幅提升了目标检测的性能和灵活性。如在网络结构、损失函数、训练策略和推理过程等方面的改进，使其在各种应用场景中表现出色。由于其高效的计算和优异的性能，YOLOv8 在自动驾驶^[7]、智能监控^[8]、工业检测^[9]等多个领域有着广泛的应用前景。基于自动驾驶的目标检测任务对实时性要求较高，因此本文选择轻量化的 YOLOv8n 模型作为基础算法。

2 基于 YOLOv8 改进的 YOLOv8-CW 算法

本文通过在 YOLOv8n 结构中融合 CBAM 注意力机制和 WIoU 损失函数，提升自动驾驶交通标志检测^[10]的性能和精度。在 Neck 部分，将 CBAM 注意力添加到每个 C2f 卷积后，使得网络能够更加有效地提取和保留重要特征，并使用 WIoU 损失函数替换原来的 CIOU 函数，使得通过对预测框和真实框之间的重叠区域进行加权处理，减少定位误差。图 1 为改进后的 YOLOv8-CW 模型结构图。

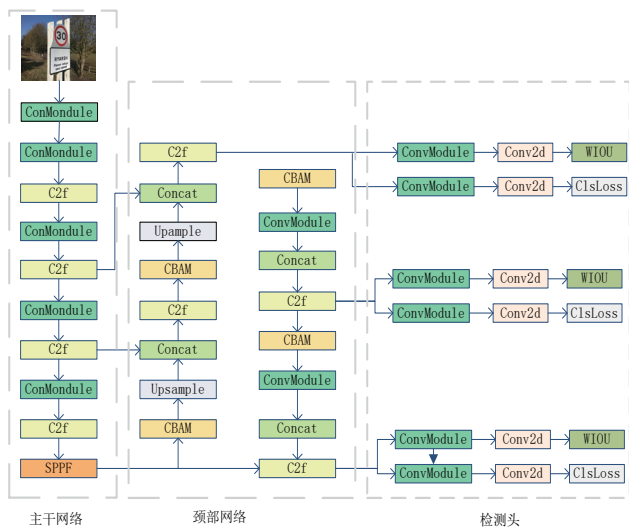


图 1 改进模型图

2.1 CBAM 模块

为解决卷积神经网络中不同通道和空间位置的特征表示，更加灵活地处理不同任务和场景下的问题，Woo 等人^[11]提出一种 CBAM 注意力机制，该模型通过同时引入通道注意力（CAM）和空间注意力（SA）两种不同的注意力机制，对输入特征图经过池化、MLP、Sigmoid 归一化等操作，动态地调整特征图的权重，最后通过注意力机制生成的特征图来完成对关键信息的关注。图 2 是 CBAM 模块的结构图。用

公式表示为：

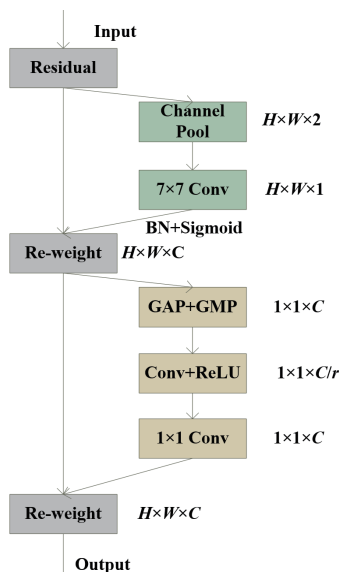


图 2 CBAM 模块结构图

$$M_c(F) = \sigma(\text{MLP}(\text{AvgPool}(F)) + \text{MLP}(\text{MaxPool}(F)))$$

$$M_c(F) = \sigma(W_1(W_0(F_{\text{avg}}^c)) + W_1(W_0(F_{\text{max}}^c))) \quad (1)$$

对输入特征图 F 分别执行平均池化和最大池化操作，所得结果分别输入多层感知机进行处理。其中， W_0 和 W_1 代表可训练的参数矩阵， $W_0(F_{\text{avg}}^c)$ 和 $W_0(F_{\text{max}}^c)$ 对应两条独立的处理路径。将两条路径的输出相加后，经 Sigmoid 函数激活，最终生成通道注意力权重 $M_c(F)$ 。

$$M_s(F) = \sigma(f^{7 \times 7}([\text{AvgPool}(F); \text{MaxPool}(F)])) \quad (2)$$

通过公式可以看出，通道注意力模块输出的特征图维度从 $H \times W$ 压缩至 1×1 。该特征图随后经过 7×7 卷积核和 Relu 激活函数进行降维处理，再通过 Sigmoid 函数进行归一化。最终，将归一化后的特征图与通道注意力输出特征图融合，生成空间注意力特征。

2.2 WIoU 损失函数

在交通标志检测中，精确的边界框定位至关重要，因为即使是轻微的误差也可能导致错误的决策或忽视重要的交通信息。且为满足在各种环境和天气条件下都能准确识别交通标志的需求。采用 WIoU (weighted intersection over union) 函数替代原有的 CIOU 损失函数。WIoU 损失函数是针对无人机航拍场景中目标检测任务设计的一种改进的损失函数，不仅能够减少低质量标签对梯度生成的不利影响，同时，通过提升预测框与真实框的重合程度，有效提升检测准确率。

WIoU 通过引入动态非单调聚焦系数 $\beta^{[12]}$ 来优化模型对小物体的检测性能，由基础损失函数 L_{WIoUv1} 、动态非单调聚焦系数和动态非单调 FM 组成。 L_{WIoUv1} 是 WIoU 的基础损失函数，由两部分组成：距离注意力分数 R_{WIoU} 和交并比损失

L_{IoU} ，如公式（3）所示。其中， R_{WIoU} 是基于边界框中心点距离的指数函数，如公式（4）所示； L_{IoU} 是交并比损失，定义为公式（5）。

$$L_{WIoUv1} = R_{WIoU} \times L_{IoU} \quad (3)$$

$$R_{WIoU} = \exp \left(\frac{\left(b_{cx}^{g'} - b_{cx} \right)^2 + \left(b_{cy}^{g'} - b_{cy} \right)^2}{c_w^2 + c_h^2} \right) \quad (4)$$

$$L_{IoU} = 1 - IoU \quad (5)$$

控制损失函数的聚焦强度由动态非单调聚焦系数 β 完成，计算公式为：

$$\beta = \frac{L_{IoU}^*}{L_{IoU}} \quad (6)$$

式中： L_{IoU}^* 是一个参考交并比损失值； L_{IoU} 是当前的交并比损失； β 的值域为 $[0, +\infty]$ 。

最后，通过应用 β ，WIoU 的损失函数进一步调整公式为：

$$L_{WIoU} = r \times L_{WIoUv1} \quad (7)$$

式中： r 是一个基于 β 的调节因子，用公式表示为：

$$r = \frac{\beta}{\delta \alpha^{\beta - \delta}} \quad (8)$$

式中： δ 和 α 是超参数，用于控制聚焦系数的变化。

3 实验结果分析

3.1 数据集

本文选择自动驾驶汽车的交通标志检测数据集，其中包括在不同场景下的交通标志，分为三大类：红绿灯、限速和停车。包含 15 个小类，分别是绿灯、红灯、限速 10 到 120 以及停车，共包含 4 969 张图像。将数据集按照 8:1:1 进行划分训练、测试和验证，部分数据集如图 3 所示。



图 3 数据集示例

3.2 评价指标

在计算机视觉领域，为更具体而全面地表达模型效果，使评价指标涵盖预测框的准确性、分类预测的精度、召回率以及整体性能的综合评估的多个方面。选择 Precision（精确

率）、Recall（召回率）、mAP50 和 mAP50-95 这 4 项指标作为结果评价。定义为：

$$AP = \int_0^1 P(r) dr \quad (9)$$

$$mAP = \frac{\sum_{i=1}^k AP_i}{K} \quad (10)$$

$$precision = \frac{True\ Positive}{True\ Positive + False\ Positive} \quad (11)$$

$$recall = \frac{True\ Positive}{True\ Positive + False\ Negative} \quad (12)$$

式中：AP 表示平均精度，通过计算 PR 曲线与坐标轴所围区域的积分面积得到。

3.3 实验结果和分析

本文进行交通标志检测的实验环境如表 1 所示，进行 100 次迭代，初始学习率为 0.001，权重衰减系数设置为 0.000 36，批量大小为 32。

表 1 实验环境

硬件 / 配置	具体参数
GPU	NVIDIA RTX 3060
OS System	Windows 11
Code language	Python 3.9

为验证本文改进的无人驾驶汽车交通标志检测算法模型的有效性，进行消融实验观察每个改进部分的实验结果，考虑到在驾驶场景中对交通标志的检测具有较高的实时性要求，因此初始模型选择为 YOLOv8n，后面改进模型部分均和初始模型保持相同的参数进行 100 次迭代实验。表 2 为 3 种模型在不同评价指标下的对比结果。

表 2 交通标志数据集上模型的评估指标

模型	CBAM	WIoU	Precision /%	Recall /%	mAP50 /%	mAP50-95 /%
YOLOv8n	×	×	95.72	85.87	93.80	80.88
YOLOv8n-CBAM	√	×	96.40	87.71	94.68	81.51
YOLOv8n-WIoU	×	√	94.27	89.32	94.76	81.62
YOLOv8n-CW	√	√	96.02	90.21	96.36	83.50

从消融实验来看，改进后的模型 YOLOv8n-CW 在不同的指标上均有提升效果，具体来说，该模型的准确率达到 96.02%，召回率为 90.21%，mAP50 为 96.36%，而 mAP50-95 为 83.50%。这些数据表明，YOLOv8n-CW 模型不仅在检测正例时的精确度较高，同时也能较好地覆盖所有的正例，减少了漏检的可能性。此外，其高 mAP50 值进一步证实了

模型在多种 IoU 阈值下都能保持优秀的性能, 展现了强大的泛化能力。尽管 mAP50-95 的值相对较低, 但仍保持在一个可接受的范围内, 说明即使在处理较为复杂的场景时, 模型也能维持一定的准确度。

4 结语

本文通过改进 YOLOv8 算法, 提出了一种基于 CBAM 注意力机制和 WIoU 损失函数的交通标志检测方法。通过在 Neck 部分的每个 Conv 后增加 CBAM 注意力机制, 使其能够自适应地调整特征图的权重, 并在各种环境和天气条件下都能准确识别交通标志。并采用 WIoU 损失函数^[13]来优化训练过程, 提升交通标志中较小尺寸目标的检测准确性, 减少定位误差。这对无人驾驶的视觉检测非常重要。

本文提出的模型与传统的 YOLOv8 模型相比, 改进后的模型在评估指标下精确度提高 1.8%, MPA50-95 提升了 2.4%。这表明, YOLOv8n-CW 模型不仅在检测正例时的精确度较高, 同时减少了漏检的可能性。未来的研究将继续探索如何进一步优化模型结构、提高实时性和部署效率、增强模型在多场景下的适应性问题^[14]。

参考文献:

- [1]DU L, SUN Y Q, CHEN S, et al. A novel object detection model based on faster R-CNN for spodoptera frugiperda according to feeding trace of corn leaves[J]. Agriculture, 2022, 12(1):248.
- [2]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. Piscataway: IEEE, 2014:23-28.
- [3]DAI J F, LI Y, HE K M, et al. R-FCN:object detection via region-based fully convolutional networks[C]//NIPS'16: Proceedings of the 30th International Conference on Neural Information Processing Systems. New York: ACM, 2016: 379-387.
- [4]HE K M, GKIOXARI G, DOLLÁR P, et al. Mask R-CNN[EB/OL].(2018-01-24)[2024-06-12].<https://doi.org/10.48550/arXiv.1703.06870>.
- [5]HUANG K, TIAN C W, SU J Y, et al. Transformer-based cross reference network for video salient object detection[J]. Pattern recognition letters, 2022, 160:122-127.
- [6]YANG Z G, ZHANG M Z, ZHANG Z M, et al. Lecture information service based on multiple features fusion[C]//2020 International Conference on Service Science (ICSS). Piscataway: IEEE, 2020: 24-26.
- [7]CAO J, ZHANG T, HOU L, et al. An improved YOLOv8 algorithm for small object detection in autonomous driving[J]. Journal of real-time image processing, 2024, 21:1-16.
- [8]BAKIRCI M. Utilizing YOLOv8 for enhanced traffic monitoring in intelligent transportation systems (ITS) applications[J].Digital signal processing, 2024, 152:104594.
- [9]TONG L, FAN C L, PENG Z B, et al. WTBD-YOLOv8: an improved method for wind turbine generator defect detection[J]. Sustainability, 2024, 16(11):4467.
- [10]SHARMA T, CHEHRIA, FOFANA I, et al. Deep learning-based road object detection for collision avoidance in autonomous driving[C/OL]//2024 IEEE World Forum on Public Safety Technology (WFPST). Piscataway:IEEE, 2024[2024-10-15]. <https://ieeexplore.ieee.org/document/10606983>.DOI:10.1109/WFPST58552.2024.00020.
- [11]WOO S, PARK J, LEE J Y, et al. CBAM: convolutional block attention module[C]//Computer Vision—ECCV 2018: 15th European Conference. Berlin: Springer, 2018:3-19.
- [12]WANG S Q, WU P Y, WU Q Q. Safety helmet detection based on improved YOLOv7-tiny with multiple feature enhancement[J/OL].Journal of real-time image processing, 2024 [2024-09-26].<https://link.springer.com/article/10.1007/s11554-024-01501-0>.
- [13]HU D A, YU M, WU X Y, et al. DGW-YOLOv8: a small insulator target detection algorithm based on deformable attention backbone and WIoU loss function[J]. IET image processing, 2024, 18(4): 1096-1108.
- [14]LIU J H, REN P. Applications and challenges of computer vision in autonomous driving[J/OL].Applied mathematics and nonlinear sciences, 2024[2024-11-12]. <https://sciendo.com/article/10.2478/amns-2024-1549>.DOI:10.2478/amns-2024-1549.

【作者简介】

郑凯东 (1964—), 男, 广东汕头人, 硕士, 副教授、硕士生导师, 研究方向: 图形学与虚拟现实、深度学习与计算机视觉、程序设计。

戴典珍 (2000—), 女, 甘肃庆阳人, 硕士研究生, 研究方向: 计算机视觉、深度学习。

(收稿日期: 2025-02-13 修回日期: 2025-06-10)