

基于改进 YOLOv8 的轻量型车辆目标检测算法

刘荣欣¹ 卢胜男¹ 刘晓天¹

LIU Rongxin LU Shengnan LIU Xiaotian

摘要

针对现有交通监控场景下车辆目标检测算法参数多、计算量大,难以在资源有限的设备中部署的问题,提出一种基于 YOLOv8 改进的轻量型车辆目标检测算法 GSE-YOLO。结合 Ghost 卷积技术,设计一种轻量型特征提取模块 C2fGhostv2,在减少计算负担的同时保证良好的特征提取能力。在颈部网络,引入 SA (shuffle attention) 注意力机制,主动选择合适的特征图权重凸显重要特征信息,减少背景对车辆检测的干扰。引入新的损失函数 EIOU,解决边界框的纵横比模糊问题,提高预测框精度。实验结果表明,在交通数据集 UA-DETRAC 上, GSE-YOLO 在检测精度没有损失的情况下,相较于原始 YOLOv8 参数量降低 36.11%,计算量降低 29.21%,更适合在计算量有限的边缘设备上部署,具有实用价值。

关键词

深度学习; 轻量型; 车辆目标检测; Ghost 卷积; shuffle attention 注意力机制; 损失函数

doi: 10.3969/j.issn.1672-9528.2024.06.019

0 引言

车辆目标检测是智能交通领域的重要研究课题,也是计算机视觉的重要研究方向。随着深度学习的发展,基于深度学习的车辆目标检测算法迅速崛起。与传统算法相比,基于深度学习的目标检测算法可以从数据中自主学习特征,不依赖人工,因此检测能力更强、泛化性更好。

目前基于深度学习的目标检测算法主要分为以 R-CNN^[1]、Faster R-CNN^[2] 为代表的双阶段算法和以 YOLO 系列为代表的单阶段算法。对于交通监控场景中车辆目标检测任务实时性的要求, YOLO 系列算法因其检测速度快的特性而备受关注。田智慧等人^[3]针对小尺度车辆检测效果差的问题,在 YOLOv3 的基础上使用 EfficientNet 作为特征提取网络,并使用 4 个特征尺度融合深层特征信息,以提升检测小尺度车辆的准确率。赵璐璐等人^[4]对 YOLOv5s 进行优化,提出了一种改进的车辆目标检测算法,在主干网络中引入 SE 注意力模块,并采用焦点损失函数 (focal loss) 解决正负样本不平衡的问题,有效地改善了密集车辆的误检和漏检情况。周飞等人^[5]针对交通监控场景,提出一种改进 YOLOv8 算法,采用 FasterNet 替换原有骨干网络,在骨干和颈部引入 SimAM 注意力机制,并添加小目标检测头,显著提高小目标检测能力。

以上研究通过将整个骨干网络替换为其他复杂网络来提升检测性能,或者通过添加小检测头、增加特征尺度来提高模型性能。虽然在检测精度上有所提升,但模型的参数量

和复杂度也随之大幅增加,难以在移动或资源有限的设备中部署,不满足交通监控场景实时性的要求。因此,本文提出 YOLOv8 的改进算法: GSE-YOLO。通过引入 Ghost 卷积技术,修改特征提取网络并设计出 C2fGhostv2 模块,以替换原始 C2f 特征提取模块,融合 SA (shuffle attention) 注意力机制,将原始损失函数 CIOU 替换为 EIOU。通过一系列改进,使原始模型轻量化,在减少参数和计算量的同时,保证检测精度。

1 改进 YOLOv8 检测模型

YOLOv8 是 YOLO 系列最新的模型,以深度和宽度为标准分为多个模型,模型参数量和计算量随着精度的提升大幅度提高。本文选择 YOLOv8n 作为基线模型,其参数量和计算量少且精度良好,满足交通监控场景的实时性要求。结构分为输入端、主干网络 (Backbone)、特征融合网络 (Neck) 和输出端 (Head) 组成。

输入端使用 Mosaic 数据增强的方式,对输入图片进行随机缩放、随机裁剪和随机排列,不但丰富了检测的数据集,还随机增加了很多小目标对象,增强了网络的鲁棒性^[6]。Backbone 由 CBS 模块、C2f 模块和 SPPF 模块组成。CBS 模块对输入信息执行卷积操作。C2f 模块使模型的梯度流更加丰富。SPPF 模块通过依次连接三个最大池化层来减少计算量,具有较低的延迟,适合对实时性要求较高的应用场景。Neck 部分对提取到的特征进行融合,使用了 FPN 和 PAN 的特征金字塔结构。Head 部分使用解耦头结构,缓解了分类和定位任务之间存在的冲突。

1. 西安石油大学 陕西西安 710065

基于 YOLOv8 改进轻量型车辆目标检测算法: GSE-YOLO。首先, 将 Ghostv2Bottleneck 结构集成到原始 C2f 模块中, 设计出轻量型特征提取模块 C2fGhostv2, 再将网络中的普通卷积 (不包括第一个) 替换为 Ghost 卷积以减少模型复杂度。其次, 在 Neck 部分进行特征融合前, 添加 SA 注意力模块, 使重要特征信息在融合时占据主要地位。最后, 将算法的 CIOU 边界框回归损失函数替换为 EIOU, 以提高检测框的精度, 加快收敛。GSE-YOLO 的整体网络结构如图 1 所示。

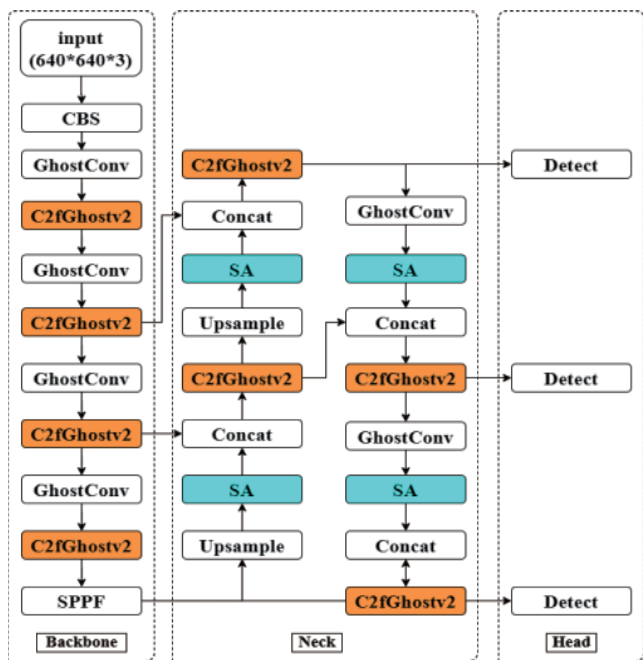


图 1 GSE-YOLO 网络结构

1.1 轻量型特征提取模块 C2fGhostv2

交通监控场景中, 车辆目标检测的实时性至关重要。然而 YOLOv8 模型的特征提取主要依赖 3×3 卷积运算, 导致模型参数和计算量较大。因此, 本文根据 Ghost 卷积, 设计了轻量高效的 C2fGhostv2 特征提取模块替代 YOLOv8 中的 C2f 模块。

Ghost 卷积能在增加特征图数量的同时, 尽可能减少参数的使用。输出特征图由两部分组成: 第一部分通过常规卷积生成, 第二部分通过线性变换生成^[7]。Ghost 卷积与常规卷积的不同之处如图 2 所示。相较于传统的卷积方法, Ghost 卷积在保持模型性能的同时, 显著减少了参数量和计算量。Ghostv2Bottleneck 由 Ghost 卷积结合解耦全连接 (DFC) 注意力机制构建而成, 具有动态校准与捕捉长距离空间信息的能力。相比于 Bottleneck, Ghostv2Bottleneck 可以使用更少的参数获得更多的特征信息。C2fGhostv2 模块使用 Ghostv2Bottleneck 代替 C2f 模块中的 Bottleneck, 有效减少 Bottleneck 中普通卷积带来的冗余计算。在不削弱特征提

取能力的前提下, 减少参数量, 保证轻量化。具体结构如图 3 所示。

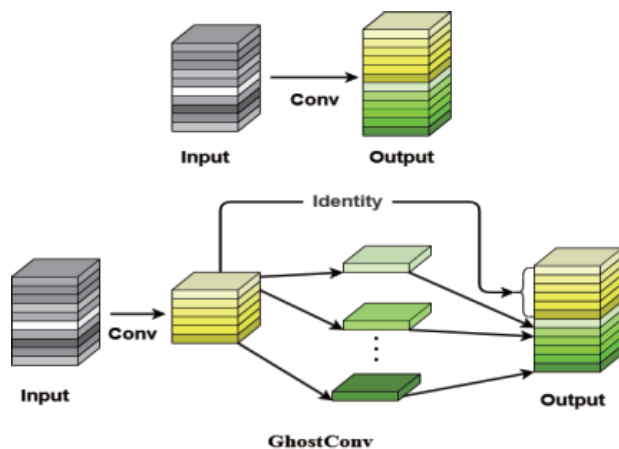


图 2 Ghost 卷积

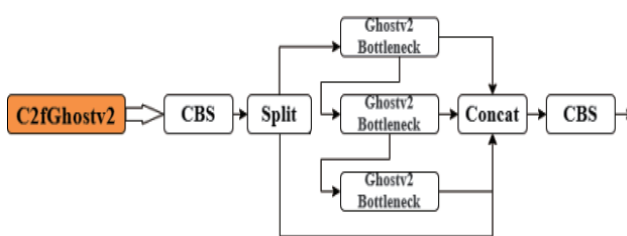


图 3 C2fGhostv2 模块

1.2 引入 SA 注意力机制

注意力机制是一种有助于网络模型学习输入信息重要性的方法, 包括空间注意力机制和通道注意力机制。SA (shuffle attention) 注意力机制引入了通道随机混合操作 (channel shuffle), 分块并行两种注意力机制, 使其高效组合减少计算量^[8]。具体结构如图 4 所示。

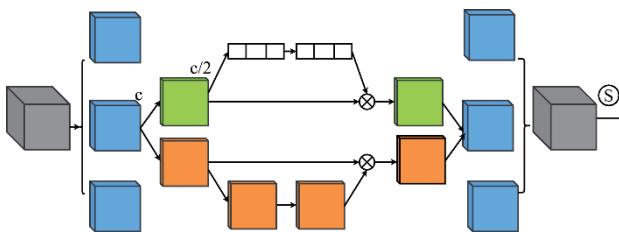


图 4 SA 注意力机制

首先将特征图分组, 接着将每一组都分为两部分, 分别使用空间和通道注意力机制; 然后将两部分特征图按通道数叠加; 最后将所有分组用通道随机混合操作, 得到输出特征图^[9]。在网络的瓶颈部分嵌入注意力机制, 可以有效地捕获位置信息和通道信息。因此, 为了弥补原始网络融合浅层信息不足的缺点, 本文在 Neck 部分进行特征融合前, 添加 SA 注意力模块。

1.3 改进损失函数

在目标检测中, 优秀的边界框回归损失函数可以显著提

升模型性能。YOLOv8 采用了 CIOU (complete intersection over union) 损失函数。相比于普通的 IOU, 它综合考虑多种影响因素如纵横比等, 因此能更好地调整预测框以匹配真实框。然而, 纵横比是相对值, 存在一定的模糊性, 回归效果不稳定。本文使用的 EIOU 损失函数在 CIOU 的基础上进行改进, 通过添加新的惩罚项, 很好地解决了上述问题。EIOU 的公式为:

$$\begin{aligned} L_{EIOU} &= L_{IOU} + L_{dis} + L_{asp} \\ &= 1 - IOU + \frac{\rho(b, b^{gt})}{(w^c)^2 + (h^c)^2} + \frac{\rho^2(w, w^{gt})}{(w^c)^2} + \frac{\rho^2(h, h^{gt})}{(h^c)^2} \end{aligned} \quad (1)$$

式中: w^c 和 h^c 表示预测框和真实框的最小外接矩形的宽度和高度, $\rho(b, b^{gt})$ 代表了预测框和真实框的中心点之间的欧氏距离^[10]。

EIOU 的惩罚项在 CIOU 的基础上进行了改进, 分别计算预测框和真实框的宽度和高度损失, 以减小宽高之间的差异, 从而加速模型的收敛。

2 实验结果与分析

2.1 实验环境与数据集

实验在 Windows10 操作系统上进行, 使用 Python3.10、CUDA11.1、PyTorch1.9.1 深度学习框架, GPU 为 NVIDIA GeForce RTX 3080ti。为验证本文的改进算法在车辆目标检测场景中的有效性, 实验采用统一的基线模型 YOLOv8, 输入图片尺寸为 640×640 , 采用 SGD 优化器, 初始学习率 0.01, 动量参数 0.937, 训练 100 轮, batchsize32。

本文所使用的数据集来自 UA-DETRAC 公开交通数据

集, 如图 5 所示。数据集包括多种天气场景和车辆类别。UA-DETRAC 数据集中同一视频下的图片数据相似度过高, 直接训练会导致数据冗余, 产生过拟合, 因此本文对原始数据集进行处理, 同一场景中的图片每 10 张提取一张, 得到 13 826 张图片作为本文的实验数据集。

2.2 评价指标

本文的实验结果评价指标为: 精确率 (P)、召回率 (R)、平均精度均值 (mAP)、模型参数量 (Params)、模型复杂度 (GFLOPs)。各指标的计算方法如下:

$$P = \frac{TP}{TP + FP} \quad (2)$$

$$R = \frac{TP}{TP + FN} \quad (3)$$

$$AP = \int_0^1 P(R) dR \quad (4)$$

$$mAP = \frac{\sum_{m=1}^M AP}{M} \quad (5)$$

式中: TP 表示正确识别的正样本数量, FP 表示将负样本错误地识别为正样本的数量, FN 表示将正样本错误地识别为负样本的数量^[11]。AP 代表 P - R 曲线下的面积, 其数值反映了分类器的性能。mAP 则是各类别 AP 的平均值, 用于整体评估分类器的性能。

2.3 实验结果分析

如图 6 所示, 经过大约 30 轮的训练, 原始模型与改进模型的损失值趋于稳定。实验结果表明, 改进模型 GSE-YOLO 相较于原始 YOLOv8 模型, 损失值更小、收敛更快。

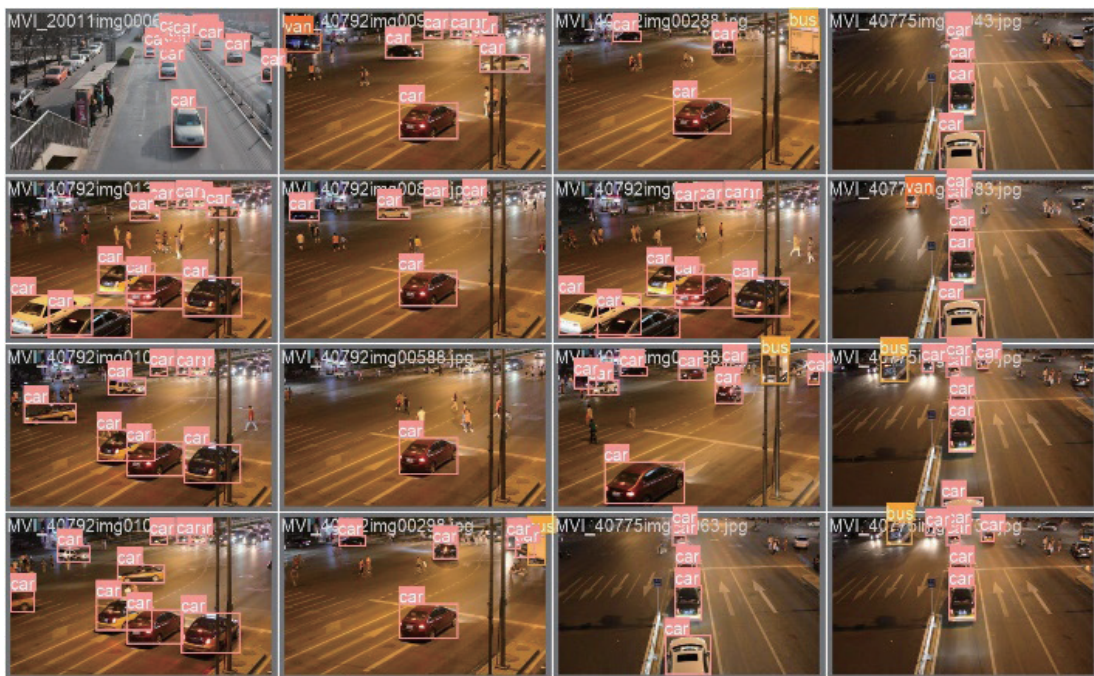


图 5 UA-DETRAC 交通数据集展示

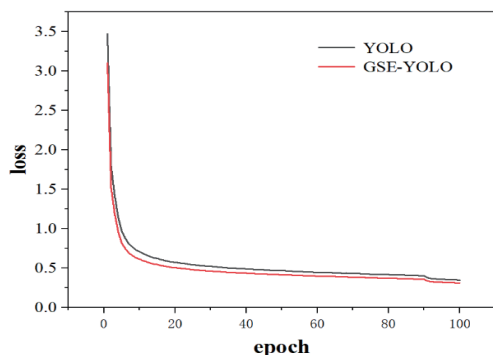


图 6 训练损失曲线

为验证改进后算法的有效性，根据所提出的改进模型设计消融实验，各项评价指标如表 1 所示。

表 1 消融实验结果

C2fGhostv2	SA	EIOU	P/%	R/%	mAP@0.5 /%	Params /MB	GFLOPs /GB
—	—	—	95.84	94.16	98.12	3.15	8.9
√	—	—	95.38	92.11	96.83	2.01	6.3
√	√	—	95.73	94.15	98.13	2.03	6.3
√	√	√	96.12	95.13	98.54	2.03	6.3

由消融实验结果可知，使用 C2fGhostv2 模块替换原始 C2f 模块并引入 Ghost 卷积后，模型的参数量和计算量显著下降，但在 P、R、mAP 三项指标上分别下降了 0.46%、2.05%、1.29%，说明 C2fGhostv2 模块和 Ghost 卷积模块在轻量化方面十分优异，但在特征提取方面略有欠缺。因此，本文引入 SA 注意力模块和 EIOU 损失函数来弥补检测精度的损失。依次加入 SA 注意力模块和 EIOU 损失函数，模型性能逐步上升，证明了模型改进的有效性。GSE-YOLO 相较于原始 YOLOv8，在参数量下降了 36.11%、计算量下降了 29.21% 的情况下，仍然提高了模型的检测精度，充分证明了本文改进模型的优越性。

3 结语

针对基于交通监控场景的车辆目标检测算法复杂度较高、计算量过大，难以在资源有限的设备中部署的问题，本文提出一种基于 YOLOv8 改进的轻量型车辆目标检测算法：GSE-YOLO。首先，在原有模型中引入轻量型特征提取模块 C2fGhostv2 和 Ghost 卷积，大幅降低参数量和计算量。然后，在 Neck 端添加 SA 注意力模块并将 CIOU 替换为 EIOU，提高检测精度。实验结果表明，与原始 YOLOv8 相比，改进算法参数量下降了 36.11%，计算量下降了 29.21%，实现了车辆目标检测模型的轻量化改进，满足交通监控场景的要求。未来的车辆目标检测研究应当集中在提高检测精度和速度的同时，应对复杂场景和多样化车辆的挑战。此外，基于端到端的无监督或弱监督学习方法也可能成为研究热点，

以减少对大量标注数据的依赖，推动车辆目标检测技术的实际应用。

参考文献：

- [1]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,2014:580-587.
- [2]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.
- [3] 田智慧, 杨奇文, 魏海涛. 基于改进的 YOLOv3 多目标小尺度车辆检测算法研究 [J]. 计算机应用与软件, 2023, 40(12): 169-175.
- [4] 赵璐璐, 王学营, 张翼, 等. 基于 YOLOv5s 融合 SENet 的车辆目标检测技术研究 [J]. 图学学报, 2022, 43(5): 776-782.
- [5] 周飞, 郭杜杜, 王洋, 等. 基于改进 YOLOv8 的交通监控车辆检测算法 [J]. 计算机工程与应用, 2024, 60(6): 110-120.
- [6] 张利丰, 田莹. 改进 YOLOv8 的多尺度轻量型车辆目标检测算法 [J]. 计算机工程与应用, 2024, 60(3): 129-137.
- [7]HAN K, WANG Y, TIAN Q, et al.GhostNet:more features from cheap operations[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition,[v.1].Piscataway: IEEE, 2020:1577-1586.
- [8]YANG Y.SA-Net:shuffle attention for deep convolutional neural networks[C]//2021 IEEE International Conference on Acoustics,Speech and Signal Processing,[v.1].Piscataway: IEEE, 2021:2235-2239.
- [9] 郭玉彬. 基于视频图像的车辆检测跟踪及行为识别研究 [D]. 北京: 北京交通大学, 2021.
- [10]ZHANG Y, REN W, ZHANG Z, et al.Focal and efficient IOU loss for accurate bounding box regression[J].Neurocomputing, 2022, 506(Sep.28):146-157.
- [11] 郭柏璋, 牟琦, 冀汶莉. 融合注意力机制的 YOLOv5 深度神经网络杂草识别方法 [J]. 无线电工程, 2023, 53(12): 2771-2782.

【作者简介】

刘荣欣（1999—），男，河南洛阳人，硕士研究生，研究方向：计算机视觉、目标检测。

卢胜男（1982—），女，江苏徐州人，博士，硕士生导师，研究方向：图像处理及机器学习、大数据技术应用。

刘晓天（1999—），男，山东济南人，硕士研究生，研究方向：模式识别、计算机视觉。

（收稿日期：2024-04-08）