

DOI:10.19651/j.cnki.emt.2415664

面向密集场景的多目标车辆检测算法^{*}

霍爱清 郭岚洁 冯若水

(西安石油大学电子工程学院 西安 710065)

摘要: 目标检测可为自动驾驶车辆提供附近目标的位置、大小和类别,但是密集场景中多目标检测仍然存在漏检、误检问题,为此该文提出了一种AD-YOLOv5车辆检测模型。首先,利用轻量型结构CBAM注意力机制对特征提取网络中的C3模块进行了优化得到C-C3模块,提高了对特征信息的获取能力,降低了对其他特征的关注度;其次,在检测头部分对分类和回归任务进行解耦,以实现更强的特征表达;然后,利用广义幂变换对IoU进行转换操作,提出鲁棒性更好的Alpha-IoU损失函数,提升了模型的检测精度并加快模型的收敛速度;最后,采用GridMask数据增强技术,增加了样本的复杂性,并在处理后的数据集上进行了实验。实验结果表明,改进后的目标检测模型的平均精度均值达到72.72%,与原YOLOv5模型相比提高了2.25%,且模型具有较高的收敛速度,通过可视化对比实验,直观展示了本文模型在密集场景能有效避免误检、漏检现象。

关键词: 目标检测;密集场景;YOLOv5算法;Alpha-IoU损失函数;CBAM;双检测头

中图分类号: TN919.8;TP391 **文献标识码:** A **国家标准学科分类代码:** 520.6040

Multi-objective vehicle detection algorithms for dense scenes

Huo Aiqing Guo Lanjie Feng Ruoshui

(College of Electronic Engineering, Xi'an Shiyou University, Xi'an 710065, China)

Abstract: Although object detection can provide the location, size and category of nearby targets for autonomous vehicles, there are still problems of missed detection and false detection in multi-object detection in dense scenes, so an AD-YOLOv5 vehicle detection model is proposed. Firstly, the C3 module in the feature extraction network is optimized to obtain the C-C3 module using the lightweight structure CBAM attention mechanism, which improves the ability to acquire feature information and reduces the attention to other features; secondly, in the detection head section, the classification and regression tasks are decoupled in order to achieve stronger feature representation; then, the generalized power transform is used to perform the transformation operation on the IoU, and the Alpha-IoU loss function with better robustness is proposed, which improves the detection accuracy of the model and accelerates the convergence speed of the model; finally, to add to the complexity of the sample, the GridMask data enhancement technique was used and experiments were carried out on the processed dataset. The experimental results show that the mean average accuracy of the improved target detection model reaches 72.72%, which is 2.25% higher than the original YOLOv5 model, and the model has a high convergence speed, and the visual comparison experiments intuitively show that the model of this paper can effectively avoid the phenomenon of misdetection and omission detection in dense scenes.

Keywords: target detection;dense scene;YOLOv5 algorithm;Alpha-IoU loss function;CBAM;double-head

0 引言

人工智能技术的发展与成熟为安全驾驶提供了可能,在自动驾驶领域,多目标检测研究具有重要的应用价值和研究意义。

Duan等^[1]提出CornerNet检测方法,从多个关键点进行检测可以提高检测精度和召回率,但不能有效检测小目标信息;在此基础上,提出了一些新的目标检测算法,例如:DETR^[2]、Swin Transformer^[3]、SSD^[4]等;Wang等^[5]提出改进的YOLOv4-Tiny算法,利用K-means聚类算法为数

收稿日期:2024-03-20

*基金项目:陕西省科技厅一般项目(2020GY-152)资助

据集生成适当的锚点,并且用改进的非最大抑制(non-maximum suppression, NMS)算法来筛选预测框,虽然提高了对目标定位的精度,但平均检测精度并未得到改善; Zhang 等^[6]提出改进的 Nanodet-Plus 模型,将扩张卷积和高效的多尺度注意力机制(excitation modulation attention, EMA)引入模型,增强了特征层之间的相关性提升了算法速率,但检测精度有所降低。欧阳继红等^[7]提出改进的 YOLOv4 算法,通过修改颈部的网络结构以此改进算法的检测精度,但网络结构复杂度高计算量大。Yao 等^[8]将 Transformer 并入 YOLOv5 算法来优化算法的骨干参数,有效提高恶劣天气下目标的检测和识别,但增加了模型参数数量会导致检测效率变低; Shi 等^[9]提出一种改进的网络模型 YOLOv5s_2E,将 Soft_NMS 与 EIoU 相结合,替代原网络的 NMS,提高对被遮挡物体的检测,但 EIoU 损失函数的计算复杂度很高降低了算法的效率; Yang 等^[10]提出 KPE-YOLOv5 算法,其采用 K-means++聚类算法重新设计锚框大小,引入 scSE 注意力模块,以解决分辨率低、特征不明显问题,但模块只关注了通道维度上的注意力无法准确识别特征信息; 张银胜等^[11]、Wang 等^[12]、Yu 等^[13]研究人员都对 YOLOv5 模型进行了相应改进,以期提升网络对特征信息的表达能力,提高了目标检测的精度值; Huang 等^[14]提出 EST-YOLOv5s 模型,在 C3 模块中集成有效信道注意力机制(efficient channel attention, ECA),并将瓶颈模块替换为 Swin Transformer Block,改善了小目标漏检问题,但 ECA 机制忽略了空间维度上的特征信息导致信息容易丢失; Li 等^[15]提出改进的 YOLOX 算法,将骨干网络替换为 CSPCrossLayer 模块且在骨干网络的输出后添加置换注意力机制(shuffle attention, SA)模块,增强对特征的提取能力和融合能力,但没有对损失函数进行改进使得模型的鲁棒性较低。

综上,尽管改进后的 YOLOv5 检测算法在训练速度和推理时间方面有很大的改善,但忽略了在特征提取时对空间维度注意力的处理、对损失函数的优化以及均未对检测头中回归和分类问题进行拆解,导致密集场景中的应用仍存在对车辆漏检和误检的问题。针对上述问题,本文基于 YOLOv5 算法,提出 AD-YOLOv5(Alpha-IoU loss with decoupled head YOLOv5, AD-YOLOv5)检测算法。具体改进如下:

1)由于密集场景的多目标车辆检测中,获取到的图像存在检测目标重叠性高和占比小的问题,因此引入 CBAM 注意力机制模块,从通道维度和空间维度上对特征信息进行提取,以获取更完整的特征信息。

2)在对信息进行识别时,引入双检测头模块,对分类和回归任务进行拆解。分类任务分析检测目标的局部信息,回归任务分析检测目标的全局信息,从而提高目标的检测精度。

3)利用 Alpha-IoU 损失函数替换原损失函数,引入 β

权衡参数,规避真实框和预测框长宽比相等时惩罚项恒为 0 导致损失函数性能降低的问题,以此提高算法的收敛速度,使模型具有更好的鲁棒性。

4)针对密集场景的数据集数量少且图像信息简单的问题,引入 GridMask 数据增强技术,丰富数据集内容,使实验结果更具说服力。

1 AD-YOLOv5 算法

本文将从特征提取网络、检测头和损失函数 3 个方面进行优化改进。

1.1 特征提取网络改进

由于密集场景车辆被遮挡后,车身在图像中所占的比重减小,以及特征信息在浅层网络更易丢失,为此在网络特征图的每一个位置上添加卷积块注意力机制模块^[16](convolutional block attention module, CBAM),从而提升网络的特征表达能力,为实现更准确的物体检测用改进后的 C-C3 模块替换原 YOLOv5 模型中的 C3 模块,图 1 为 C-C3 模块结构图。

如图 1 所示,输入的特征图进入 C-C3 模块后,将会分为两路,右路先经过 ConvBNSiLU 卷积层再经过由 n 个 BottleNeck 串联的倒残差层,左路只经过一个 ConvBNSiLU 卷积层,再将两路叠加,经过一个 ConvBNSiLU 卷积层。C-C3 结构中的 ConvBNSiLU 模块均为 1×1 卷积核,起到改变维度和减少参数的作用。BottleNeck 由两个 ConvBNSiLU 模块组成,第 1 个 ConvBNSiLU 为 1×1 卷积核,可以缩减一半的通道数,第 2 个为 3×3 卷积,作用是将通道数翻倍。降低维度使卷积核提升理解特征信息的能力,升高维度使卷积核获得更多的特征信息。最后,将处理后的特征图输入 CBAM 注意力机制模块。

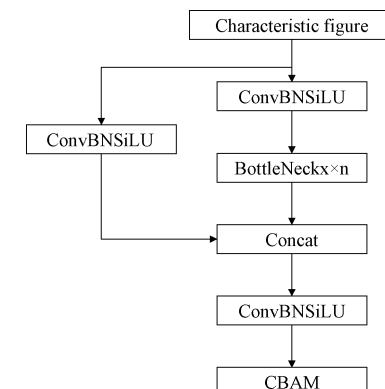


图 1 C-C3 模块结构图

如图 2 所示,CBAM 是由特征通道(channel)和特征空间(spatial)组成的注意力机制。具体而言,单个图像特征层输入到通道注意力模块(channel attention module, CAM)后,首先进行全局最大池化和全局平均池化以获取

两个特征向量,这两个特征向量随后被输入到共享的全连接层,处理后将两者相加的结果经过 Sigmoid 激活函数使特征图每个通道的权重归一化到 0~1,将归一化后的权重和输入特征图相乘。接着,将 CAM 的输出传递给空间注意力模块(spatial attention module, SAM),SAM 以通道为单位进行最大池化和平均池化,并将两者的结果进行特征联合再通过一个卷积得到 $1 \times w \times h$ 的特征图。最终,通过一个 sigmoid 函数,得到包含 SAM 的特征图。在两个维度上进行注意力分配机制增强了模型性能的提升效果。实现过程如式(1)和(2)所示。

$$F' = M_c(F) \otimes F \quad (1)$$

$$F'' = M_s(F') \otimes F' \quad (2)$$

式中: F 为输入特征图; F' 为经通道注意力处理后的特征图; F'' 为经空间注意力处理后的特征图; $M_c(F)$ 为特征通道模块; $M_s(F')$ 为特征空间模块。

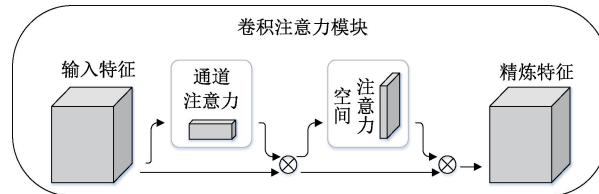


图 2 CBAM 注意力机制模块流程图

1.2 检测头改进

在 YOLOv5 模型中,检测头模块的作用是对网络提取的特征图进行多尺度目标检测。该模块使用耦合的权重同时执行分类和回归任务,即在一个 1×1 卷积层中完成。然而,由于分类和回归任务之间的耦合,可能导致任务在执行时发生冲突,降低密集场景下对目标车辆检测的准确率。为此通过解耦分类和回归任务,将它们作为独立的分支进行计算,以有效提高检测效果。

经过解耦操作后的双检测头示意图如图 3 所示。输入特征图经过 1×1 卷积层进行降维,减少通道数至 256。本文提出的解耦头设计是将分类和回归分离,为了减少通道数为 128,两条并行的分类和回归分支经过两个平行的 3×3 卷积层,分别提取类别特征和位置特征,然后两条分支经过 1×1 卷积层完成分类和回归任务,这样可以提高目标检测的准确性。对检测头进行解耦操作,可以降低由分类和回归之间产生的差异所导致的预测偏差,有效提升模型的检测精度。

1.3 损失函数优化

损失函数用来评价模型的预测值和真实值不一样的程度。在目标检测中,损失函数的使用主要是在模型的训练阶段,损失函数越小,模型的性能越好。

1) YOLOv5 算法损失函数

YOLOv5 模型中采用了完整交并比损失函数^[17](complete intersection over union,CIoU)。CIoU 损失函数考虑了真实框和预测框中心点间的距离以及两个框的最小

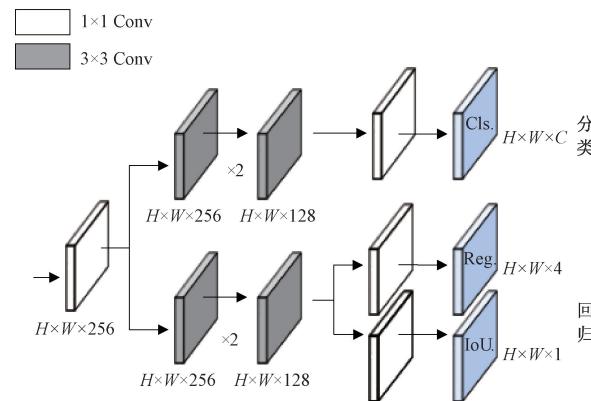


图 3 双检测头示意图

包裹框的对角线距离,引入了衡量长宽比的一致性参数,以更精准地度量真实框和预测框之间的相似性。CIoU 损失函数的计算方法相比传统 IoU 损失函数更为复杂,但这也使得模型在训练过程中能够更好地理解真实框和预测框的准确位置和形状,其表达式如式(3)、(4)所示。

$$\left\{ \begin{array}{l} CIoU = IoU - \frac{\rho^2(B, b)}{c^2} - \alpha v \\ \alpha = \frac{v}{1 - IoU + v} \\ v = \frac{4}{\pi^2} (\arctan \frac{w^{gt}}{h^{gt}} - \arctan \frac{w}{h})^2 \end{array} \right. \quad (3)$$

$$LossCIoU = 1 - CIoU = 1 - IoU + \frac{\rho^2(B, b)}{c^2} + \alpha v \quad (4)$$

式中: IoU 为真实框和预测框交集与并集的比值; $\rho^2(B, b)$ 为真实框和预测框中心点的距离; c 为真实框和预测框最小外接矩形的对角线距离; α 为权重参数; v 为真实框和预测框长宽比例差值的归一化参数。

由式(3)和(4)可知,如果真实框和预测框的长宽比是相同的,那么长宽比的惩罚项恒为 0,会对损失函数的性能产生较大影响。公式中 w 和 h 是一对相反数,也就是说, w 和 h 不能同时增大或减小,这显然是不合理的。为此,本文提出了 α 幂指化交并比损失函数(alpha intersection over union,Alpha-IoU)。

2) Alpha-IoU Loss

Alpha-IoU Loss 对 CIoU 损失函数中的 IoU 项和惩罚项进行转换操作,目的是针对不同交并比损失函数目标的损失和梯度实现动态调整。利用广义幂变换进行转换操作,解决回归分析和方差分析中的数据不满足正态分布的问题。Alpha-IoU 在原损失函数的基础上,不仅加强对 IoU 值目标的关注度,而且调节用于广义幂变换的 α 值可以提高模型的准确度和可靠性。

原 YOLOv5 中损失函数的公式为 $L_{IoU} = 1 - IoU$,在原函数的基础上添加 α 幂化指标并进行广义幂变换,变换后的表达式为:

$$L_{\alpha-IoU} = \frac{(1 - IoU^{\alpha})}{\alpha}, \alpha > 0 \quad (5)$$

通过选择不同的 α 参数值, Alpha-IoU Loss 可以呈现多种形式变体,因此可以将其简化为式(6)。

$$L_{\alpha-IoU} = \begin{cases} -\log(IoU), & \alpha \rightarrow 0 \\ 1 - IoU^{\alpha}, & \text{其他} \end{cases} \quad (6)$$

为保证 α 参数的有效性 α 通常不趋近于 0,引入惩罚项 P ,得到式(7),并且 $\alpha_1 > 0, \alpha_2 > 0$,其中 α_2 对结果影响不大,因此令 $\alpha_2 = \alpha_1$ 。

$$L_{\alpha-IoU} = 1 - IoU^{\alpha_1} + P^{\alpha_2}(B, b^{gt}) \quad (7)$$

综上,Alpha-IoU Loss 损失函数的表达式为式(8)、(9), β 是权衡参数,目的是权衡长宽比例造成的损失和 IoU 部分造成的损失的平衡因子。

$$L_{\alpha-IoU} = 1 - IoU^{\alpha_1} + \frac{P^{\alpha_2}(B, b^{gt})}{c^{\alpha_2}} + (\beta v)^{\alpha} \quad (8)$$

$$\beta = \frac{v}{(1 - IoU) + v} \quad (9)$$

式中: IoU 为真实框和预测框交集与并集的比值;

$\frac{\rho^{\alpha_2}(B, b^{gt})}{c^{\alpha_2}}$ 为防止过拟合的正则项; α 为幂化指标; v 为真实框和预测框长宽比例差值的归一化参数。

1.4 AD-YOLOv5 算法的整体结构

利用改进后的 C-C3 模块、双检测头模块和 Alpha-IoU 损失函数替换原 YOLOv5 模型中对应的模块,提出了 AD-YOLOv5 改进算法,算法的网络结构如图 4 所示。

如图 4 所示,AD-YOLOv5 算法使用的主干特征提取网络为 CSPDarknet^[18],采用原网络中的 Focus 结构和 SPP 结构。首先,利用本文提出的 C-C3 模块替换原网络中的 C3 模块,且加入 CBAM 注意力机制负责对通道和空间上的特征权重进行调整,可以更好的提取出特征图中的特征信息。通过这样的处理,模型可以更加精准地突出有效目标,从而提高检测的准确性。其次,解耦回归任务与分类任务,对目标类别和位置进行精细化预测,降低因权重共享而引起的检测错误。最后,利用 Alpha-IoU 损失函数替换原模型的 CIoU 损失函数,不仅能够更准确地评估目标框的质量,还能够提高模型在复杂场景下的性能。

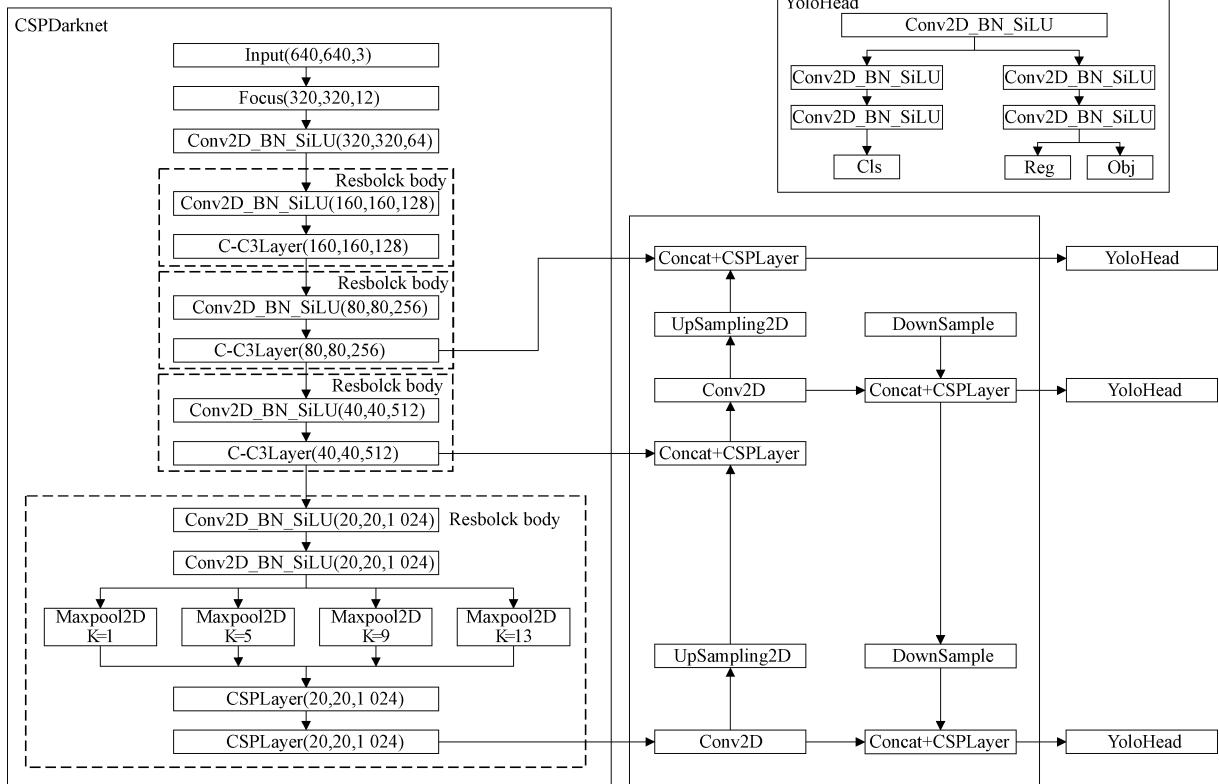


图 4 AD-YOLOv5 算法网络结构图

2 数据增强

本文研究的问题是建立在密集场景下的车辆检测,由于可供模型使用的公交车数据集数量较少,且数据集中多数车辆未被遮挡,导致数据集过于简单。因此,对数据集

进行 GridMask 数据增强^[19],目的是调整形状、大小和遮盖数量等参数控制增强效果,防止出现数据集单一和训练过程中出现过拟合问题。

常用的数据增强方法有随机擦除^[20](random erasing, Reras)、隐藏寻找^[21](hide-and-seek, HaS)、Cutout^[22]等,这

些方法都会随机选择掩盖区域,导致出现可判别区域全部删除或者全部保留现象。不同信息擦除增强方法的对比效果如图 5(a)所示,GridMask 的改进效果如图 5(b)所示。其原理是生成结构化的格子,然后随机移除一个不连续的格子图像信息,从而确保只会部分遮盖图像,避免了对重要信息进行完全遮盖或完全不遮盖的情况。因此,本文使用 GridMask 增强方法对数据集进行数据增强。

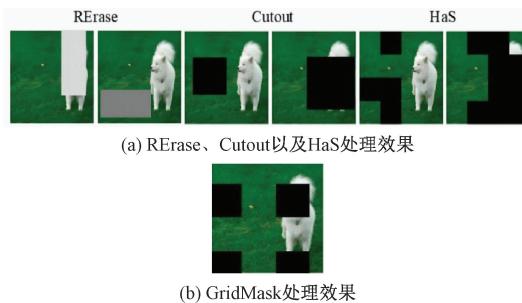


图 5 数据增强处理效果

GridMask 方法利用有规律的正方形区域对图片进行掩码,设置不同参数值,可确保仅部分区域被遮盖,从而降低了完全遮挡重要信息的可能性,其计算公式为:

$$\tilde{x} = x \times M \quad (10)$$

式中: $x \in R^{H \times W \times C}$ 为输入图像; $M \in \{0, 1\}^{H \times W}$ 为应丢弃像素的二进制掩码; \tilde{x} 为处理后的结果。

当 $M=1$ 时表示输入图像的像素会被保留,否则会被丢弃。如图 6 所示,被删除区域的密度和大小由 4 个参数(r, d, δ_x, δ_y)控制, r 是一个单元中灰色区域短边的比例, d 决定了删除区域的大小, δ_x 和 δ_y 是删除区域的边长。

3 实验结果与分析

3.1 实验平台搭建

本文实验环境的配置和参数设置如表 1 所示。

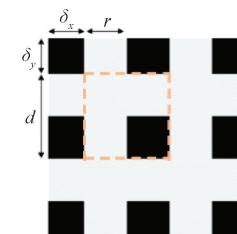


图 6 掩码示意图

表 1 试验环境的配置和参数设置

配置项目	配置参数
操作系统	Windows10
CPU	Intel Core i5-10200H@2.4 GHz
GPU	INVIDIA GeForce GTX 1650Ti
编程语言	Python3. 8
深度学习框架	Pytorch1. 7. 1
编程环境	NVIDIA CUDA11. 0、Cudnn8. 0. 5

3.2 检测结果与分析

1) 改进算法消融实验

本文在对 YOLOv5 网络模型的改进过程中,引入了轻量级的 CBAM 注意力机制、解耦头以及 Alpha-IoU 损失函数,目的是提升算法性能和优化模型结构。为了验证 3 个改进模块对原 YOLOv5 模型的提升效果,设计了 7 组消融对比实验。为保证消融实验结果的可靠性,所有实验项的环境配置均与表 1 中的设置一致,并在同等的置信度、学习率和训练周期下进行。

消融实验结果如表 2 所示,从参数量(parameter/M)、浮点计算次数(giga floating-point operation per second, GFLOPs/G)以及平均精度均值(mean average precision, mAP/%)三方面对结果进行评估。

表 2 消融实验结果

模型	GridMask	C-C3	Alpha-Loss	Decoupled Head	P/M	GFLOPs/G	mAP/%
YOLOv5					6.98	15.74	68.79
方法 1	✓				6.98	15.74	70.53
方法 2	✓	✓			7.16	13.88	70.91
方法 3	✓		✓		6.98	15.74	71.18
方法 4	✓			✓	13.29	18.24	71.09
方法 5	✓		✓	✓	13.44	18.22	71.95
AD-YOLOv5	✓	✓	✓	✓	13.70	17.55	72.67

由表 2 可知:对数据集进行方法 1 数据增强后使平均精度均值提升了 1.74%。后续的方法 2~6 均以数据增强

为基础,所有平均精度均值的提升值均与方法 1 进行比较。方法 2 中,利用 C-C3 模块替换原模型中的 C3 模块,

使得 mAP 值提高了 0.38%，浮点计算次数损失了 1.86 G；方法 3 优化了模型中的损失函数，用 Alpha-Loss 损失函数替换原有 CIoU-Loss 损失函数，使 mAP 值提升了 0.65% 且参数量和浮点计算次数未发生改变；方法 4 利用解耦操作后的双检测头替换原有检测头，虽然参数量增加了 6.4 M，但 mAP 值获得了 0.56% 的提升；方法 5 中同时将 Alpha-Loss 损失函数和解耦头引入模型中，mAP 值增大了 1.42%。

本文改进的 AD-YOLOv5 模型的参数量达到 13.70 M，mAP 值达到了 72.66%，相较于原 YOLOv5 模型，参数量仅增大了 6.72 M，但平均精度值增加了 2.14%。

2) 对比实验与分析

(1) 损失函数对比实验

原 YOLOv5 模型的损失函数曲线图如图 7(a) 所示，AD-YOLOv5 模型的损失函数曲线图如图 7(b) 所示。图 7 中的横轴表示一代训练，训练周期的范围设置为 0~80 个，纵轴表示损失值。由图 7 可知，两种模型的损失值都会先迅速下降，再逐渐趋于平稳。但可以观察到，AD-YOLOv5 减少的损失值大于原 YOLOv5 模型。与原 YOLOv5 模型相比，本文提出的 AD-YOLOv5 模型的回归过程具有更平稳且收敛速度更快的特点。

(2) 多种算法对比实验

本小节将 AD-YOLOv5 算法与其他 7 种常用算法进

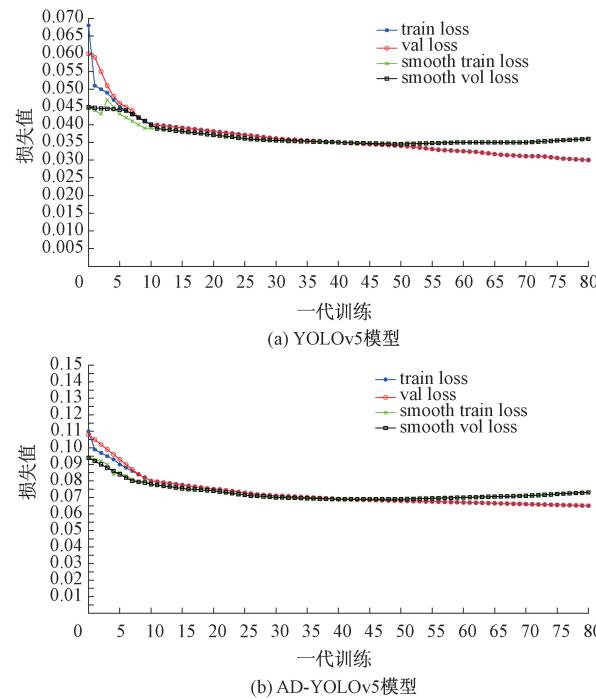


图 7 YOLOv5 和 AD-YOLOv5 模型损失曲线对比

行比较，验证 AD-YOLOv5 算法的有效性。实验结果如表 3 所示。

表 3 对比实验结果

目标检测算法	检测精度/%			mAP/%	FPS/帧
	car	bus	truck		
YOLOv4-tiny ^[5]	75.69	62.98	44.51	61.06	25
SSD ^[4]	76.11	65.21	45.43	62.25	18
NanoDet-plus ^[6]	78.98	68.75	48.55	65.42	29
YOLOv4 ^[7]	79.4	69.33	51.01	66.58	25
YOLOv5 ^[8]	82.65	72.89	55.89	70.47	33
CenterNet ^[1]	83.79	71.95	57.41	71.05	17
YOLOX ^[15]	83.98	72.75	56.87	71.2	32
AD-YOLOv5	84.52	74.23	59.41	72.72	32

由表 3 中的数据得知，本文提出的 AD-YOLOv5 模型的检测精度、mAP 和 FPS 相较于其他算法均有所提升。与轻量级的 YOLOv4-tiny、SSD 和 NanoDet-plus 模型相比 mAP 值得到了很大提高，分别为 11.66%、10.47% 和 7.3%。与 YOLOv4、YOLOv5、CenterNet 以及 YOLOX 模型相比 mAP 值平均提高了 2.89%。

综上，AD-YOLOv5 模型在实验中体现出更好的检测效果。由表 3 中显示的推理时间 FPS 值可以看到，该模型

具有较高的 FPS 值，能够达到实时检测的要求。

(3) 可视化对比实验

为了直观展示 AD-YOLOv5 算法的检测效果，图 8 显示了 YOLOv5 与 AD-YOLOv5 模型在多个不同场景（典型路面场景包括车型相似、夜间小目标、道路黑暗和车辆遮挡）下对车辆目标检测的可视化结果。

如图 8 所示，在对相似车型进行检测时两种模型均表现出较好的检测效果，但在车辆密集的黑暗环境和有车辆

遮挡的场景下,本文提出的 AD-YOLOv5 模型表现出较好的检测效果,原 YOLOv5 模型出现了漏检和错检现象。实验结果表明,该模型可以在多种场景中使用并具有较好的检测效果。



图 8 不同模型下的可视化结果

4 结 论

为解决密集场景下车辆检测准确度低的问题,基于 YOLOv5 模型的网络结构,提出使用含有 CBAM 注意力机制的 C-C3 模块,目的是获取更多更精准的特征信息;提出使用双检测头解耦检测过程中的分类和回归任务,目的是提升对特征信息的处理速度和准确度;提出使用 Alpha-IoU Loss 损失函数,目的是提升模型的鲁棒性和收敛性。使用被 GridMask 增强方法处理过的数据集进行实验。结果表明,在密集道路上的车辆检测过程中本文提出的 AD-YOLOv5 算法模型表现出较高的准确性,相较于原 YOLOv5 模型,检测精度提高了 2.25%,检测时的每秒帧数达到了 32 帧,达到了目标检测的实时检测要求。可视化实验进一步证实了所提出的算法具有可行性和有效性。

参考文献

- [1] DUAN K, BAI S, XIE L, et al. Centernet: Key point triplets for object detection[C]. Proceedings of the IEEE International Conference on Computer Vision, 2019.
- [2] CARION N, MASSA F, SYNNAEVE G, et al. End-to-end objection with transformers[C]. Proceedings of the European Conference on Computer Vision, 2020.
- [3] LIU Z, LIN Y, CAO Y, et al. Swin transformer: Hierarchical vision transformer using shifted windows[J]. ArXiv preprint arXiv:2103.14030, 2021.
- [4] 孙备,左震,吴鹏,等.面向无人艇环境感知的改进型 SSD 目标检测方法[J].仪器仪表学报, 2021, 42(9): 52-61.
- [5] WANG L, ZHOU K, CHU A, et al. An improved light-weight traffic sign recognition algorithm based on YOLOv4-Tiny[C]. IEEE Access, 2021.
- [6] ZHANG Z, DENG A, CAO X. Remote sensing ship detection based on feature fusion [C]. 2023 5th International Conference on Frontiers Technology of Information and Computer(ICFTIC), 2023.
- [7] 欧阳继红,王梓明,刘思光.改进多尺度特征的 YOLO_v4 目标检测方法[J].吉林大学学报(理学版), 2022, 60(6): 1349-1355.
- [8] YAO J, FAN X, LI B, et al. Adverse weather target detection algorithm based on adaptive color levels and improved YOLOv5[J]. Sensors, 2022, 22: 8577.
- [9] SHI T, DING Y, ZHU W. YOLOv5s_2E: Improved YOLOv5s for aerial small target detection[J]. IEEE Access, 2023, 11: 80479-80490.
- [10] YANG R, LI W, SHANG X, et al. KPE-YOLOv5: An improved small target detection algorithm based on YOLOv5[J]. Electronics, 2023, 12: 817.
- [11] 张银胜,杨宇龙,胡宇翔,等.改进视觉注意力网络的水下目标检测算法[J].国外电子测量技术, 2023, 42(12): 132-143.
- [12] WANG Y, ZHAO L, MA Y, et al. Multiscale YOLOv5-AFAM-Based infrared dim-small-target detection[J]. Applied Sciences, 2023, 13: 7779.
- [13] YU Z, LEI Y, SHEN F, et al. Application of improved YOLOv5 algorithm in lightweight transmission line small target defect detection [J]. Electronics, 2024, 13: 305.
- [14] HUANG M, YAN W, DAI W, et al. EST-YOLOv5s: SAR image aircraft target detection model based on improved YOLOv5s [J]. IEEE Access, 2023, 11: 113027-113041.
- [15] LI C, XU A, ZHANG Q, et al. Steel surface defect detection method based on improved YOLOX [J]. IEEE Access, 2024, 12: 37643-37652.

- [16] OOI Y K, IBRAHIM H, MAHYUDDIN M N. Enhanced dense space attention network for super-resolution construction from single input Image[J]. IEEE Access, 2021, 9: 126837-126855.
- [17] LI Y, ZHANG J, HU Y, et al. Real-time safety helmet-wearing detection based on improved yolov5 [J]. Computer Systems Science and Engineering, 2022, 43(3): 1219-1230.
- [18] 解宇敏,张浪文,余孝源,等.可见光-红外特征交互与融合的YOLOv5目标检测算法[J/OL].控制理论与应用,2024;1-9. <http://kns.cnki.net/kcms/detail/44.1240.TP.20230511.1643.024.html>.
- [19] LI P, ZHAO A, FAN Y, et al. Research on underwater robust object detection method based on improved YOLOv5s [C]. 2023 IEEE International Conference on Mechatronics and Automation(ICMA), 2023.
- [20] ZHONG Z, ZHENG L, KANG G, et al. Random erasing data augmentation [J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 34(7): 13001-13008.
- [21] SINGH K K, LEE Y J. Hide-and-seek: Forcing a network to be meticulous for weakly-supervised object and action localization[C]. Proceedings of the 2017 IEEE International Conference on Computer Vision, 2017.
- [22] DEVRIES T, TAYLOR G W. Improved regularization of convolutional neural networks with cutout[J]. ArXiv preprint arXiv: 1708. 04552, 2017.

作者简介

霍爱清(通信作者),博士,教授,主要研究方向为导向钻井控制、人工智能与深度学习等。

E-mail:aqhuo@xsyu.edu.cn

郭岚洁,硕士研究生,主要研究方向为深度学习与计算机图像处理等。

E-mail:1349314399@qq.com

冯若水,硕士,主要研究方向为人工智能与目标检测等。

E-mail:410907643@qq.com