

文章编号:1671-6833(2025)05-0001-08

基于改进 YOLOv7 的高风险区工程车辆识别算法

张震¹, 肖宗荣², 李友好³, 黄伟涛³

(1. 郑州大学 电气与信息工程学院, 河南 郑州 450001; 2. 郑州大学 计算机与人工智能学院, 河南 郑州 450001; 3. 河南汇融油气技术有限公司, 河南 郑州 450001)

摘要:为解决高风险区域工程车辆施工时对天然气管道的安全威胁问题,特别是重型车辆可能带来的物理冲击与环境干扰,提出一种基于改进 YOLOv7 的工程车辆识别算法。以 6 种施工现场常见的自卸车、压路车、搅拌机、叉车、挖掘机和装载机等车型为研究对象,利用自定义数据集进行训练,数据集涵盖多种环境和角度的图像,确保模型效能。首先,在 YOLOv7 头部网络中引入了 CBAM 注意力机制并在最大池化层结构中增加了改进的 GAM 注意力机制,提升模型对关键图像特征的关注度,从而提高目标检测的准确性;其次,采用 DySample 动态上采样器替换最近邻插值上采样模块,提高检测精度;最后,提出了一种改进的 SPPCSPC 模块,提高特征提取效率,降低计算成本,加速推理过程。这些改进使得模型在图像质量低、目标距离远等挑战下仍能维持高检测精度。实验结果表明:所提算法在自定义工程车数据集上的精确度 P 、召回率 R 、 $mAP@0.5$ 、 $mAP@0.5:0.95$ 分别为 97.7%、94.7%、98.6%、90.4%;与 YOLOv7 算法相比, P 、 R 、 $mAP@0.5$ 、 $mAP@0.5:0.95$ 分别提升了 1.3 百分点、1.4 百分点、1.4 百分点、3.7 百分点。

关键词:高风险区;工程车辆;YOLOv7;注意力机制;上采样器;特征提取

中图分类号:TP391

文献标志码:A

doi:10.13705/j.issn.1671-6833.2025.02.019

近年来,随着建筑工地和工程场地数量的不断增加,工程车辆在施工现场的活动也变得更加频繁。尽管工程车辆的使用提高了施工效率,但同时也带来了一些安全隐患和风险。尤其是在高风险区域,如建筑物密集的地方,人员聚集的区域,以及设备众多、作业空间狭窄的场所,工程车辆的操作可能对施工现场造成严重威胁。因此,通过对施工现场工程车辆的实时监测和预警,可以有效地降低工程车辆可能带来的风险和损失,为施工现场的安全施工提供有力保障^[1]。

深度学习目标检测算法通过卷积神经网络自动提取目标的高级语义特征,在检测性能上表现出色。当前主流的深度学习目标检测算法主要分为两类:两阶段目标检测算法和一阶段目标检测算法^[2-3]。两阶段目标检测算法主要是先生成候选框,再对这些候选框进行进一步的分类和位置回归处理,通常能够获得更高的检测精度。施工场景目标检测研究中使用的两阶段目标检测算法有 R-CNN^[4]、Faster

R-CNN^[5]等。两阶段目标检测算法虽然可以获得良好的检测精度,但其检测速度较慢,无法满足车辆检测任务中实时性的要求。一阶段目标检测算法则是直接通过网络模型进行目标位置和类别的预测,具有快速推理和网络结构简单等优点。一阶段目标检测算法有 SSD^[6]、RetinaNet^[7]、YOLO 系列^[8-9]等。

基于上述研究背景,本文针对施工现场工程车辆检测的特殊需求,提出了一种改进的一阶段目标检测算法优化方案。本文以 YOLOv7 算法^[10]为基础网络模型,引入改进的全局注意力机制(global attention mechanism, GAM)和卷积块注意力模块(convolutional block attention module, CBAM),采用 DySample 动态上采样器替代最近邻插值上采样模块,并提出改进的结合跨阶段部分卷积的空间金字塔池化(spatial pyramid pooling with cross stage partial convolutions, SPPCSPC)模块,这些改进方法显著提升了算法在复杂施工环境下的检测目标精度和鲁棒性。

收稿日期:2024-11-02;修订日期:2024-12-20

基金项目:河南省重点研发专项项目(231111211600);河南省重大公益专项项目(201300311200)

作者简介:张震(1966—),男,河南郑州人,郑州大学教授,博士,博士生导师,主要从事计算机视觉研究, E-mail: zhangzhen66@126.com。

引用本文:张震,肖宗荣,李友好,等. 基于改进 YOLOv7 的高风险区工程车辆识别算法[J]. 郑州大学学报(工学版), 2025,46(5):1-8. (ZHANG Z, XIAO Z R, LI Y H, et al. Construction vehicles recognition algorithm based on improved YOLOv7 in high risk areas[J]. Journal of Zhengzhou University (Engineering Science), 2025,46(5):1-8.)

1 相关工作

在目标检测领域,尤其是在车辆检测方面,近年来涌现出了众多研究成果。这些研究主要致力于提升检测算法的精度、实时性和鲁棒性,以适应复杂多变的环境。对于二阶段目标检测算法,李松江等^[11]提出了改进的 Cascade RCNN 算法,通过优化特征金字塔网络,显著增强了算法对小目标和遮挡目标的检测能力;周康等^[12]改进了 Faster R-CNN 的特征金字塔网络,使多维度特征图的信息得到更充分的利用,从而提升了小目标的识别效果。这些研究展示了特征金字塔网络在目标检测中的重要作用。对于一阶段目标检测算法,李国进等^[13]引入了改进的 Inception 模块,这一创新显著提高了对小目标车辆的检测能力;胡习之等^[14]结合 Camshift 和 YOLOv4 算法,提出了一种新的车辆检测算法,该算法通过计算图像帧间的汉明距离,为实时车辆检测提供了新的思路;Zhang 等^[15]则针对无人机视角下的行人和车辆检测问题提出了 Drone-YOLO 算法,该算法在主干网络中采用 RepVGG 替换传统卷积,显著增强了网络学习多尺度特征的能力。这些方法在特征提取和识别精度方面取得了较好的效果,但在复杂环境下仍面临准确率和实时性不足的挑战。

在此基础上,一阶段目标检测算法中 YOLOv7 模型由于具有良好的性能和高效的检测精度使其成为车辆检测任务中一个不错的选择。其模型结构主要包括输入端(Input)、主干网络(Backbone)、头部

网络(Head)和输出端(Output) 4个部分。在输入端,YOLOv7 通过对图像进行缩放、填充、归一化、数据增强等处理,提高了图像的清晰度。在 Backbone 中,通过多层次的特征提取及 SPP^[16]、CSPNet^[17]等模块的组合,完成对输入图像的关键特征提取,确定了检测的准确性。在 Head 部分,模型将特征转换为目标类别、位置和置信度的高效输出,通过特征融合和锚框生成等操作,实现目标检测。在输出端,模型通过非极点最大值抑制(NMS)^[18]等方法优化检测结果,满足了实际应用的需求。

本文以 YOLOv7 为基础网络模型,提出了改进 YOLOv7 算法。首先,由于施工现场背景复杂,在检测时容易受到环境干扰的影响,引入了 GAM 注意力^[19]模块和 CBAM 注意力^[20]模块,提升模型对图像中关键区域的关注,从而提高检测精度;其次,针对最近邻插值上采样模块在处理高分辨率图像时可能导致细节丢失问题,采用 DySample 动态上采样器^[21]对其进行替换,能够更好地保留图像细节,进一步提升精度检测;最后,为了避免特征图过度平滑,提出改进的 SPPCSPC 模块,在保持感受野的同时,减少计算量和参数量,提高了模型的推理速度和检测效果。改进后的 YOLOv7 模型在精确率、召回率和综合性能方面均取得了显著提升,验证了本文方法在实际应用中更高的可靠性和更好的实用性。

2 改进的 YOLOv7 模型

改进的 YOLOv7 网络结构如图 1 所示。

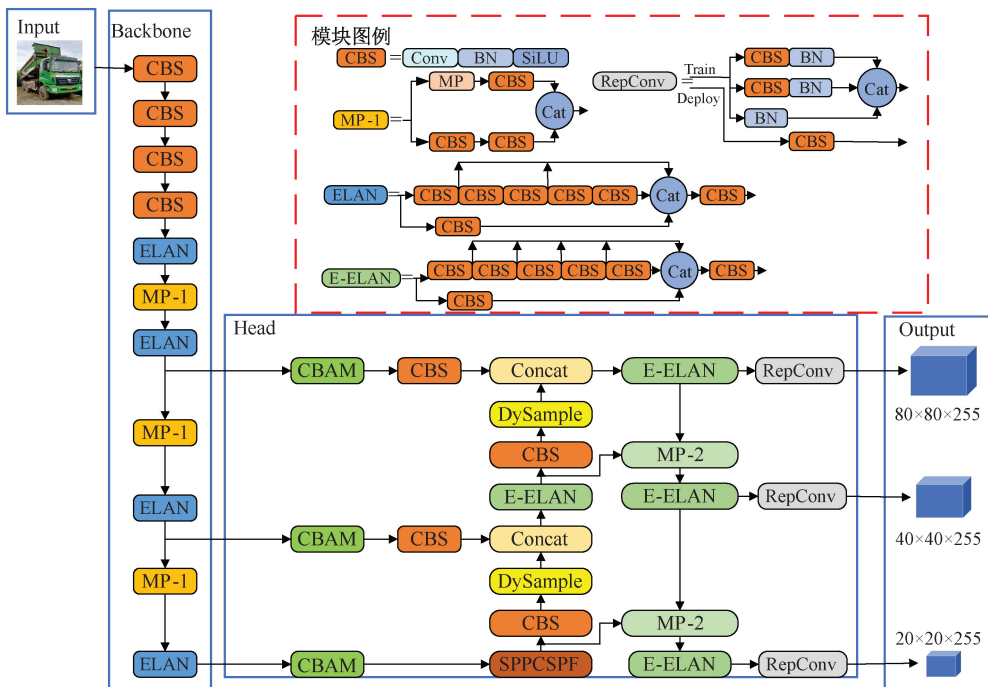


图 1 改进的 YOLOv7 网络结构

Figure 1 Structure of the improved YOLOv7 network

2.1 注意力机制

在高风险区域的工程车辆检测中,将注意力机制添加到 YOLOv7 模型中能够显著增强特征表示能力,提高小目标的检测性能,提升对复杂背景的鲁棒性,优化特征融合,并降低计算成本。其通过关注关键区域和特征,抑制无关背景的干扰,使模型在目标检测任务中更加精确、可靠和高效,从而实现准确的检测效果。本文在 Head 网络中引入了 CBAM 注意力机制,并在 MP 模块中增加了改进的 GAM 注意力机制,改进后的 MP 模块结构如图 2 所示。CBAM 机制结合了通道和空间注意力,其结构如图 3 所示。CBAM 注意力机制包含两个核心部分:通道注意力模块和空间注意力模块。通道注意力模块通过全局平均池化和全局最大池化生成不同的特征描述,并通过共享的全连接层和 Sigmoid 函数生成通道注意力映射,以重新加权输入特征。在空间注意力模块中,通过全局池化生成单通道特征描述,并经过卷积和 Sigmoid 生成空间注意力映射,重新加权空间特征。通过这一方法,CBAM 能够动态调整特征的通道和空间权重,提升模型的表情能力,增强了模型对关键特征的关注,从而提高目标检测的准确性和鲁棒性。

从而使模型能够更有效地学习并提取特征,使模型在处理不同类型和规模的目标时表现更加稳定,从而提高目标检测的准确性。

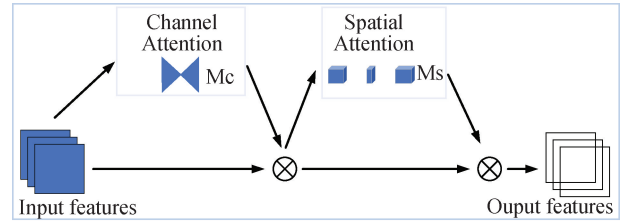


图 4 GAM 结构

Figure 4 Structure of GAM

2.2 DySample 动态上采样器

DySample 动态上采样器能够根据输入特征图的上下文信息自适应调整采样策略,生成更加细致和精确的上采样结果。相比传统的最近邻插值上采样方法,DySample 动态上采样器在以下方面展现出显著优势:更好地保留目标细节和边缘特征;有效减少特征图的信息损失和几何失真;显著提升模型在复杂背景和小目标检测任务中的性能表现。因此,DySample 不仅具备更高的准确性,还拥有更强的鲁棒性,有助于提升模型在实际应用中的性能。

在目标检测任务中,DySample 通过上采样操作调整输入特征图的大小,使其与原始图像的尺寸相匹配,从而实现对不同尺度和空间分布目标的检测。传统的上采样方法通常依赖双线性插值^[22],但这些方法存在局限性,会导致关键图像细节丢失。此外,传统的基于内核的上采样过程需要大量的计算和参数开销,不利于网络架构的轻量化。在拍摄的施工场景中,工程车图像相对较小,有时会出现像素失真等问题,导致识别任务中的细粒度细节丢失和特征学习困难。为了应对这一问题,本文采用了 DySample 动态上采样器,这是一种高度轻量且有效的动态上采样器,旨在增强对低分辨率图像或较小目标的检测能力,同时减少误报和漏报的情况。DySample 利用基于点的采样方法和学习采样的视角进行上采样,完全避免了耗时的动态卷积操作和额外的子网络。DySample 需要的计算资源更少,并且可以在不增加额外负担的情况下增强图像分辨率,从而以最小的计算成本提高模型的效率和性能。

DySample 的网络结构如图 5 所示。其采样集 δ 由原始采样网格 O 和生成的偏移 G 组成。偏移通过线性+像素洗牌方法生成,其中偏移的范围可以由静态和动态因素确定。具体来说,以静态因子采样方法为例,给定大小为 $C \times H \times W$ 的特征图和上采样因子 s ,特征图首先通过输入和输出通道

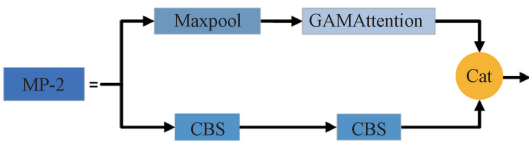


图 2 改进后的 MP 模块结构

Figure 2 Structure of the modified MP module

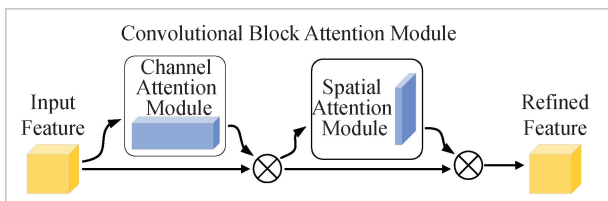


图 3 CBAM 结构

Figure 3 CBAM structure

GAM 注意力模块通过全局外围信息提升特征表达能力,主要包括全局信息提取和门控机制两个部分,GAM 结构如图 4 所示。通过全局平均池化和最大池化提取上下文信息,再利用门控机制重新加权通道特征,使网络聚焦重要特征,抑制相关信息,从而增强模型的表情能力。在 YOLOv7 的 MP 模块中加入 GAM 注意力机制,有助于提升目标检测的准确性和鲁棒性。对 GAM 进行了改进,将通道和空间注意力机制的激活函数从 ReLU 替换为 SiLU,

分别为 C 和 $2s^2$ 的线性层;其次,使用像素洗牌方法将其重塑为 $2g \times sH \times sW$;最后,生成尺寸为 $C \times sH \times sW$ 的上采样特征图。

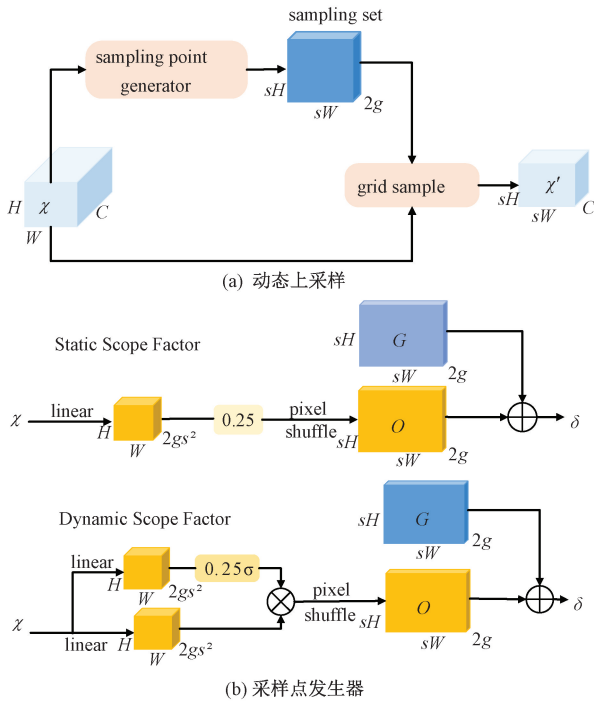


图5 DySample网络结构

Figure 5 Structure of DySample network

在图5中,输入特征、上采样特征、生成的偏移和原始网格分别表示为 X 、 X' 、 O 和 G 。在图5(a)中,采样集由采样点生成器生成,输入特征通过网格采样函数重新采样;在图5(b)中,采样集由生成的偏移与原始网格位置相加得到, σ 表示 sigmoid 函数, $2g$ 表示波形核的组数。

2.3 SPPCSPF 网络

为了减少使用的内存和计算资源,降低推理阶段的计算复杂度以及提升目标检测的准确性和模型的鲁棒性,将基准模型中的 SPPCSPC 网络^[23]改进为 SPPCSPF 网络,SPPCSPC 与 SPPCSPF 网络结构对比如图6所示。

在训练阶段,使用多支路残差结构进行特征提取;在推断阶段,将卷积层和批量归一化层合并,并将3个分支整合为单一路径。这样可以将所有训练得到的参数等效为一个 3×3 大小的卷积核层,从而提高推断速度并降低内存访问成本。

将原有的卷积模块替换为 RepConv,可以将训练过程中增加参数整合成一个更高效的卷积操作,从而减少训练过程中的内存占用,使得模型在推理阶段能够更快速地处理特征图,提升了检测速度和准确性。考虑到使用结构重参数化会增加训练成本,将 SPPCSPC 网络进行了一定的改进,

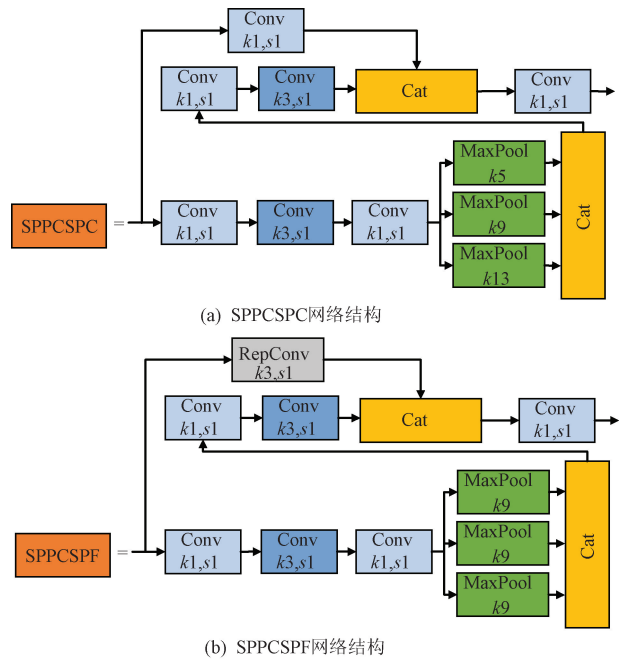


图6 SPPCSPC 与 SPPCSPF 网络结构对比
Figure 6 Comparison of network structure SPPCSPC and SPPCSPF

将原有的3个池化操作内核大小从 5×5 、 9×9 、 13×13 统一为3个 9×9 的最大池化层。这种设计不仅简化了池化操作,也减少了复杂度,提高了训练效率。

3 实验与结果分析

3.1 实验数据集

由于当前没有公开的工程车辆数据集,本文选择自定义数据集进行实验。其中图像的来源包括 robflow 网站、网上图片及视频等。本文使用的数据集共划分为训练集、验证集和测试集3部分,分别包含7303张、904张和885张图像,考虑到工地施工的工程车辆类别很多,将数据集划分为6个类别,分别是自卸车、压路车、搅拌车、叉车、挖掘机和装载车。每个类别中训练集、验证集、测试集中图像数量的比例大致为 $8:1:1$ 。

3.2 实验平台

本文利用 PyTorch 深度学习框架,在 Windows 11 家庭中文版 23H2 操作系统下,使用 12th Gen Intel(R) Core(TM) i7-12700H 2.30 GHz 处理器进行自定义数据集的训练和测试。显卡为 NVIDIA GeForce RTX 3060,显存为 6 GB,系统内存为 16 GB,CUDA 版本号为 11.7,PyTorch 版本号为 2.0.1。

实验中训练参数设置:输入图像尺寸为 640×640 ,batch size 为 4,epoch 设置为 100。

3.3 评价指标

本文采用精确率 P 、召回率 R 、 $mAP@0.5$ 和 $mAP@0.5:0.95$ 作为评价模型性能的主要指标。精确率 P 表示在所有被预测为正样本中实际为正样本的比例。召回率 R 表示在所有实际为正样本中被预测为正样本的比例。 P - R 曲线作为辅助评价指标,用于评估工程车辆检测模型的表现,展示了模型在不同分类阈值下的预测精度 P 和召回率 R 之间的关系。 mAP 代表所有类别的平均精度,其值是以召回率为横坐标、精确率为纵坐标绘制的 P - R 曲线下的面积。精确率 P 、召回率 R 、平均精度 AP 和平均精度均值 mAP 的计算公式如下:

$$P = \frac{TP}{TP + FP} \times 100\%; \quad (1)$$

$$R = \frac{TP}{TP + FN} \times 100\%; \quad (2)$$

$$AP = \int_0^1 P(R) dR \times 100\%; \quad (3)$$

$$mAP = \frac{\sum_{i=1}^n AP_i}{n} \times 100\%。 \quad (4)$$

式中: TP 表示被正确分类为正例的正样本数量; FP 表示被错误分类为正例的负样本数量; FN 表示被错误分类为负例的正样本数量; n 为数据集中数据类别数目。

3.4 实验结果与分析

为评估所提算法的有效性,在自定义数据集上进行了消融实验,保持其他参数指标一致,以验证本

文方法对模型检测性能的影响。改进点消融实验如表 1 所示,采用控制变量法逐步对 YOLOv7 模型进行改进,评价这些模块对提高模型性能的贡献。

表 1 中 Exp. 1 为原始 YOLOv7 算法的实验结果,作为后续 15 组实验的对比基准。经过表 1 中这些改进操作,各实验的部分评价指标相较于基准实验都有一定的提升,添加所有改进操作后得到 Exp. 16,相较于基准实验的 P 、 R 、 $mAP@0.5$ 和 $mAP@0.5:0.95$ 分别提升了 1.3 百分点,1.4 百分点,1.4 百分点和 3.7 百分点,进一步说明了本文算法的有效性。

为进一步验证本文算法的有效性,将其与 Faster R-CNN^[5]、SSD^[6]、YOLOv5^[24]、YOLOv7^[10]、YOLOv7-tiny^[25] 以及 YOLOv8^[26] 进行比较,并将实验结果进行整理,不同算法检测结果对比如表 2 所示。

由表 2 可知,本文算法在自制数据集上进行检测, $mAP@0.5$ 和 $mAP@0.5:0.95$ 分别可达 98.6% 和 90.4%。相比于其他模型,本文算法在 R 、 $mAP@0.5$ 和 $mAP@0.5:0.95$ 均取得了较好的表现。

为了更好地验证本文算法对于工地工程车辆的检测性能,从测试集上选取部分图像进行测试,YOLOv7 与本文算法检测结果对比如图 7 所示。

高风险区域是指可能对天然气管道造成物理冲击和环境干扰的施工场所,其特征包括重型车辆的闯入、施工活动的复杂性等。虽然数据集来自多个现场,但筛选时重点选择高风险特征明显的图像。针对待测图像,划定了可调整的自定义高风险区域,

表 1 改进点消融实验

Table 1 Improved point ablation experiments

序号	GAM	CBAM	DySample	SPPCSPF	P	R	$mAP@0.5$	$mAP@0.5:0.95$
Exp. 1					96.4	93.3	97.2	86.7
Exp. 2	√				97.2	93.7	97.4	86.6
Exp. 3		√			94.9	94.6	97.4	87.1
Exp. 4			√		97.0	94.3	97.3	84.7
Exp. 5				√	96.1	94.4	98.2	89.5
Exp. 6	√	√			96.9	93.9	97.9	88.8
Exp. 7	√		√		97.8	93.3	97.3	86.8
Exp. 8	√			√	95.8	94.5	97.5	87.6
Exp. 9		√	√		96.4	93.4	97.4	87.3
Exp. 10		√		√	96.6	94.8	98.3	89.0
Exp. 11			√	√	98.1	91.1	96.3	82.4
Exp. 12	√	√	√		96.2	93.8	97.9	88.9
Exp. 13	√	√		√	97.8	94.2	97.5	88.0
Exp. 14	√		√	√	97.2	92.7	96.3	82.0
Exp. 15		√	√	√	97.6	94.4	98.1	88.4
Exp. 16	√	√	√	√	97.7	94.7	98.6	90.4

注:√表示使用相应方法。

图7中各图的 Detection Area 为自定义的划定风险区,只有该区域内的目标才会被检测到。通过图7中置信度的数值大小可以对模型识别的准确程度进行评估。由图7可知,在工地上各类车辆在相同图像下,相对于 YOLOv7 算法,本文算法的目标框具有更高的置信度,对小目标检测效果也表现得更好。原 YOLOv7 算法可能存在误检的情况,但本文算法却能以较高的置信度识别出检测目标,这也进一步验证了本文算法的检测表现优于原模型。

表2 不同算法检测结果对比

Table 2 Comparison of detection results of different algorithms %

算法	P	R	$mAP@0.5$	$mAP@0.5:0.95$
Faster R-CNN	70.2	92.8	93.1	69.3
SSD	86.4	88.0	92.5	57.4
YOLOv5	96.1	94.6	97.9	86.6
YOLOv7	96.4	93.3	97.2	86.7
YOLOv7-tiny	94.8	91.0	95.8	80.2
YOLOv8	98.0	94.6	97.9	89.4
本文算法	97.7	94.7	98.6	90.4

挖掘机与
自卸车检
测结果



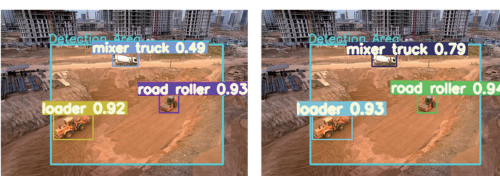
压路机检
测结果



叉车检测
结果



装载车、
搅拌机与
压路车检
测结果



(a) YOLOv7检测

(b) 本文算法

图7 YOLOv7与本文算法检测结果对比

Figure 7 Comparison of detection results of YOLOv7 and this algorithm

4 结论

本文以自卸车、压路车、搅拌机、叉车、挖掘机和装载车6种施工现场常见的车型为研究对象,设计了一种基于改进YOLOv7的高风险区工程车辆识别算法。该算法在特征提取和检测精度方面较原模型有所提升。改进过程中,在Head网络中引入CBAM注意力机制,在最大池化层添加改进的GAM注意力机制,增强了特征表达能力并提升了检测精度,同时也提升了其在复杂场景下的检测性能;在Head网络中将最近邻插值上采样模块替换成DySample动态上采样模块,使模型更有效地保存了细节和边缘信息,减少了信息丢失和失真,对于处理小目标和复杂场景表现更加出色;将原有的SPPCSPC网络改进成SPPCSPF网络,由于将原有卷积块替换成RepConv和统一池化内核大小,使得其在内存使用、计算资源和检测精度等方面都有一定提升。在相同实验环境下,本文算法与原YOLOv7相比,精确度 P 、召回率 R 、 $mAP@0.5$ 和 $mAP@0.5:0.95$ 分别提升了1.3个百分点、1.4个百分点、1.4个百分点和3.7个百分点。

通过实验验证,结果表明本文算法对划定区域的工程车辆表现出更好的检测能力,对小目标检测也能表现出更好的检测效果,且具有较高的精度和召回率,基本能够满足在高风险区对工程车辆进行实时检测的要求。

参考文献:

- [1] 胡皓,郭放,刘钊.改进YOLOX-S模型的施工场景目标检测[J].计算机科学与探索,2023,17(5):1089-1101.
HU H, GUO F, LIU Z. Object detection based on improved YOLOX-S model in construction sites[J]. Journal of Frontiers of Computer Science and Technology, 2023, 17(5): 1089-1101.
- [2] 张震,李浩方,李孟洲,等.改进YOLOv3算法与人体信息数据融合的视频监控检测方法[J].郑州大学学报(工学版),2021,42(1):28-34.
ZHANG Z, LI H F, LI M Z, et al. Video surveillance detection method based on improved YOLOv3 algorithm and human body information data fusion[J]. Journal of Zhengzhou University (Engineering Science), 2021, 42(1): 28-34.
- [3] 李柯泉,陈燕,刘佳晨,等.基于深度学习的目标检测算法综述[J].计算机工程,2022,48(7):1-12.
LI K Q, CHEN Y, LI J C, et al. Survey of deep learning-based object detection algorithms[J]. Computer En-

- gineering, 2022, 48(7): 1-12.
- [4] 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.
- [5] REN S Q, HE K M, GIRSHICK R, et al. Faster RCNN: towards real-time object detection with region proposal networks [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137-1149.
- [6] LIU W, ANGELOV D, ERHAN D, et al. SSD: single shot multibox detector [C] // Proceedings of the European Conference on Computer Vision. Berlin: Springer, 2016: 21-37.
- [7] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42(2): 318-327.
- [8] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, realtime object detection [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2016: 779-788.
- [9] 张震, 王晓杰, 晋志化, 等. 基于轻量化 YOLOv5 的交通标志检测 [J]. 郑州大学学报 (工学版), 2024, 45(2): 12-19.
- ZHANG Z, WANG X J, JIN Z H, et al. Traffic sign detection based on lightweight YOLOv5 [J]. Journal of Zhengzhou University (Engineering Science), 2024, 45(2): 12-19.
- [10] 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. Piscataway: IEEE, 2023: 7464-7475.
- [11] 李松江, 吴宁, 王鹏, 等. 基于改进 Cascade RCNN 的车辆目标检测方法 [J]. 计算机工程与应用, 2021, 57(5): 123-130.
- LI S J, WU N, WANG P, et al. Vehicle target detection method based on improved Cascade RCNN [J]. Computer Engineering and Applications, 2021, 57(5): 123-130.
- [12] 周康, 朱宗晓, 徐征宇, 等. 改进 Faster R-CNN 的道路目标检测 [J]. 计算机与数字工程, 2022, 50(4): 750-756.
- ZHOU K, ZHU Z X, XU Z Y, et al. Road target detection based on improved Faster R-CNN [J]. Computer Engineering, 2022, 50(4): 750-756.
- [13] 李国进, 胡洁, 艾矫燕. 基于改进 SSD 算法的车辆检测 [J]. 计算机工程, 2022, 48(1): 266-274.
- LI G J, HU J, AI J Y. Vehicle detection based on improved SSD algorithm [J]. Computer Engineering, 2022, 48(1): 266-274.
- [14] 胡习之, 魏征, 周文超. 融合 Camshift 与 YOLOv4 车辆检测算法 [J]. 机床与液压, 2021, 49(11): 70-74.
- HU X Z, WEI Z, ZHOU W C. Fusion of Camshift and YOLOv4 vehicle detection algorithm [J]. Machine Tool & Hydraulics, 2021, 49(11): 70-74.
- [15] ZHANG Z X. Drone-YOLO: an efficient neural network method for target detection in drone images [J]. Drones, 2023, 7(8): 526.
- [16] HE K, ZHANG X Y, REN S Q, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition [J]. Pattern Analysis and Machine Intelligence, 2015, 37(9): 1904-1916.
- [17] WANG C Y, LIAO H Y M, WU Y H, et al. CSPNet: a new backbone that can enhance learning capability of CNN [C] // 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. Piscataway: IEEE, 2020: 1571-1580.
- [18] 唐靓, 余明慧, 武明虎, 等. 基于改进 YOLOv5 的绝缘子缺陷检测算法 [J]. 华中师范大学学报 (自然科学版), 2022, 56(5): 771-780.
- TANG J, YU M H, WU M H, et al. Insulator defect detection algorithm based on improved YOLOv5 [J]. Journal of Central China Normal University (Natural Sciences), 2022, 56(5): 771-780.
- [19] LIU Y C, SHAO Z R, NICO H. Global attention mechanism: retain information to enhance channel-spatial interactions [EB/OL]. (2021-12-10) [2024-08-16]. <https://doi.org/10.48550/arXiv.2112.05561>.
- [20] WOO S, PARK J, LEE J, et al. CBAM: convolutional block attention module [EB/OL]. (2018-07-17) [2024-08-16]. <https://doi.org/10.48550/arXiv.1807.06521>.
- [21] LIU W Z, LU H, FU H T, et al. Learning to upsample by learning to sample [EB/OL]. (2023-08-29) [2024-08-16]. <https://doi.org/10.48550/arXiv.2308.15085>.
- [22] LI Y J, MA C Y, LI L P, et al. Lightweight tunnel obstacle detection based on improved YOLOv5 [J]. Sensors, 2024, 24(2): 395.
- [23] 张徐, 朱正为, 郭玉英, 等. 基于 cosSTR-YOLOv7 的多尺度遥感小目标检测 [J]. 电光与控制, 2024, 31(4): 28-34.
- ZHANG X, ZHU Z W, GUO Y Y, et al. Multi-scale remote sensing small target detection based on cosSTR-YOLOv7 [J]. Electronics Optics & Control, 2024, 31(4): 28-34.
- [24] FAHAD J, JANIBUL A S, ABU N M, et al. Detection of mold on the food surface using YOLOv5 [J]. Current Research in Food Science, 2021, 4: 724-728.

- [25] SHE F F, HONG Z Y, ZENG Z Q, et al. Improved traffic sign detection model based on YOLOv7-tiny [J]. IEEE Access, 2023, 11: 126555–126567.
- [26] WANG F, WANG H Y, QIN Z Y, et al. UAV target detection algorithm based on improved YOLOv8 [J]. IEEE Access, 2023, 11: 116534–116544.

Construction Vehicles Recognition Algorithm Based on Improved YOLOv7 in High Risk Areas

ZHANG Zhen¹, XIAO Zongrong², LI Youhao³, HUANG Weitao³

(1. School of Electrical and Information Engineering, Zhengzhou University, Zhengzhou 450001, China; 2. School of Computer and Artificial Intelligence, Zhengzhou University, Zhengzhou 450001, China; 3. Henan Huirong Oil and Gas Technology Co., Ltd., Zhengzhou 450001, China)

Abstract: To address the safety risks posed by construction vehicles operations in highrisk areas near natural gas pipelines, particularly the physical impacts and environmental disturbances caused by heavy vehicles, in this study an improved YOLOv7-based construction vehicles recognition algorithm was proposed. Six common types of construction vehicles including dump trucks, rollers, mixers, forklifts, excavators, and loaders were selected as the research objects. A custom dataset, containing images captured in various environments and angles, was used to train the model, ensuring its performance. Firstly, the CBAM attention mechanism was introduced into the YOLOv7 head, and an improved GAM attention mechanism was added to the max pooling layer to enhance the model's focus on key image features and improve detection accuracy. Secondly, the DySample dynamic upsampling module replaced the nearest neighbor interpolation, boosting precision. Finally, an improved SPPCSPC module was designed to enhance feature extraction efficiency, reduce computational costs, and accelerate inference. These modifications could enable the model to maintain high detection accuracy even in challenging scenarios such as low-quality images or distant targets. Experimental results demonstrated that the proposed algorithm achieved a precision P of 97.7%, recall R of 94.7%, $mAP@0.5$ of 98.6%, and $mAP@0.5:0.95$ of 90.4%. Compared to the original YOLOv7 algorithm, these metrics improved by 1.3, 1.4, 1.4, and 3.7 percentage points, respectively.

Keywords: highrisk areas; construction vehicles; YOLOv7; attention mechanism; upsampling; feature extraction