改进特征提取的YOLOv7算法优化研究

任富军 沈晓萍 郭灵燕 陆叶鑫 翟同悦 疏玉兵[通讯作者]
嘉兴职业技术学院海盐学院 浙江 嘉兴 314036

摘 要 针对YOLOv7算法在复杂场景下检测目标受到遮挡或检测背景受到干扰等情况下导致的检测精度低、检测不稳定等问题，本文在特征提取阶段，提出一种增加注意力机制的检测方法来优化YOLOv7的算法模型，通过在骨干网络处添加CBAM注意力机制来提高检测的精度及稳定性。

关键词 YOLOv7；目标检测；特征提取；注意力机制；CBAM

Research on YOLOv7 Algorithm Optimization for Improved Feature Extraction

Ren Fu-jun, Shen Xiao-ping, Guo Ling-yan, Lu Ye-xin, Zhai Tong-yue, Shu Yu-bing Corresponding Author
Haiyan College, Jiaxing Vocational and Technical College, Jiaxing 314036, Zhejiang Province, China

Abstract To address the issues of low detection accuracy and instability in the YOLOv7 algorithm when detecting targets in complex scenarios due to occlusion or background interference, this paper proposes an optimized YOLOv7 model by introducing an attention mechanism in the feature extraction stage. Specifically, the CBAM attention mechanism is integrated into the backbone network to enhance detection accuracy and stability.

Keywords YOLOv7; Target Detection; Feature Extraction; Attention Mechanism; CBAM

在计算机视觉技术架构中，目标检测作为推动智能系统商业化的关键任务，其技术成熟度对智能监控、自动驾驶等领域的系统性能起着决定性作用。然而，在实际工程应用中，由于环境复杂，目标检测颇具难度[1]。复杂的背景信息可能会妨碍模型的特征学习，降低YOLOv7的检测性能，进而导致漏检和误报现象。注意力机制通过引导模型识别重要信息并抑制干扰，为解决复杂场景检测问题提供了有效途径。CBAM凭借其轻量级结构和二维特征优化能力表现突出，能够实现特征通道和空间维度的动态权重调整，从而增强模型从复杂场景中识别和提取特征的能力[2]。

因此，本文提出将CBAM集成至YOLOv7的主干网络中，提高算法的检测精度和稳定性，改进复杂场景下的检测缺陷，为复杂场景下的可靠应用提供技术支持。

1 传统YOLOv7算法

YOLOv7作为2022年提出的单阶段目标检测算法，在COCO数据集上实现了检测精度与推理速度的突破性平衡，其核心优势源于对网络结构的精细化设计、梯度流优化及计算效率的精准控制[3]。YOLOv7网络结构如图1所示。

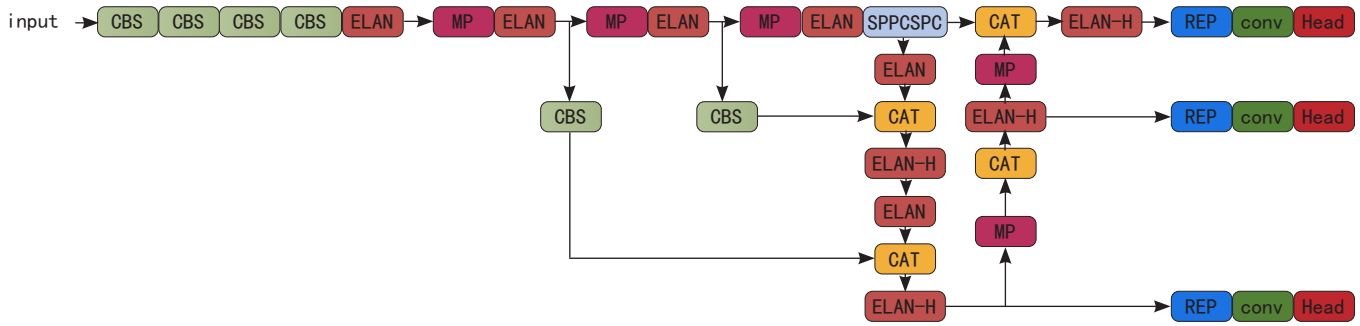


图1 YOLOv7网络结构图

2 改进特征提取网络的YOLOv7算法

本章在YOLOv7模型的基础上，重新构建新的主干网络，添加CBAM注意力机制。

2.1 CBAM注意力机制

CBAM是一种用于增强卷积神经网络性能的注意力机制模块，主要目标是通过在CNN中引入通道注意力和空间注意力来

[基金项目] 2025年嘉兴职业技术学院校立科研项目，项目名称：基于YOLOv7的车辆停车状态识别算法研究，项目编号：jzyy202532。

提高模型的感知能力，从而在不增加网络复杂性的情况下改善性能。

2.2 在YOLOv7中添加CBAM注意力模块

伪代码如下。

```
// 核心模块
DEFINE MODULE CBS(c1, c2, k=3, s=1): // Conv+BN+SiLU
RETURN Conv2 d(c1,c2,k,s) + BatchNorm + SiLU()
DEFINE MODULE CBAM(c): // 通道+空间注意力
// 通道注意力(CA)
avg_p, max_p = AdaptiveAvgPool2d(1)(x), AdaptiveMaxPool2 d(1)(x)
ca = Sigmoid( Conv2 d(Conv2 d(avg_p, c//16) + Conv2 d(max_p, c//16), c) )
x = x * ca
// 空间注意力(SA)
avg_s, max_s = mean(x, dim=通道, keepdim=True), max(x, dim=通道, keepdim=True)
sa = Sigmoid( Conv2 d(concatenate(avg_s, max_s), 1, 7) )
RETURN x * sa
DEFINE MODULE ELAN(c1, c2): // 特征聚合
x1,x2 = CBS(c1,c2//2,1), CBS(c1,c2//2,1)
x3,x4 = CBS(c2//2,c2//2,3), CBS(c2//2,c2//2,3)
RETURN CBS(concatenate(x1,x2,x3,x4), c2, 1)
DEFINE MODULE MP(c1): // 最大池化下采样
RETURN MaxPool2 d(2,2)
DEFINE MODULE SPPCSPC(c1, c2): // 空间金字塔池化
x = CBS(c1, c2//2, 1)
pools = [x, MaxPool2 d(5,1,2)(x), MaxPool2 d(9,1,4)(x), MaxPool2 d(13,1,6)(x)]
RETURN CBS(concatenate(pools), c2, 1)
DEFINE MODULE CAT(x_list): // 通道拼接
RETURN concatenate(x_list, dim=通道)
DEFINE MODULE UP(x, s=2): // 最近邻上采样
RETURN Upsample(x, scale_factor=s, mode=”nearest”)
DEFINE MODULE REP(c1, c2): // RepVGG
RETURN RepConv(c1,c2,3)
DEFINE MODULE Head(c): // 检测头（框+类别+置信度）
RETURN Conv2 d(c, (类别数+4)*3, 1)
// 前向传播
FUNCTION YOLOv7_CBAM_Forward(input):
// Backbone（插入CBAM）
x = CBS(input,32) → CBS(64) → CBS(128) → CBS(256)
x = ELAN(x,256) → CBAM(x) # ELAN1 + CBAM
b1 = CBS(x,128) # 分支1
x = MP(x) → ELAN(512) → CBAM(x) # ELAN2 + CBAM
b2 = CBS(x,256) # 分支2
x = MP(x) → ELAN(1024) → CBAM(x) → MP(x) → ELAN(2048) # ELAN3+CBAM+ELAN4
x = SPPCSPC(x,2048)
// Neck特征融合+检测头
x = CAT([x, UP(x)]) → ELAN-H(1024)
h1 = Head(REP(x,512)) # 大目标
x = UP(x) → CAT([x,b2]) → ELAN-H(512)
h2 = Head(REP(x,256)) # 中目标
x = MP(x) → CAT([x,b1]) → ELAN-H(256)
h3 = Head(REP(x,128)) # 小目标
RETURN [h1,h2,h3]
```

改进后的YOLOv7网络结构如图2所示。

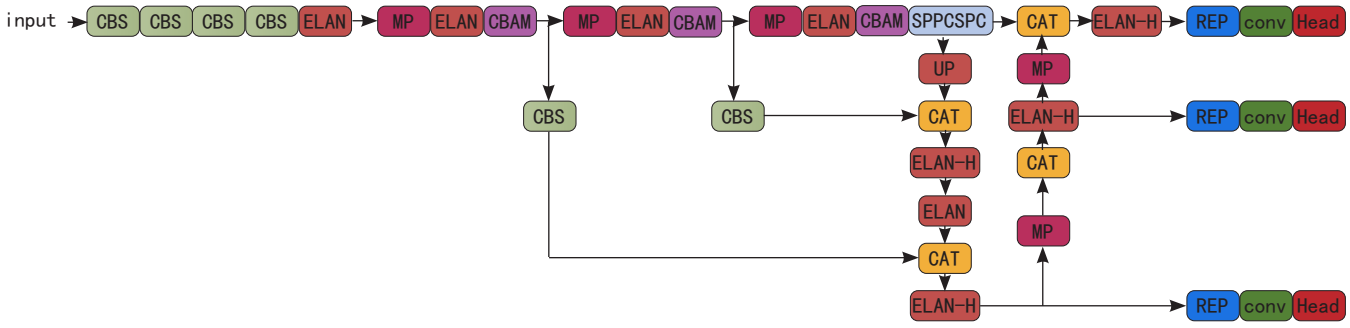


图2 改进后的YOLOv7网络结构

3 实验结果与分析

3.1 实验环境及平台

实验环境：AMD Ryzen 7 5800H with Radeon Graphics 3.20 GHz，Windows 10操作系统，16 GB内存，1 T硬盘。

3.2 数据集

本实验采用智能停车场系统通常使用的Parking Lot Project数据集，Parking Lot Project是一套多任务导向的停车位相关开源数据集，由学术机构、开源社区共同维护，核心用于停车位检测、分割、占用分类等智能停车场相关算法的训练与评估[4]，数据集结构如表1所示。

表1 数据集结构

表名	字段	说明
images	image_id	图像唯一标识
	file_path	图像存储路径
	image_type	图像类型
	image_size	图像尺寸
	image_format	图像格式
	capture_time	采集时间戳
detection_annotations	annotation_id	检测标注唯一标识
	image_id	关联图像ID
	bbox_coords	停车位边界框坐标
	class_id	类别ID
	parking_status	停车位状态
	annotation_format	标注格式
segmentation_annotations	annotation_id	分割标注唯一标识
	image_id	关联图像ID
	mask_path	分割掩码文件路径
	seg_coords	停车位轮廓坐标
	mask_format	掩码格式
classification_annotations	patch_id	预分割patch唯一标识
	image_id	关联全景图ID
	patch_size	patch图像尺寸
	parking_status	停车位状态
	patch_coords	patch在全景图中的坐标
metadata	meta_id	元数据唯一标识
	camera_id	摄像机编号
	camera_angle	摄像机视角
	parking_type	停车场类型
	weather	采集天气
	lighting	光照条件

3.3 评价指标

评价指标主要用召回率（Recall）和精确率（Precision）衡量准确度。精准率指反映预测的精准性，可衡量模型误判正类的概率.召回率体现模型对正类样本的捕获能力，用于评估模型漏判正类的概率。平均精度均值mAP通过综合考量不同IoU阈值下的模型精度表现，能全面反映模型在多目标、多尺度场景下的检测性能，规避单一阈值评估的结果偏差，是衡量目标检测模型综合能力的关键指标[5-7]。

召回率和精确率公式如式（1）和式（2）所示：

$$Recall = \frac{TP}{TP + FN} \quad (1)$$

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

3.4 实验结果与分析

从实验数据可以看出，本文改进后的算法在复杂场景下检测目标受到遮挡或检测背景受到干扰等情况下的检测精度有了一定程度的提高，如表2所示。

表2 对比实验

Models	Precision（%）	Recall（%）	mAP@0.5（M）
YOLOv7	83.6	69.2	77.2
本文	83.9	69.7	77.5

4 结束语

针对YOLOv7算法在复杂场景下对目标物体进行检测时，因目标受到部分或完全遮挡、检测背景存在复杂纹理或动态干扰、光照条件变化剧烈等不利因素影响，而导致检测精度显著下降、检测结果波动较大、模型鲁棒性不足等问题。本文在特征提取阶段提出一种引入注意力机制的改进检测方法，以优化YOLOv7算法模型的结构与性能，通过在骨干网络的关键位置嵌入CBAM注意力机制模块，增强模型对重要通道特征的响应能力，抑制冗余或干扰信息，从而有效提升复杂环境下目标检测的准确性、稳定性及泛化性能。

参考文献

[1]李海涛,贺自帅,赵长明,等.基于图像识别算法的无人机影像烤烟识别研究[J].现代农业科技,2025(24):145-148.

[2]黄子甜,兰康睿,郑泊文,等.恶劣环境下的道路目标检测算法研究[J].电子测量与仪器学报,2025,39(10):269-277.

[3]王焱辉.基于环视视觉下的停车位及障碍物多任务检测算法技术研究[D].合肥:合肥工业大学,2024.

[4]昌新萌.基于YOLO网络的交通车辆目标检测方法研究[D].昆明:云南大学,2023.

[5]张卓阳.基于深度学习的室外停车场车位检测研究[D].西安:长安大学,2022.

[6]刘伟.路边停车场景下基于嵌入式平台的目标检测研究与应用[D].成都:电子科技大学,2023.

[7]刘丹,王建华,袁骋,等.基于YOLOv7-coupling的城市街景车辆检测方法[P].中国专利:CN120125976A,2025-06-10.

（本文编辑 曹 阳）