

DOI:10.19651/j.cnki.emt.2210035

基于改进 YOLOv5 的复杂场景多目标检测^{*}

强 栋 王占刚

(北京信息科技大学信息与通信工程学院 北京 100101)

摘要:针对多目标图像检测环境复杂、目标物位置数据冗余且长宽高数据大小不一的问题,利用神经网络算法可以有效提高不同类目标物并行检测的准确度和稳定性,提出一种基于改进 YOLOv5 网络的多目标检测方法。首先依据不同目标物的空间尺度大小,改进模型的特征融合方法,添加多尺度特征检测层以减小多目标检测时的误差,同时增加自适应特征增强模块(adaptive feature adjustment),降低网络的误检率与漏检率;然后使用 K-means++ 算法估计候选框,获得更优的框参数;最后在损失函数中使用 EIOU(efficient IOU loss)做优化。实验表明:改进后的方法 mAP(mean average precision)达到 76.48%,相比经典 YOLOv5 网络提升了 3.2%,小尺寸目标物检测准确度均值增加 6.3%。改进方法网络延续 YOLOv5 网络的轻量高效,对于多尺度目标物检测获得更优的检测精度,能够实现更准确的实时多目标检测。

关键词:神经网络;多目标检测;YOLOv5;自适应特征增强;损失函数优化

中图分类号: TP391.4 **文献标识码:** A **国家标准学科分类代码:** 510.4050

Improved YOLOv5 complex scene multi-target detection

Qiang Dong Wang Zhangang

(School of Information and Communication Engineering, Beijing Information Science and Technology University, Beijing 100101, China)

Abstract: Aiming at the problems of complex multi-target image detection scenes and redundant target position data with different length, width and height, the neural network algorithm can effectively improve the accuracy and stability of parallel detection of different types of targets. A multi-target detection method based on the improved YOLOv5 network is proposed. First, according to the spatial scale of different objects, the feature fusion method of the model is improved, and a multi-scale feature detection layer is added to reduce the error of multi-target detection. At the same time, Adaptive Feature Adjustment module is added to reduce the false detection rate and missed detection rate of the network; then K-means++ algorithm is used to estimate the candidate frame to obtain better frame parameters; finally, Efficient IOU Loss is used in the loss function for optimization. Experiments show that the mean average precision of the improved method reaches 76.48%, which is 3.2% higher than the classic YOLOv5 network, and the average detection accuracy of small-sized objects increases by 6.3%. The improved method network continues the lightweight and high-efficiency of the YOLOv5 network, obtains better detection accuracy for multi-scale target detection and can achieve more accurate real-time multi-target detection.

Keywords: neural network; multi-target detection; YOLOv5; adaptive feature enhancement; loss function optimization

0 引 言

在计算机视觉和深度学习技术蓬勃发展的今天,作为图像识别和增强现实领域核心技术之一的多尺度目标检测算法也在不断发展,伴随着各个行业与计算机及人工智能联合愈发紧密,多目标检测技术在人脸识别、海关安检、消

防预警、公安监控、车辆检测,医学仪器检查以及军事等领域取得了广泛的应用。目标检测主要是通过对特定图像进行处理,利用相关算法提取图像的特征指标,从而获得指定目标物在图像中的位置、长宽高范围并且识别此目标物^[1]。相较于传统算法,深度学习在目标检测中的应用使得检测效率和速率有极大的提升,利用该技术在输入图像和视频

收稿日期:2022-05-20

^{*} 基金项目:国家重点研发计划课题(2018YFC1800203)、北京市科技创新服务能力建设-基本科研业务费(市级)(科研类)(PXM2019_014224_000026)项目资助

中可以锁定目标物的精确位置,获得高精度目标参数并且识别该目标物的类别^[2]。

多目标检测往往是提升目标检测效果的有效方法和必要步骤^[3],但由于多目标图像检测同时包含多个特征尺度不一的目标物,位置数据准确性不高,多种目标物特征信息复杂,识别判定困难,所以需要引入性能更强和效率更高的深度学习算法^[4]。有别于单一目标检测,多目标检测算法利用卷积神经网络算法提取目标物复杂多样的图像特征信息,其特征特异性减弱,可同时适用于多种目标物的识别,提高了检测算法的有效性和稳定性^[5]。

基于一个单独 end to end 回归卷积神经网络的 YOLO 目标检测算法^[6]于 2015 年提出后发展迅速。相较于 R-CNN 和 Faster CNN 算法,YOLO 深度学习的训练和检测均是在一个单独的卷积网络中进行,其算法不事先标定目标检测预选区域(region proposal),而是直接将输入图像在 YOLO 网络层中进行训练和检测,最终输出由 24 个卷积层和 2 个全连接层处理后的图像特征信息。随着 YOLO 算法的发展迭代,又更新了 mAP 值更高的 v2、v3、v4、v5 版本的改进算法^[7]。

YOLO 目标检测算法采用设计简洁的模型,检测效率比较高,不过对于更新迭代后的 YOLOv5 算法,其将全局特征图中的多个特征信息采用类似的权重进行计算,使得多个特征信息对模型结果产生类似的影响。但是,在现实场景下,目标附近通常伴随着相互关联密切的上下文语义,模型需将关注点集中于关键语义且省略关联不大的背景语义,不应该同等对待^[8];其次由于现实环境多目标检测场景中存在数目较大的多尺度目标物,单一目标物信息匮乏,且多尺度目标在不同角度现实场景中呈现的位置信息差异较大,目标物空间尺度不一,从而使得检测精度不高^[9];另外,采用拍摄设备采集现实环境多目标复杂场景的图像,位置集中的多目标间彼此易发生空间重叠情况。综上所述,应提出一种适用于复杂现实场景下的多尺度目标物检测模型,以满足实际应用场景需求。

针对多目标图像检测环境复杂、目标物位置数据冗余且宽高数据大小不一、目标物识别判定信息繁多的特点,本文在 YOLOv5 算法的基础上加以创新。首先,为了提升空间尺度较小的目标物的检测精度,结合空间注意力与通道注意力,引入自适应特征增强模块(Adaptive Feature Adjustment, AFA),从而获得在空间和通道两个维度上的权重参数,用以补充空间尺度较小的目标物特征信息,减弱大尺度目标对小目标的影响,并且在原主干网络架构中添加多尺度特征检测层以减小多目标检测时的模型误差;其次,在候选框尺寸选择时使用 K-means++ 算法进行估计,提高候选锚框的尺寸匹配度;最后,在损失函数中针对多目标检测使用 EIOU 代替 GIOU,提高模型归一化性能。

1 YOLOv5 网络算法

YOLOv5 是 YOLO 最新迭代的单一端到端神经网络

目标检测算法,分为 YOLOv5s、YOLOv5m、YOLOv5l 和 YOLOv5x 四种不同检测尺度的算法模式^[10],本文主要选用 YOLOv5s 作为多目标检测模型,其网络结构如图 1 所示,主要包括输入端、骨干网络 Backbone、特征网络层 Neck 和头输出端 4 部分。其中输入端主要是数据预处理阶段,引入 Mosaic 模糊数据扩展技术和多尺度自适应候选框技术对输入数据进行压缩、延展、统一标准等多项处理^[11]。其次骨干网络主要包括专注力结构 Focus 和跨阶段局部网络(cross stage partial network, CSPnet),目的是提升网络在获取图像特征数据时的专注度,从而更高效地融合上下立体塔状层级的特征信息,便于检测识别空间范围不同的多个目标^[12]。网络结构末尾的头输出端主要是对融合后的特征信息进行识别判定,经损失函数优化归一和 softmax 层后输出多个目标物的空间位置信息、范围大小和判定类别^[13]。头输出端的损失函数是广义交叉联合损失函数 GIOU 和二元交叉熵损失函数,极大地改善了目标检测精确性。

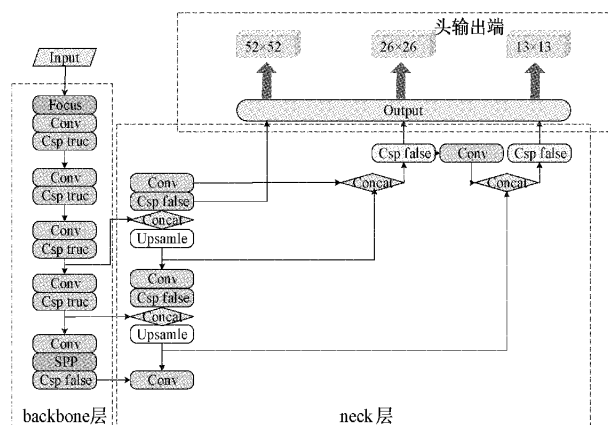


图 1 YOLOv5 网络结构

2 改进 YOLOv5 算法

2.1 多尺度特征图融合检测

在迭代版本 YOLOv5 的核心网络结构中,原始定义了 3 种采样率不同的特征检测层来处理输入图像中尺寸不一的多个目标物^[14],模型网络首先对经数据预处理的输入图像通过各个特征检测层进行 3 次下采样,采样速率分别为原始速率的 8 倍、16 倍和 32 倍,之后得到 3 种不同大小的特征信息图,用来作为 SPP、FPN 和 PAN 模块组成的特征图融合网络结构的输入。根据特征塔状排列网络结构的原理,不同采样率特征检测层卷积后的特征图所包含的语义和特征信息不同,采样率越高卷积网络层数越大的特征图感受野越大,包含信息越丰富,但经深层次网络后会损失小目标的部分位置信息和空间范围;而采样率越低卷积网络层数越小的特征图感受野越小,原始图像中小目标的细节特征信息保留较完整,但其缺乏图像的全局特征信息。

针对多目标图像检测场景,由于输入图像包含许多待

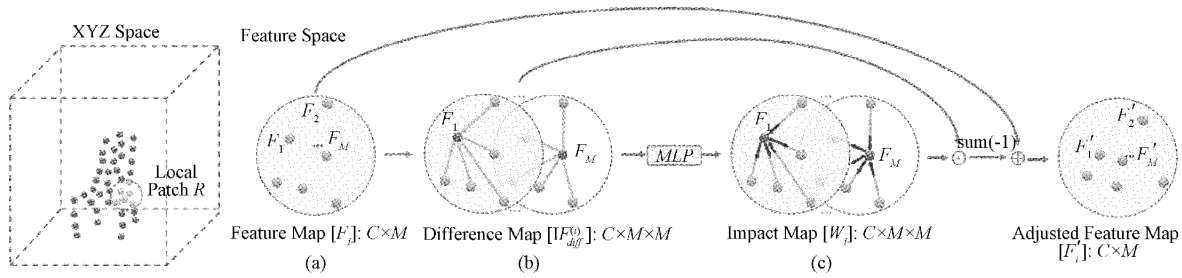


图 3 自适应特征增强模块结构

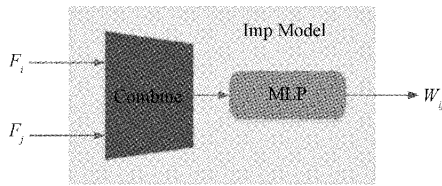


图 4 f_{imp} 模块结构

不适用于本文的复杂场景多尺度目标数据,如果使用原算法的候选框来学习训练,不但会降低网络收敛速度,而且会使得模型的目标检测精度下降。

目标物候选框是锚定待检测物边界的框区域,为获取针对性的初始锚框,可以利用 K-means++ 算法来推导候选框和实际目标的对应关系,使得检测候选框 Bounding box 的空间位置尽可能趋近于实际目标的空间位置。实际目标样本与聚类中心的对应形式如图 5 所示,其中,虚线框表示实际目标样本区域,实线框表示聚类中心区域。

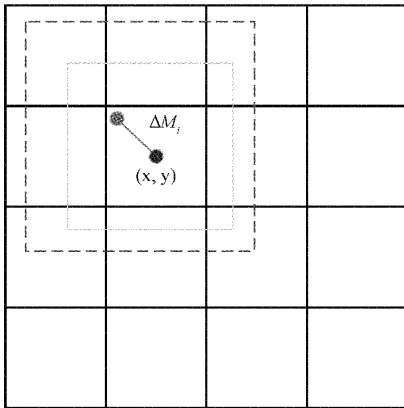


图 5 样本与聚类中心空间对应关系

对于 K-means 算法随机选择首个聚类中心的弊端改进效果显著,具体是起初选取任意一个样本作为算法的首个聚类中心,对实际目标样本的聚类可引入 K-means++ 策略计算出样本点对应于初始聚类中心的相对坐标,并将该样本归类到与其相距最小的聚类中心所对应的类别中。同时计算出每个样本被设定为下一个聚类中心的概率,选择概率值最大的样本作为下一个中心^[15],其中概率的计算公式为:

$$p = \frac{\Delta M^2}{\sum_{i=1}^n \Delta M_i^2} \quad (3)$$

其中, ΔM 表示每一个样本点到当前聚类中心的距离,在进行计算时针对每一个目标样本,可基于已有的聚类目标更新迭代当前聚类中心,不断进行迭代计算直至全部样本目标都经判别定义到各个聚类中心的类别中,结果可挑选出 L 个聚类类别。虽然 K-means++ 算法采取初始聚类中心由样本中心决定的策略会提高算法推导的时间成本,然而 L 均值的具体数值在决定恰当的聚类中心后将立即趋于收敛,故该算法事实上有效缩短了推导时间,且推导出的候选框与本文数据集中多尺度目标的位置信息的误差较小,从而使复杂场景下多尺度目标物的检测更加高效。

为了平衡计算精度和效率,本文设置 5 层特征检测图层,5 种不同采样率下的特征图与候选框范围的关系如表 1 所示。

表 1 特征图与候选框范围关系

特征图尺度	候选框范围		
13×13	[129,101]	[162,204]	[396,332]
26×26	[40,73]	[76,82]	[116,53]
52×52	[18,12]	[15,32]	[24,46]
104×104	[7,10]	[10,16]	[15,21]
208×208	[4,6]	[6,9]	[10,13]

表 1 中显示的 15 个候选框范围的相差较大,故在模型的不同尺度特征检测层中选取对应的候选框,聚类算法性能比较好。

2.3 损失函数

在原 YOLOv5 网络架构中,采用广义边界回归损失函数 GIOU 作为头输出端的损失函数,使用二元交叉熵损失函数在输出端 softmax 层中归一化计算目标物类别概率和目标物空间位置范围,计算公式为:

$$GIOU = \frac{C \cap D}{C \cup D} - \frac{H - (C \cup D)}{H} \quad (4)$$

$$Loss_{GIOU} = 1 - GIOU$$

式中: $Loss_{GIOU}$ 为目标位置损失函数; C 和 D 分别为实际目标区域和预测候选框区域面积; $C \cap D$ 为边框交集面积; $C \cup D$ 为边框并集面积; H 为能包围预测候选框和实际目标区域最小的外接框面积。

针对候选框和实际目标区域完全重合或者候选框在实际目标区域的内部, IOU 和 GIOU 的损失函数值 $Loss$ 趋于一致, 则此时 GIOU 等同于 IOU, 始终不能区分候选框和实际目标区域的位置关系, 同时会导致这种情况下边界回归损失函数 $Loss$ 收敛很慢, 难以计算数值。之后提出采用 CIOU 损失函数取代 GIOU 损失函数来保证候选框与实际目标区域的长宽纵横比尽可能接近, 从而加快边界回归损失函数 $Loss$ 的收敛速度。但一旦收敛到候选框和实际目标区域的长宽比值达到线行比例时, 则候选框再进行收敛回归时长度和宽度就不能同时增大或者减少。因此本文提出使用 EIOU 回归损失函数^[16]来解决 CIOU 存在的问题, 在 CIOU 的惩罚项基础上将候选框和实际目标区域长宽比值的影响因子拆分开, 分别计算候选框和实际目标区域的长宽数值。EIOU 回归损失函数的表达式为:

$$L_{EIOU} = L_{IOU} + L_{dis} + L_{asp} = 1 - IOU + \frac{\rho^2(b, b^{gt})}{C_w^2} + \frac{\rho^2(w, w^{gt})}{C_w^2} + \frac{\rho^2(h, h^{gt})}{C_h^2} \quad (5)$$

其中, C_w 和 C_h 是闭包候选框和实际区域最小外接矩形的长度和宽度。式中包含 3 个部分: 候选框和实际区域的重叠损失 L_{IOU} , 候选框和实际区域的中心距离损失 L_{dis} , 候选框和实际区域的长宽损失 L_{asp} 。EIOU 损失函数的前两部分继承 CIOU 中的原理, 而新加入的长宽损失 L_{asp} 可直接使候选框与实际区域的长宽之差达到最小, 使得回归损失函数收敛速度更快, 从而有效提高在复杂多目标检测场景下的检测精确性, 优化了由边框之间相互关系所引起的收敛问题。

3 实验与结果分析

3.1 实验环境与数据集

本文进行实验的操作系统及开发框架为 Windows10 64 位 + Python 3.7.13 + Pytorch 1.7 + CUDA 10.1 + OpenCV 3.4.2; 实验环境的硬件设备为 CPU: AMD Ryzen 7 4800H, GPU: NVIDIA GeForce RTX 2060, 显存: 6 G, 内存: 16 G。

实验所使用的多目标检测数据集为开源数据集 NEU-

DET, 首先对目标数据集采取预处理, 选取六个类别的图像标签当作学习和训练的对象。实验所用数据集总共有 937 张图片, 其中训练集、测试集与验证集分别以 8:1:1 的比例选取图片, 并且两者数据保证互不相干, 相互独立。

在对数据进行标签预处理后, 将输入 YOLOv5 检测网络进行学习和训练的数据集样本图像尺度范围设置为 640×420 , batch-size 的值为 16, 配置初始学习率为 0.01, 采用学习率余弦退火策略对学习率进行更新, 余弦退火超参数设置为 0.2, 使用 SGD 函数优化器, 将第一个动量因子为 0.937。权重损失递减系数为 0.0005。

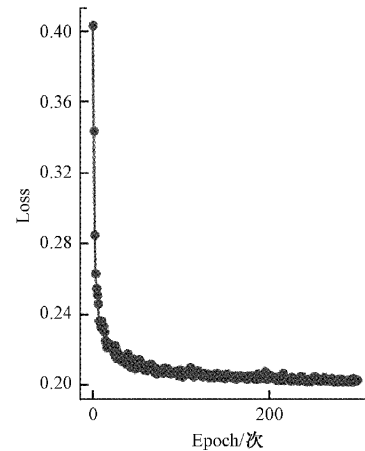


图 6 检测算法训练损失值

改进的 YOLOv5 检测算法训练损失值如图 6 所示, 学习训练 300 个 epoch 后, 损失函数逐渐归一化至平稳, 变化幅度逐渐减弱, 证明学习率设置合理, $Loss$ 值最终收敛于 0.21~0.22 之间, 多目标检测网络最终达到收敛状态。

3.2 消融实验

本文需检验在改进算法增添的各个模块与策略的有效性, 因而针对自适应特征增强模块 AFA、K-means++ 算法和改进损失函数模块来设置消融实验, 目的是验证每个模块在其余变量不发生改变时对复杂场景多目标检测模型效率的影响。设置实验时采用 YOLOv5s 模型作为对比模型, 样本图像尺寸大小设置为 640×420 , 学习训练 300 个 epoch, 其结果如表 2 和图 7 所示。

表 2 消融实验验证结果

方法	mAP@50/%	mAP@50:95/%	P/%	R/%	模型参数/m
a YOLOv5s	59.41	40.07	54.26	53.84	7.08
b CMBA_YOLOv5s	60.05	41.83	55.08	57.47	7.36
c AFA_YOLOv5s	65.57	45.29	59.39	60.75	8.62
d AFA_K-means++_YOLOv5s	68.33	50.48	65.03	64.28	8.63
e AFA_K-means++_EIOU_YOLOv5s	72.47	52.59	68.83	71.09	8.62

根据表 2 中所呈现的数据进行分析, 将自适应特征增强 AFA 模块引入至 YOLOv5s 的检测模型中, 实验数据 b 和 c 表明 AFA 模块的检测效率更加优秀, mAP50 值比

CBAM 提高了 5.52%, 且模型权重数增多了 1.26 m, mAP50 值对比初始 YOLOv5s 提高了 6.16%, 虽然在模型参数数量上有所增加, 但引入 AFA 模块后能有效提高

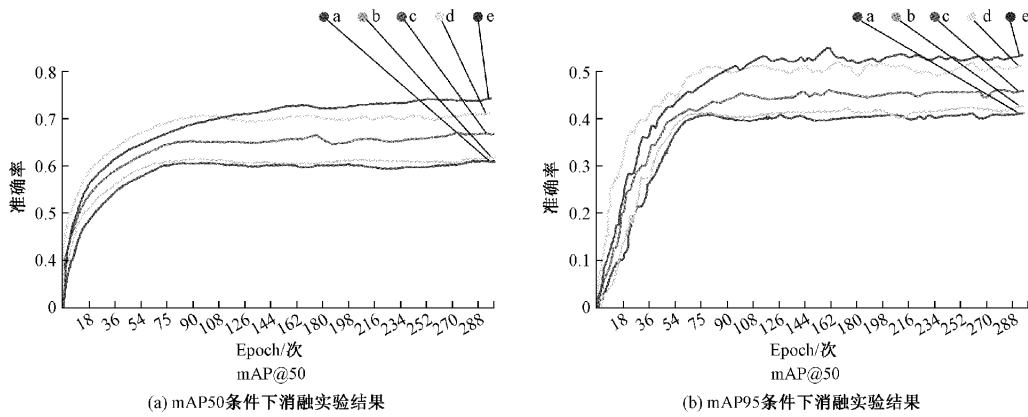


图 7 检测算法训练过程

YOLOv5s 网络模型的检测精准度,在复杂现实场景多尺度目标检测时有效解决因环境变量多变而引起的目标物误检、漏检等缺陷;再选择 K-means++ 算法提升候选框的空间范围契合度,模型权重数未产生较大改变,且 mAP50 值提升了 2.76%,使候选框与目标样本的空间契合度得到明显提升;另外,采用优化后的损失函数 EIOU 与模型结合应用在目标检测上,优化后模型参数量也基本不发生太大变化, mAP50 值在之前的基础上提升了 0.43%。根据实验数据,实验结果如图 7 所示,详细表现出各模块改进前后的 YOLOv5s 模型在学习进程中目标检测精确度数值的变化。其中,增加自适应特征增强模块 AFA 后的模型曲线在原使用 CBAM 模块的模型曲线上方,结果表明经 AFA 模块的改进 YOLOv5s 模型平稳性更好、精确度更高。另外引入 AFA 模块前的精确度到达拐点的耗时比引入后更久,同时在实验进行到 100 次 Epoch 后实验数据呈现出更平稳的态势,由此可得出结论,优化后的 YOLOv5s 模型能够迅速稳定收敛。经实验数据进行详细分析后,可证明本文改进检测模型在 YOLOv5s 模型的基础上提高了多尺度目标检测精度,具有较高的准确率,优化了复杂现实场景检测模型,卓有成效地满足实际使用的设计需求。

3.3 对比实验

1) K-means++

相较于原 YOLOv5s 模型,可进行对比实验来检验文中提出 K-means++ 算法提升候选框的空间契合度能否满足检测模型在复杂场景多尺度目标检测的精确度。对 YOLOv5s 检测模型与改进算法后的模型在统一变量环境下独立进行同等 epoch 次数的测试,实验结果如表 3 所示。

表 3 对比实验测试结果

模型	聚类方法	Pr/%	mAP/%	FPS
原 YOLOv5s	K-means	75.23	77.69	30.6
原 YOLOv5s	K-means++	77.02	78.26	30.8
改进 YOLOv5s	K-means	77.45	79.08	34.8
改进 YOLOv5s	K-means++	78.37	79.93	35.1

从表 3 中能得出结论,将原 YOLOv5s 模型中候选锚框的聚类中心选择策略由 K-means 替换为 K-means++ , mAP 值提升了 0.57%,FPS 提升了 0.65%。由于改进算法后的 YOLOv5s 模型对比原 YOLOv5s 模型的检测策略发生改变,检测准确度也会有所不同。在改进 YOLOv5s 模型的选择策略上引入 K-means++ 算法和引入 K-means 算法做对比, mAP 值提升了 0.85%,FPS 提升了 0.86%,且对比原 YOLOv5 检测模型, mAP 值高出了 2.24%。综上所述,引入 K-means++ 聚类算法推导出的候选框的空间范围更契合,应用于文中的复杂现实环境多尺度目标检测效果显著,可使检测模型获取了更高的检测精确率。

2) 特征增强模块对比

为了进一步验证本文改进自适应特征增强模块 AFA 的性能,将其分别与不同特征增强模块进行对比实验,结果如表 4 所示, YOLOv5s_CBAM 表示使用原始 CBAM 中压缩率为 16 的两层全连接层进行特征聚合; YOLOv5s_CBAM_FC 表示使用一层全连接层进行特征聚合; YOLOv5s_AFA 表示使用改进后的自适应特征增强模块 AFA 进行特征聚合。

表 4 不同特征增强模块对比结果

类别	模型参数/m	mAP/%
YOLOv5s	7.08	65.47
YOLOv5s_CBAM	7.35	66.38
YOLOv5s_CBAM_FC	31.43	72.06
YOLOv5s_AFA	8.61	72.84

由表 4 的数据可得出结论,在 YOLOv5s 模型中引进 CBAM 特征增强模块后,参数量增加了 0.27 m, mAP 提升了 0.91%,从而证实引入特征增强模块能够使网络愈发精确地检测多尺度目标;而在 YOLOv5s 模型引入使用一层全连接层聚合的 CBAM 模块时,模型参数量增加了 24.08 m, mAP 提升了 5.68%,精准度比仅增加 CBAM 模块时更高,这意味着在特征增强模块中减少空间特征压缩更有助

于多尺度目标检测;最后,在 YOLOv5s 模型结构中引进 AFA 模块后,虽然增加了 1.53 m 的模型参数,却使得 mAP 比原 YOLOv5s 网络高出了 7.37%。以上测试可说明,改进后的自适应特征增强模块 AFA 并行结合特征空间增强和信息增强,能够显著地提升模型检测的精确度,更适用于复杂场景下多目标检测。

3) 损失函数对比

本文进行对比实验后可说明在损失函数上用 EIOU 替代能否使改进后的 YOLOv5s 模型愈发迅速平稳地收敛。对比实验结果如下,将采用广义边界回归损失函数 GIOU 的检测模型作为对比对象,与采用 EIOU 回归损失函数的模型进行对照试验,损失函数曲线对比图如图 8 所示。

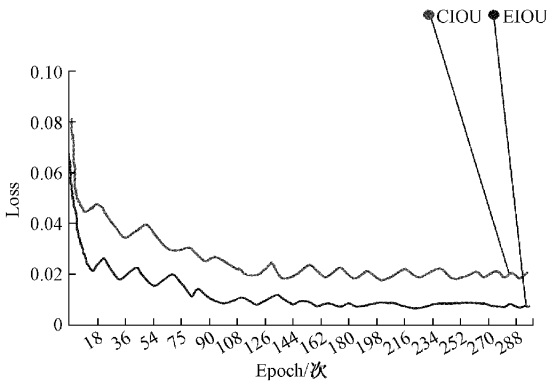


图 8 损失函数曲线对比图

分析图 8 可以看出,60 次 Epoch 后,两不同损失函数模型都开始逐渐收敛,GIOU 模型收敛后的损失值在 0.019~0.022 之间波动,并且其 Loss 值普遍高于 EIOU,损失值波动严重;而 EIOU 模型收敛后,其损失波动范围在 0.009~0.010 之间,虽然其训练过程仍有振荡的现象,但是损失函数的数值减小了很多,函数值波动区间也降低,从而使模型收敛地愈发迅速平稳,在使得模型有效收

敛的前提下,极大程度地提升了模型的平稳性,适用于复杂现实场景中多尺度目标检测。

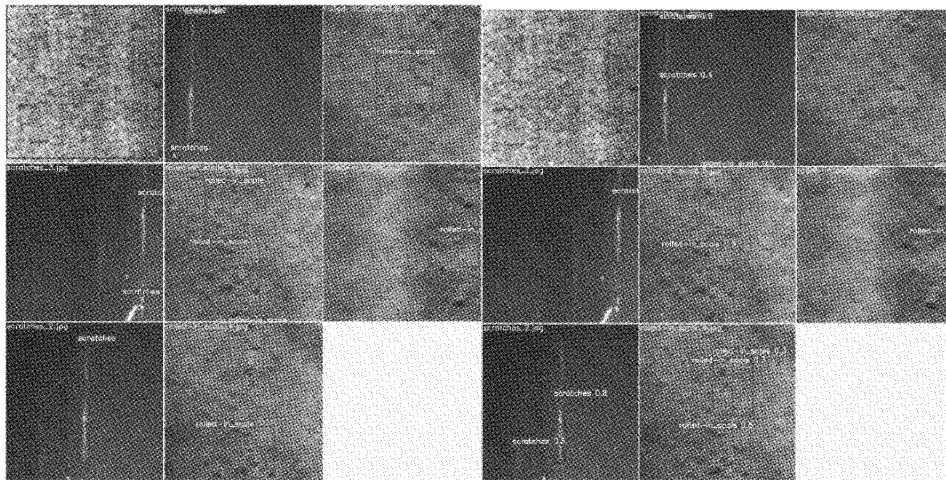
4) 不同模型对比

为更细致地评估改进算法的检测精度和效率,将改进算法与主流神经网络目标检测算法在统一数据集上作横向对比实验。改进前后各项评价指标对比如表 5 所示。

表 5 改进前后各项评价指标对比表

类别	Precision/%	mAP/%	时间/s	FPS
Faster R-CNN	75.41	65.07	0.08	12.6
AIZOO	80.05	69.83	0.025	46.8
YOLOv5	81.27	73.29	0.028	40.7
本文算法	86.33	76.48	0.034	35.2

改进算法和原算法做对比,其 mAP 改善了 3.2%,包括尺寸较小目标物的召回率提高 11.7%,目标检测准确度均值增加 6.3%;空间范围尺寸较大的中型目标物的召回率和检测准确度均值亦得到改善,大型尺寸目标物的召回率和检测准确度均值也仅仅减少了 0.3%和 0.4%。由于添加两层对象是尺寸较小物体的采样结果为 104×104 和 208×208 的特征提取层,在新增特征提取层上将出现 $(104 \times 104 + 208 \times 208) \times 3 = 33\ 696$ 个候选框,小型候选框的结果增加会对空间尺度较大目标的检测输出产生一定的影响。目标检测实验最终结果数据表明,改进 YOLO 算法在轻微影响空间尺度较大目标检测准确度均值的条件下,改善了多种尺寸较小目标物和中型目标的准确度和 mAP 值。同时检测效率为 35FPS,满足了实时检测的基本要求。YOLO 算法与改进后算法的多目标检测情况对比如图 9 所示,可见改进算法对比原算法识别出了更多的中小目标物,召回率和平均精确度性能都有所增强,可见复杂多目标场景检测实际效果不错。使用该算法对复杂场景下多目标物图像进行检测,效果如图 10 所示,目标检测情况良好。



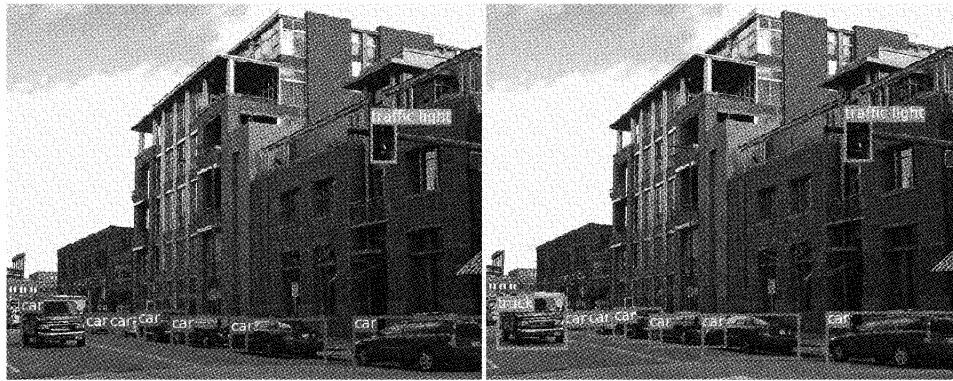


图 9 检测算法结果对比图



图 10 多目标图像检测效果图

4 结 论

针对多目标检测图像场景复杂,检测目标数量多且尺度不一导致检测精确度不高的问题,本文在 YOLOv5 算法的基础上增加多尺度特征图融合检测模块以便更好地识别空间范围不一的多个目标;在主干网络架构中加入特征偏置增强机制以减小 YOLO 目标检测网络模型的参数误差;并且在损失函数中针对多目标检测优化其交叉二元熵函数,提高模型归一化性能。改进后的 YOLO 算法精确率更高,稳定性更强,满足在复杂场景下多目标检测的性能要求,提高了 3.2% 的 mAP。

参考文献

- [1] GEETHAPRIYA S, DURAIMURUGAN N, CHOKKALINGAM S P. Real-time object detection with Yolo[J]. International Journal of Engineering and Advanced Technology(IJEAT), 2019, 8(5): 92-98.
- [2] JUN W, YU L, KE W, et al. Skeleton based temporal action detection with YOLO[J]. Journal of Physics: Conference Series, 2019, DOI: 10.1088/1742-6596/1237/2/022087.
- [3] DU Z X, YIN J Y, YANG J. Expanding receptive field YOLO for small object detection [J]. Journal of Physics: Conference Series, 2019, DOI: 10.1088/1742-6596/1314/1/012202.
- [4] XU Q W, LIN R Z, YUE H, et al. Research on small target detection in driving scenarios based on improved Yolo network [J]. IEEE Access, 2020, DOI: 10.1109/access.2020.2966328.
- [5] 李秀智,李家豪,张祥银,等. 基于深度学习的机器人最优抓取姿态检测方法[J]. 仪器仪表学报, 2020, 41(5): 108-117.

- [6] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 779-788.
- [7] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. Yolov4: Optimal speed and accuracy of object detection[J]. ArXiv Preprint, ArXiv: 2004.10934, 2020:332-343.
- [8] 冯文宇,张宇豪,张堃,等.复杂环境下课堂多人状态检测算法研究[J].电子测量与仪器学报,2021,35(6):53-62.
- [9] 刘素行,吴媛,张军军.基于 YOLO v3 的交通场景目标检测方法[J].国外电子测量技术,2021,40(2):116-120.
- [10] 仇娇慧,贝绍轶,尹明锋,等.基于改进 YOLOv5s 的齿轮表面缺陷检测[J].现代制造工程,2022,(3):104-113.
- [11] 牟亮,赵红,李燕,等.基于梯度压缩的 YOLO v4 算法车型识别[J].工程科学学报,2022,44(5):940-950.
- [12] 朱晓春,陈子涛.基于改进型 YOLO v5 算法的安全帽佩戴检测[J].南京工程学院学报(自然科学版),2021,19(4):7-11.
- [13] 程婧怡,段先华,朱伟.改进 YOLOv3 的金属表面缺陷检测研究[J].计算机工程与应用,2021,57(19):252-258.
- [14] 郭磊,王邱龙,薛伟,等.基于改进 YOLOv5 的小目标检测算法[J].电子科技大学学报,2022,51(2):251-258.
- [15] 邢晋超,潘广贞.改进 YOLOv5s 的手语识别算法研究[J/OL].计算机工程与应用,2022,06(12):1-15.
- [16] ZHANG Y F, REN W, ZHANG Z, et al. Focal and efficient IOU loss for accurate bounding box regression[J]. ArXiv Preprint,2021, ArXiv:2101.08158.

作者简介

强栋, 硕士, 主要研究方向为移动增强现实、多目标识别等。

E-mail:1239208384@qq.com

王占刚, 博士, 副教授, 主要研究方向为虚拟现实与三维模拟、智慧感知与信息处理等。

E-mail:wangzg@bistu.edu.cn