

DOI:10.19651/j.cnki.emt.2520467

面向家庭服务机器人的改进 YOLOv10 生活物品检测算法^{*}

葛晓亮^{1,2} 左云波^{1,2} 陈 赛^{1,2} 王少红^{1,2} 苗卓然^{1,2}

(1. 北京信息科技大学现代测控技术教育部重点实验室 北京 100192; 2. 北京信息科技大学机电工程学院 北京 100192)

摘要: 针对智能家庭服务机器人在生活物品识别任务中,目标存在的多尺度、多类别及密集分布等复杂检测挑战,提出一种改进的 YOLOv10-DLC 目标检测算法,设计特征融合多样分支模块(C2f-DBB),提高算法网络对不同尺度目标的检测能力;嵌入快速空间金字塔池化与大核分离卷积注意力相结合的模块(SPPF-LSKA),提升算法网络对特征空间结构的理解能力,突出目标关键特征;引入内容感知特征重组(CARAFE)算子提升网络在图像上采样过程中的细节重建精度。实验分析表明,改进后的 YOLOv10-DLC 算法有效提高了生活物品目标检测的准确性和效率,其平均精度均值(mAP)值达到 94.3%,较原算法网络模型提升 3.6%。

关键词: 深度学习;生活物品;目标检测;改进的 YOLOv10-DLC

中图分类号: TN29 **文献标识码:** A **国家标准学科分类代码:** 520.6040

Improved YOLOv10 algorithm for daily object detection oriented towards home service robots

Ge Xiaoliang^{1,2} Zuo Yunbo^{1,2} Chen Sai^{1,2} Wang Shaohong^{1,2} Miao Zhuoran^{1,2}

(1. Key Laboratory of Modern Measurement and Control Technology, Ministry of Education, Beijing Information Science and Technology University, Beijing 100192, China; 2. Mechanical Electrical Engineering School, Beijing Information Science and Technology University, Beijing 100192, China)

Abstract: Against the complex detection challenges posed by the multi-scale, multi-category, and dense distribution of objects in the daily item recognition task for intelligent household service robots, an improved YOLOv10-DLC object detection algorithm is proposed: By designing a diverse branch module for feature fusion (C2f-DBB), the algorithm network's ability to detect objects of different scales is enhanced; By adopting a combined module of fast spatial pyramid pooling and large kernel separable convolution attention (SPPF-LSKA), the algorithm network's ability to understand the spatial structure of features is improved, and the key features of objects are highlighted; By using the content-aware feature reassembly (CARAFE) operator, the accuracy of detail reconstruction during image upsampling in the network is improved. Experimental analysis shows that the improved YOLOv10-DLC algorithm has improved the accuracy of object detection, with its mAP value reaching 94.3%, an increase of 3.6% compared with the original algorithm network model.

Keywords: deep learning; daily objects; object detection; improved YOLOv10-DLC

0 引言

根据国家统计局数据,我国 60 岁及以上老年人口已突破 3 亿,人口老龄化趋势显著加剧。在此背景下,家庭服务机器人作为缓解养老压力、提升老年群体生活质量的重要技术载体,受到广泛关注^[1-2]。物品识别与抓取是家庭服务

机器人的核心功能之一,直接影响其实用性。然而,家庭场景具有物品种类繁多、摆放位置随机、光照条件复杂等特点,现有目标检测算法在面对小尺度物体密集分布、多类别物品混杂等情况时存在漏检率高、误检严重等问题^[3-4],难以满足实际应用需求。因此,研究适用于家庭场景的高精度物品检测算法具有重要的理论价值和现实意义。

收稿日期:2025-12-25

^{*} 基金项目:国家重点研发计划课题项目(2020YFB1713203)资助

点,为目标检测算法带来严峻的挑战。针对上述难点,本文对YOLOv10算法做出如下改进。改进后的YOLOv10-DLC网络结构如图2所示。

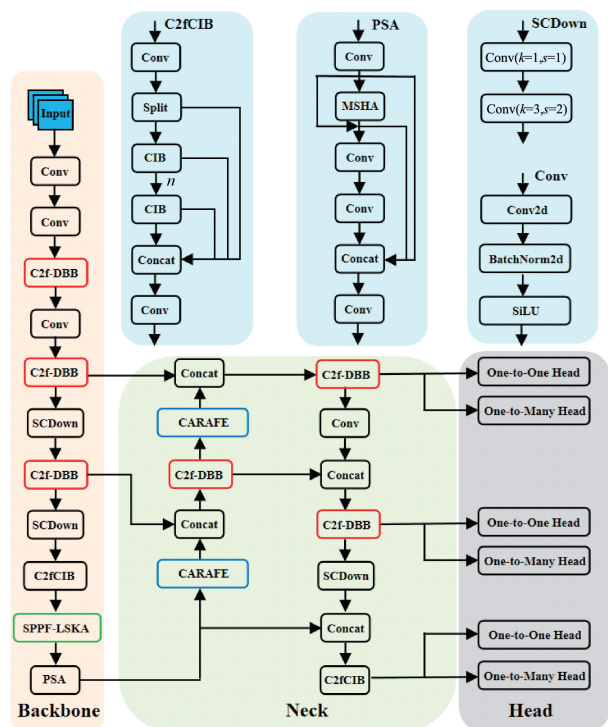


图2 改进后的YOLOv10-DLC网络结构

Fig. 2 Improved YOLOv10-DLC network architecture

本文在网络前端设计C2f-DBB模块,通过多分支结构和重参数化技术强化密集场景中物体的多尺度局部特征提取,构建强判别性的特征基底;在网络中层集成SPPF-LSKA引入大核注意力机制,拓展感受野的同时实现特征重加权,有效抑制生活场景中的背景噪声并增强小目标的突出性;在网络解码端融入CARAFE内容感知上采样,其能基于增强特征精准预测重组核,在保持生活物品高频细节的同时修正了边界形变并缓解目标粘连。通过这样的“表征增强—上下文聚焦—细节重构”的系统优化链路,形成显著的协同增益效果,有效的提高了密集场景中物体的检测成功率。

1) C2f-DBB 模块优化

在实际环境中,生活物品往往呈现出尺度差异显著、形态各异且分布密集的特点,这对算法捕捉多尺度特征的能力提出了极高要求。原有的C2f模块虽然通过堆叠BottleNeck瓶颈模块来提取特征,但其内部的跨层连接叠加信息粒度较高,层间缺乏直接的交互与关联,导致特征依赖性较差,难以充分提取并融合不同尺度的目标信息,容易造成漏检。为了解决这一问题,增强网络对多尺度目标的感知与表达能力,本文引入多样分支结构,设计了特征融合多样分支模块(C2f-DBB)。

图3(a)所示为YOLOv10中C2f的整体模块结构,由

常规卷积层、批量归一化层、激活函数层、Split拆分层及 N 个BottleNeck瓶颈模块构成。图3(b)所示为BottleNeck瓶颈模块,其中残差块根据网络位置动态添加或删除。

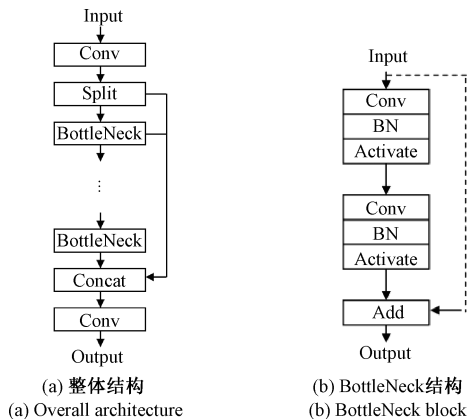


图3 C2f结构

Fig. 3 C2f structure

多样分支模块(diverse branch block, DBB)^[16]采用了多并行融合结构,训练时通过多分支路径获取不同的感受野和多尺度特征,显著增强了模型的多尺度信息表达能力,在尺度信息的融合上更为丰富,如图4(a)所示。推理时,多分支通过多种结构重参数化方法合并为单个卷积,实现推理等价转换,避免了多分支融合导致的推理性能降低,如图4(b)所示。

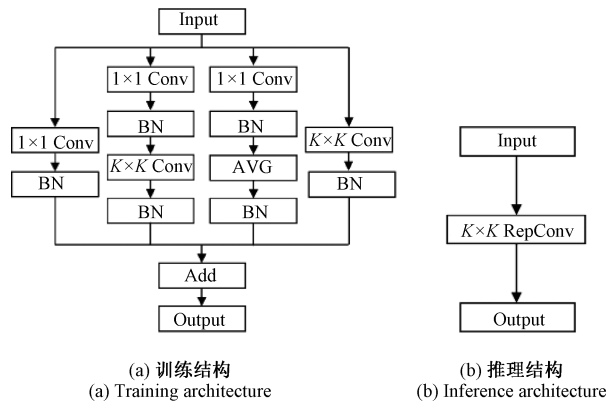


图4 DBB模块结构

Fig. 4 DBB module structure

本文用DBB结构替换原C2f模块BottleNeck中的卷积层设计出C2f-DBB模块,如图5(a)和(b)所示,该模块能够在同一卷积层上实现多尺度特征的并行融合,细化多尺度信息,通过扩大网络的感受野与强化特征融合,精准捕捉各类目标的关键特征,降低漏检率。

2) SPPF-LSKA 模块优化

快速空间金字塔池化模块(spatial pyramid pooling fast, SPPF)是YOLOv10网络中的关键组件,在SPPF层中,其使用了3个连续的最大池化操作提取特征,但最大池化仅能捕获最大值,会导致细微特征丢失。对于小尺度目

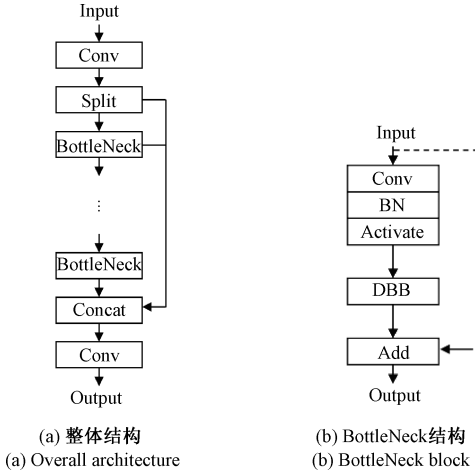


图 5 C2f-DBB 结构

Fig. 5 C2f-DBB module structure

标物体,由于其特征常常被复杂背景遮挡或不易被识别,问题尤为突出。为了解决上述问题,将大核分离卷积(large separable kernel attention, LSKA)注意力机制^[17]集成到 SPPF 层中,形成如图 6 所示的优化后的 SPPF-LSKA 结构模块。

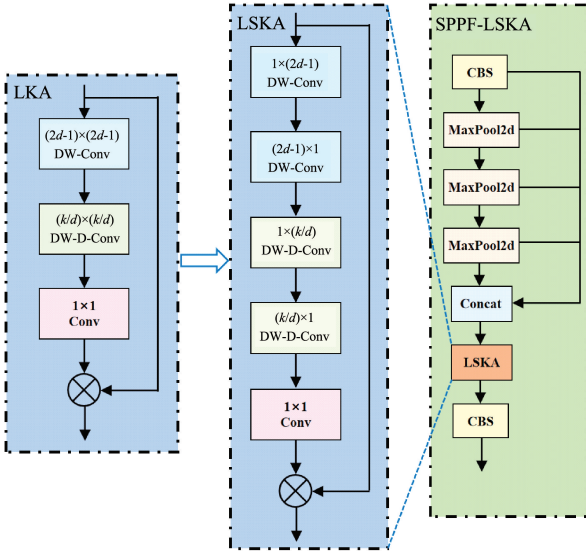


图 6 SPPF-LSKA 模块结构

Fig. 6 SPPF-LSKA module structure

LSKA 注意力机制的相关计算如式(1)~(4)所示。

$$\bar{Z}^C = \sum_{H,W} W_{(2d-1) \times 1}^C * \left(\sum_{H,W} W_{1 \times (2d-1)}^C * F^C \right) \quad (1)$$

$$Z^C = \sum_{H,W} W_{\frac{k}{d} \times 1}^C * \left(\sum_{H,W} W_{1 \times \frac{k}{d}} * \bar{Z}^C \right) \quad (2)$$

$$A^C = W_{1 \times 1} * Z^C \quad (3)$$

$$\bar{F}^C = A^C \otimes F^C \quad (4)$$

其中, $\bar{Z}^C$ 表示深度卷积后的输出,求和符号下方的 H 和 W 表示特征图中的高度和宽度, A^C 表示注意力图,表示向下取整操作, d 表示扩展率, k 是核大小, $*$ 表示卷积, $\otimes$

表示 Hadamard 运算, F^C 表示输入特征图, C 表示输入通道数。

根据上述公式和原理图可知,LSKA 将大核注意力(large kernel attention, LKA)的前两层分解为 4 层,每层有两个一维卷积层。LSKA 注意力机制首先使用一维卷积核进行空间分离卷积,然后利用得到的特征图生成注意力图,最后将注意力图应用于原始特征图,此过程有助于网络更高效地处理大范围的空间信息。因此,该改进方式在几乎不增加计算负担的情况下,丰富了跨多尺度特征的感受野信息,改善了特征融合。

3) CARAFE 上采样

内容感知特征重组模块(CARAFE)^[18]解决了传统特征上采样方法在保留细节和重建语义信息方面的局限性。其主要功能是对输入特征图进行上采样,同时确保语义信息的完整性。

如图 7 所示,对于输入特征图 $H \times W \times C$,内核预测模块会生成一个大小为 $\sigma H \times \sigma W \times \sigma C$ 的上采样特征图,其中 $\sigma \in \mathbb{Z}$ 。

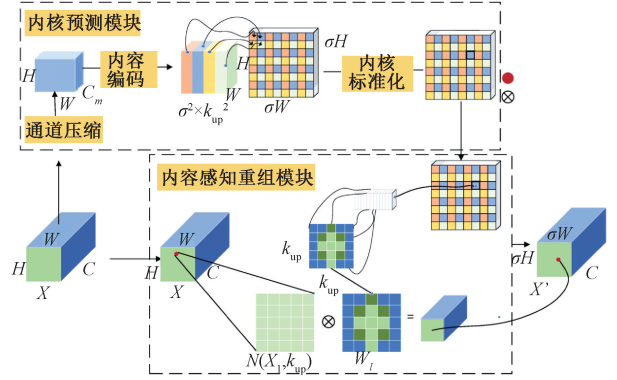


图 7 CARAFE 原理

Fig. 7 CARAFE schematic

输出特征图中的目标点坐标对应于原始坐标的坐标关系如式(5)所示。

$$\begin{cases} l = (i, j), l' = (i', j') \\ i = \frac{i'}{\sigma}, j = \frac{j'}{\sigma} \end{cases} \quad (5)$$

其中, σ 代表下采样率, l 是上采样前的坐标, l' 是上采样后的坐标, (i, j) 表示采样前的原始坐标, (i', j') 表示采样后的坐标。

以示例点为例,内核预测模块对目标进行内容编码、重组内核及归一化处理,内容感知重组模块确定该点生成大小为 $k_{up} \times k_{up}$ 的重组内核,并且以该点为中心的区域 $N(X_1, k_{up})$ 中对所有点执行特征重建,最终的输出结果如式(6)和(7)所示。

$$X'_{l'} = \sum_{n=-r}^r \sum_{m=-r}^r W_{l'(n,m)} * X_{(i+n, j+m)} \quad (6)$$

$$r = \frac{k_{up}}{2} \quad (7)$$

其中, $\mathbf{W}_{l'(n,m)}$ 表示该点的重组内核, $\mathbf{X}'_{l'}$ 是输出后的特征图, $X_{(i+n,j+m)}$ 是相邻区域的像素点, k_{up} 是重组内核的大小。

因此,将 CARAFE 算子集成到 YOLOv10 中可以使模型更好地整合来自不同区域的语义信息能力,并增强模型对物体细节特征的提取和重建能力。

2 实验设计与分析

2.1 实验环境配置

实验采用的环境配置如表 1 所示。本文所有对比实验均在此配置下进行,以保证结果的可比性。

表 1 实验环境配置

参数	配置
CPU	Intel(R) Core(TM) i7-12700
GPU	NVIDIA GeForce RTX 3090
系统环境	Windows 10
语言	Python 3.7
学习框架	Pytorch 1.13.1
加速环境	CUDA 11.6

2.2 数据集构建与统计分析

1)数据集构建与划分

为验证模型在家庭服务机器人典型应用场景中的检测性能,本文构建了面向家庭场景的生活物品检测数据集。依据目标使用场景将数据集划分为厨房用品、文具用品、个人护理、食品与日用杂物等 5 类,涵盖球体、长方体、圆柱体等多形态目标,能够用于评估模型对多尺度、多形态物体的表征与定位能力。

数据来源包括相机实拍与网络采集,覆盖不同拍摄角度、光照条件以及单目标/多目标摆放情形。为增强模型对真实环境扰动的鲁棒性,采用随机翻转、缩放、噪声与模糊等增强策略,流程如图 8 所示。数据集共标注 9 000 余张图像,并按 8:1:1划分为训练集、验证集与测试集。

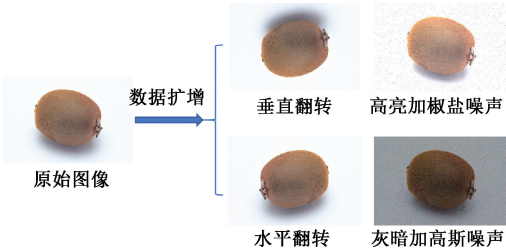


图 8 数据增强示意图

Fig. 8 Data enhancement

2)数据集规模与类别统计

表 2 汇总了数据集按场景类别统计的图像数与实例数信息。其中,“图像数”指包含该场景类别实例物品的图像

数量(各类别之间不互斥),此处图像数之和会大于数据集总图像数。“实例数”指在该场景类别下的实例物品标注框的总数量。

表 2 自建生活物品数据集统计信息

Table 2 Statistics of the proposed daily-object dataset

场景类别	物品类别	图像张数	实例个数
厨房用品类	芥末酱、汤勺等	2 317	11 457
文具用品类	笔、橡皮等	2 194	14 263
个人护理类	沐浴露、药瓶等	2 583	10 129
食品类	香蕉、猕猴桃等	2 645	12 984
日用杂物类	固体胶、网球等	1 892	5 280
总计		9 264	54 113

3)数据集复杂度量化

为定量刻画生活物品数据集的复杂度,基于检测框统计目标尺度、图像密集度等指标参数。

(1)目标尺度,设图像分辨率为 $W \cdot H$,标注框宽高为 w, h ,则目标尺度 s 定义为:

$$s = (w \cdot h) / (W \cdot H) \tag{8}$$

将目标尺度划分为 3 类,Small ($s < 0.01$)、Medium ($0.01 \leq s < 0.05$)与 Large ($s \geq 0.05$),分别占比为 25.8%、56.0%和 18.2%。

(2)图像密集度,以单张图像的目标物品实例数 n 表示该张图像的密集程度,统计 n 的均值和分位数 P_{95} 分别为 5.84 和 10。

上述统计表明,本数据集具有显著的多尺度分布与高密度特征,能充分检验算法在复杂场景下的鲁棒性。

2.3 算法评价指标

本文涉及到的评价指标有精确率(precision, P)、召回率(recall, R)、平均精度(average precision, AP)、计算量(giga floating-point operations per second, GFLOPs)以及帧处理速率(frames per second, FPS)等。如式(9)~(12)所示。

$$P = \frac{TP}{TP + FP} \tag{9}$$

$$R = \frac{TP}{TP + FN} \tag{10}$$

$$AP = \int_0^1 p(x) dx \tag{11}$$

$$mAP = \frac{1}{N} \sum_{i=1}^N AP_i \tag{12}$$

式中:TP 表示检测到且确实存在的生活物品的样本数量;FP 表示事实上不存在但被检测为生活物品的样本数量;FN 表示事实上存在但被检测为不存在的生活物品的样本数量;N 表示类别数。

2.4 基于改进的 YOLOv10-DLC 目标检测消融效果对比

基于上述实验环境,设定以下的训练参数:训练轮数

(epoch)为 100,批处理大小(batch-size)为 64,图片输入尺寸为 640×640,初始学习率为 0.01,在此参数配置下对 YOLOv10 网络模型进行训练。为进一步量化说明不同改进方法对网络

性能的影响,分别对单个改进和整合改进进行消融实验对比,实验结果如表 3 所示。其中,CD 代表 C2f-DBB 模块,SL 代表 SPPF-LSKA 模块,CA 代表 CARAFE 模块。

表 3 不同改进方法实验效果对比
Table 3 Comparison of different improvement methods

组别	CD	SL	CA	$P/\%$	$R/\%$	mAP50/%	GFLOPs	FPS/fps
1	—	—	—	88.7	86.9	90.7	6.5	317
2	✓	—	—	89.8	88.5	91.7	6.5	345
3	—	✓	—	89.3	87.6	91.3	6.5	332
4	—	—	✓	90.1	88.7	92.2	6.8	289
5	✓	✓	—	90.8	89.5	92.9	6.5	351
6	✓	—	✓	91.9	90.8	93.8	6.8	304
7	—	✓	✓	91.6	90.4	93.5	6.8	296
8	✓	✓	✓	92.5	91.2	94.3	6.8	310

由表 3 可知,单独引入 C2f-DBB(组 2)或 SPPF-LSKA (组 3)时,模型在维持 6.5 的 GFLOPs 低计算负载的前提下,mAP50 分别提升了 1.0%与 0.6%,并且 FPS 也得到了提升。表现为 C2f-DBB 通过训练期的多分支获取密集场景中物品更丰富的尺度信息,并在推理阶段等价折叠为单分支结构,从而在增强表征能力的同时避免了明显的推理开销;SPPF-LSKA 利用大核分解卷积注意力扩大了空间感受野,有效抑制了密集场景的背景干扰,改善了漏检问题,并使召回率提升 0.7%。引入 CARAFE(组 4)虽然轻微增加了计算量,但其内容感知上采样机制有效缓解了密集场景中物品的特征错位与边界模糊,并且带来 1.5%的 mAP50 精度增益。

此外,组合改进呈现出了更显著的协同效果。以 C2f-DBB 与 CARAFE 的组合(组 6)为例,其联合提升幅度的 3.1%显著高于二者单项增益之和的 2.5%。表现为 C2f-DBB 提取的高判别性语义特征在通过 CARAFE 进行尺度变换时,得益于自适应重组核的高保真映射,避免了传统插值带来的特征平滑与混叠,实现了优质特征生成与无损传输的高度互补。最终,3 个模块的协同组合(组 8)在仅增加少量参数的情况下,以 310 的 FPS 高实时性达到了 94.3%的最优精度,验证了该改进方案的高效性与鲁棒性。

为了展示改进前后的 YOLOv10 算法性能变化情况,本文绘制了部分改进方案的精确率、召回率和平均精度曲线图,分别如图 9~11 所示。

其中,YOLOv10n、YOLOv10-D、YOLOv10-L、YOLOv10-C 和本文分别对应为 YOLOv10 的基础模型、单独引入 C2f-DBB 模块的模型、单独引入 SPPF-LSKA 模块的模型、单独引入 CARAFE 上采样算子的模型和本文最终的改进模型。由图 9~11 可知,改进后的 YOLOv10-DLC 算法模型在精确率、召回率和平均精度方面提升相对明显,表现

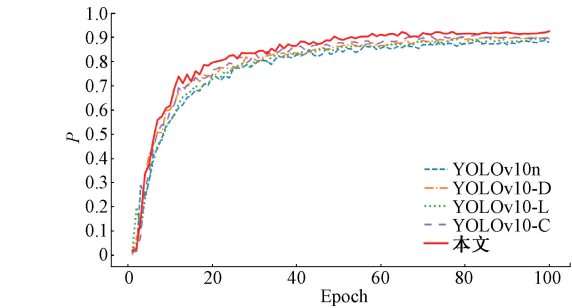


图 9 精确率对比曲线
Fig. 9 Comparison curve of accuracy

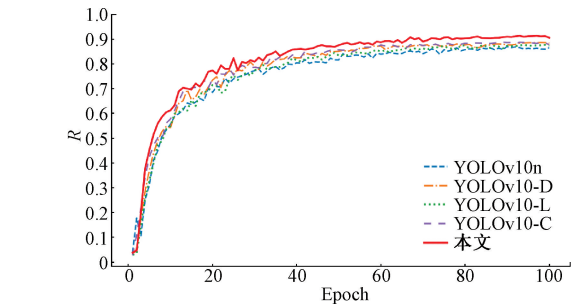


图 10 召回率对比曲线
Fig. 10 Comparison curve of recall rate

较好。

为了更直观地对比 YOLOv10 原始模型与改进后的 YOLOv10-DLC 模型的检测效果,在测试集中随机抽取一组样本进行可视化效果对比,如图 12 和 13 所示。

图 12 中出现漏检情况,在密集场景下,处于角落或轻微遮挡状态的 1 个药瓶和 2 个橡皮等小目标物体未能被成功检测。相对地,图 13 中,这些物体不仅被成功检测,而且多数物体的置信度分数也有所提高。

通过检测效果对比,可以清晰地看到改进后的

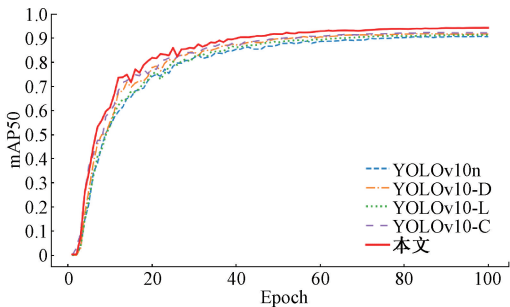


图 11 平均精度对比曲线

Fig. 11 Comparison curve of average accuracy

表 4 不同算法对比实验

Table 4 Comparison of different algorithms

模型	mAP50/%	FPS/fps
SSD	78.5	125
Faster R-CNN	82.2	78
YOLOv5s	85.3	242
YOLOv8n	88.4	284
YOLOv10n	90.7	317
YOLOv11n	91.4	310
YOLOv10-DLC(本文)	94.3	310

实验分析表明,改进的 YOLOv10-DLC 算法在精确度方面优于其他原始算法。与最新的 YOLOv11n 模型相比,在 FPS 帧率持平的情况下,mAP 值高出 2.9%。

同时,为进一步验证本文方法在家居生活物品检测方面的优越性,在相同的数据集和实验环境下,将与本文应用场景接近的文献[12-13]进行对比,结果如表 5 所示。由表 5 可知,本文改进后的方法在检测精度和速度方面的表现较好。

表 5 不同方法对比结果

Table 5 Comparison of different methods

模型	mAP50/%	FPS/fps
文献[12]	91.2	293
文献[13]	89.9	271
本文方法	94.3	310

此外,对于堆叠遮挡场景的目标检测,如大物体压覆小物体等,目前仍是图像识别领域的难点。考虑到家庭服务机器人有更高的自由度,当出现堆叠遮挡时,家庭服务机器人能通过调整视角或重新打乱物品位置等方式以应对遮挡等复杂场景,且对于如何更高效的解决堆叠遮挡场景的检测问题,也将是后续的研究重点。

3 结 论

针对智能家庭服务机器人生活物品识别任务中,目标存在的多尺度、多类别及密集分布等复杂检测挑战,现有的目标检测方法效果不佳。本文提出了一种面向家庭服务机器人的改进的 YOLOv10 生活物品检测算法。设计 C2f-DBB 模块、SPPF-LSKA 模块以及 CARAFE 上采样算子,对 YOLOv10 目标检测算法进行改进,旨在增强目标检测模型对多尺度、多类别物体的识别能力。通过进行消融实验表明,改进后的 YOLOv10-DLC 模型与改进前的 YOLOv10 模型相比,在检测速度几乎不变的情况下,检测精度 mAP 增长了 3.6%,实现了速度与精度的优化平衡。此外,通过与其他目标检测算法及其他现有方法进行对比分析,进一步验证了改进的 YOLOv10-DLC 算法的有效



图 12 原模型检测效果

Fig. 12 Original model detection effect

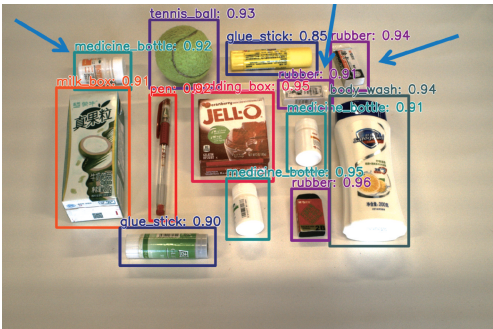


图 13 改进后模型检测效果

Fig. 13 Improved model detection performance

YOLOv10-DLC 模型因其更宽广的感受野、卓越的特征融合能力以及对关键细节特征的专注和重建能力,在密集场景中对不同尺度和类别的多目标物体特征实现了更精细的计算和理解,检测结果精度较高且更为稳定。

2.5 基于改进的 YOLOv10-DLC 目标检测对比其他算法

为了更好地评估改进后的 YOLOv10-DLC 算法性能,并且再次验证本文改进的目标检测算法相较于其他目标检测算法在检测精度和速度上的性能表现,选取了 SSD、Faster R-CNN、YOLOv5s、YOLOv8n、YOLOv10n、YOLOv11n 以及本文提出的改进的 YOLOv10-DLC 算法模型,在同一数据集上展开训练。以平均精度和帧率作为评价指标,对这些算法的性能进行评估,实验结果如表 4 所示。

性,在密集场景下多尺度和多类别物体检测方面具有重大应用价值和现实意义。

本文方法对密集场景下非重叠的多尺寸和多类别物体检测有较好的效果,但对于堆叠场景检测仍是一个挑战,这将是未来工作的进一步研究重点。

参考文献

- [1] 孙伟平. 智能机器人的主体地位, 权利, 责任和义务问题[J]. 华东师范大学学报(哲学社会科学版), 2024, 56(1): 1-2, 175.
Sun Weiping. Subjectivity, rights, responsibility and obligation of intelligent robots [J]. Journal of East China Normal University (Humanities and Social Sciences), 2024, 56(1): 1-2, 175.
- [2] 陶永, 刘海涛, 王田苗, 等. 我国服务机器人技术研究进展与产业化发展趋势[J]. 机械工程学报, 2022, 58(18): 56-57.
Tao Yong, Liu Haitao, Wang Tianmiao, et al. Research progress on technology and industrialization development trend of service robots in China[J]. Journal of Mechanical Engineering, 2022, 58(18): 56-57.
- [3] 石杰, 周亚丽, 张奇志. 基于改进 Mask RCNN 和 Kinect 的服务机器人物品识别系统[J]. 仪器仪表学报, 2019, 40(4): 216-218.
Shi Jie, Zhou Yali, Zhang Qizhi. Service robot object recognition system based on improved Mask RCNN and Kinect [J]. Chinese Journal of Scientific Instrument, 2019, 40(4): 216-218.
- [4] Yan K Y, Wang J, Wang J, et al. An efficient improved YOLOv5-based method for detecting iron waste in ores[J]. Instrumentation, 2025, 12(2): 37-39.
- [5] Ren S Q, He K M, Girshick R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016, 39(6): 1139-1142.
- [6] Yang F C, Huang L D, Tan X W, et al. FasterNet-SSD: A small object detection method based on SSD model[J]. Signal, Image and Video Processing, 2024, 18(1): 174-177.
- [7] Wang G, Chen Y F, An P, et al. UAV-YOLOv8: A small-object-detection model based on improved YOLOv8 for UAV aerial photography scenarios[J]. Sensors, 2023, 23(16): 7190.
- [8] Guan S T, Lin Y M, Lin G Y, et al. Real-time detection and counting of wheat spikes based on improved YOLOv10[J]. Agronomy, 2024, 14(9): 1936.
- [9] Badgujar C M, Poulouse A, Gan H. Agricultural object detection with you only look once(YOLO) algorithm: A bibliometric and systematic literature review [J]. Computers and Electronics in Agriculture, 2024, 223: 109090.
- [10] 张瑶, 张俊三, 马俊朋, 等. CAFR-YOLO: 基于 YOLOv8 的多尺度目标检测算法[J/OL]. 计算机工程, 2025-11-18. <https://doi.org/10.19678/j.issn.1000-3428.0252821>.
Zhang Yao, Zhang Junsan, Ma Junpeng, et al.

- CAFR-YOLO: Multi-scale object detection algorithm based on YOLOv8 [J/OL]. Computer Engineering, 2025-11-18. <https://doi.org/10.19678/j.issn.1000-3428.0252821>.
- [11] 文思予, 张上, 张朝阳, 等. 基于改进 YOLOv8s 的多尺度检测算法[J]. 现代电子技术, 2024, 47(15): 134-136, 138.
Wen Siyu, Zhang Shang, Zhang Zhaoyang, et al. Multi-scale detection algorithm based on improved YOLOv8s[J]. Modern Electronics Technique, 2024, 47(15): 134-136, 138.
- [12] 任梦哈, 赵海燕, 宋佳智. 基于改进 YOLO11 的生活垃圾检测模型[J]. 电子测量技术, 2026, 49(1): 247-256.
Ren Menghan, Zhao Haiyan, Song Jiazi. Domestic waste detection model based on improved YOLO11[J]. Electronic Measurement Technology, 2026, 49(1): 247-256.
- [13] 田海峰, 邱茂顺, 张维健. 基于 YOLO-DA 的商品识别算法[J]. 曲阜师范大学学报(自然科学版), 2025, 51(3): 75-77, 79-80.
Tian Haifeng, Qiu Maoshun, Zhang Weijian. Commodity recognition algorithm based on YOLO-DA[J]. Journal of Qufu Normal University (Natural Science), 2025, 51(3): 75-77, 79-80.
- [14] 赵宗扬, 康杰虎, 吴斌, 等. 基于 FRL-Net 的高鲁棒性多尺度小样本轨道入侵异物检测方法研究[J]. 仪器仪表学报, 2024, 45(1): 245-249.
Zhao Zongyang, Kang Jiehu, Wu Bin, et al. Research on the high robust multi-scale few-shot railway intrusion obstacles detection method based on FRL-Net[J]. Chinese Journal of Scientific Instrument, 2024, 45(1): 245-249.
- [15] 顾杨海, 李富, 陈德基, 等. 基于多尺度特征融合与交互的路侧目标检测算法[J]. 电子测量技术, 2024, 47(23): 155-159, 161.
Gu Yanghai, Li Fu, Chen Deji, et al. Roadside object detection algorithm based on multi-scale feature fusion and interaction[J]. Electronic Measurement Technology, 2024, 47(23): 155-159, 161.
- [16] Ding X H, Zhang X Y, Han J G, et al. Diverse branch block: Building a convolution as an inception-like unit [C]//IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021: 10884-10886, 10890.
- [17] Tie J, Zhu C G, Zheng L, et al. LSKA-YOLOv8: A lightweight steel surface defect detection algorithm based on YOLOv8 improvement [J]. Alexandria Engineering Journal, 2024, 109: 207-209, 212.
- [18] Wang J Q, Chen K, Xu R, et al. Carafe: Content-aware reassembly of features [C]//IEEE/CVF International Conference on Computer Vision, 2019: 3008-3010, 3015-3016.

作者简介

葛晓亮, 硕士研究生, 主要研究方向为图像处理和机械臂抓取技术。

E-mail: gxl886688@163.com

左云波(通信作者), 博士, 副教授, 硕士生导师, 主要研究方向为机器视觉检测技术。

E-mail: zuoyunbo@bistu.edu.cn