

基于改进 YOLOv8n 的散养鸭蛋识别与定位

龚东军^{1,2}, 林卫国^{1,3*}, 杨浩^{1,3}, 陈瑞^{1,3}, 王巧华^{1,3}, 王树才^{1,3}

(1. 华中农业大学工学院, 武汉 430070; 2. 武汉软件工程职业学院/武汉开放大学, 武汉 430205;

3. 农业农村部长江中下游农业装备重点实验室, 武汉 430070)

摘要:为解决鸭蛋捡拾机器人对中远距离散养鸭蛋识别困难与定位精度差等问题, 该研究先以 YOLOv8n 原始模型为基础, 采用添加针对小目标检测的 P2 检测头的升级模型 YOLOv8n-P2, 使模型能够在目标特征保存更完整的大尺寸特征图进行检测, 并引入卷积块注意力机制模块 (convolutional block attention module, CBAM) 与 Involution 内卷算子, 增强特征图不同通道之间的联系, 提高模型网络在通道维度上的感知能力和对小目标的检测能力。其次, 在图像目标识别的基础上, 提出了一种基于机器人操作系统 (robot operating system, ROS) 的视觉定位方法, 通过机器视觉识别与坐标变换 (transform, TF), 实现了鸭蛋捡拾机器人搭载深度相机在运动过程中对目标鸭蛋的高精度实时视觉定位。试验结果表明, 改进模型 YOLOv8n-P2-CBAM-INV 的权重文件大小为 20.80 MB, 每秒浮点计算量为 51.4 G, 单张图像平均检测用时 14.5 ms, 精确率和召回率分别为 98.30% 和 96.60%, 相较于原始模型 YOLOv8n, 在测试集上平均精度均值提高了 7.1 个百分点, 该模型在 6.0 m 距离测试, 平均精度均值达到 98.0%, 相较于原始模型 YOLOv8n 提高了 29.4 个百分点。最后, 通过在 0.5~3.0 m 范围的移动机器人视觉定位试验表明, 该定位方法可实现移动捡蛋机器人在实验室内对中远距离鸭蛋的高精度实时运动定位, 最大定位误差可控制在 0.03 m 以内。此改进模型可部署到移动边缘计算图像处理平台, 不仅可为移动捡蛋机器人提供支持, 也可为类似的自主移动采收作业机器人提供一定借鉴。

关键词: 目标检测; 视觉定位; 鸭蛋; 识别; YOLOv8n; 机器人操作系统

doi: 10.11975/j.issn.1002-6819.202407156

中图分类号: S126; TP391.4

文献标志码: A

文章编号: 1002-6819(2025)-01-0161-10

龚东军, 林卫国, 杨浩, 等. 基于改进 YOLOv8n 的散养鸭蛋识别与定位[J]. 农业工程学报, 2025, 41(1): 161-170. doi: 10.11975/j.issn.1002-6819.202407156 <http://www.tcsae.org>

GONG Dongjun, LIN Weiguo, YANG Hao, et al. Recognition and localization of free-range duck eggs based on improved YOLOv8n[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2025, 41(1): 161-170. (in Chinese with English abstract) doi: 10.11975/j.issn.1002-6819.202407156 <http://www.tcsae.org>

0 引言

中国是世界上第一大蛋鸭养殖和消费国, 具有悠久的蛋鸭养殖历史。蛋鸭养殖业是禽蛋产业的重要组成部分, 是中国农业的支柱产业之一。近年来, 中国蛋鸭存栏量已经达到 3 亿羽左右, 并且数量仍然保持稳定增长趋势, 截至 2022 年, 中国年产鸭蛋总量约 600 万 t, 占世界总量 90% 以上^[1-3]。中国蛋鸭养殖产业体量大, 但存在分散养殖、自动化程度低等问题。目前, 国内大部分蛋鸭养殖企业仍依赖人工捡拾鸭蛋, 工作效率低且劳动强度大, 还可能因人员频繁进出鸭场会给蛋鸭健康带来风险。鉴于上述问题, 开发一种能够自动化完成散养鸭蛋捡拾工作的机器人显得尤为重要。苑进等^[4-5]开展了禽蛋自动捡拾机械手的系统设计和农业机器人应用相关

技术研究, 阳康等^[6-8]开展了基于 ROS (robot operating system) 轻量化机械臂系统研究开发。同时, 畜禽养殖作业机器人的相关研究也为鸭蛋捡拾机器人的研发提供了有益的参考^[9-10]。特别是吕惠芳^[11]开展了基于单目视觉的散养鸭蛋识别与定位研究, 为实现鸭蛋的准确识别与定位提供了关键技术参考。

在图像目标检测领域, 基于深度神经网络的目标检测方法在解决复杂环境下的目标识别难题方面取得了显著进展。SSD (single shot multi-box detector)、DSSD (deconvolutional single shot detector) 等算法提高了目标检测任务的效果和效率^[12-13]。REDMON 等^[14-16]提出 YOLO (you only look once) 目标检测算法取得良好的效果。近年来, YOLO 系列目标检测算法改进研究进一步提高了目标检测任务的效果和效率^[17-20]。祝志慧等^[21]针对鸭场中复杂情境下鸭蛋的识别问题, 在 YOLOv7 主干网络加入卷积注意力机制模块 (convolutional block attention module, CBAM) 提高模型网络对目标特征的敏感度, 降低环境中干扰因素对鸭蛋识别精度的影响, 优化后的网络训练出的模型平均精度达到 85.2%, 与 YOLOv7 原始模型相比, 对光线昏暗、覆盖遮挡等复杂情况下的鸭蛋识别的平均精度值提高了 8.1 个百分点。此外, 多位学者基于 YOLOv8n 网络模型在农作物果实、农畜产品

收稿日期: 2024-07-17 修订日期: 2024-11-06

基金项目: 国家自然科学基金项目 (32372426); 中央高校基本科研业务费专项资金资助项目 (2662022GXJ004); 武汉市产学研成果项目 (CXZY202016)

作者简介: 龚东军, 博士生, 讲师, 研究方向为农业机器人自动化。

Email: dongjun.81@163.com

※通信作者: 林卫国, 博士, 副教授, 研究方向为农业机械化技术及装备。

Email: linweiguo@mail.hzau.edu.cn

的目标检测领域,进行相关改进研究以提高检测速度和准确性,取得了良好的识别效果^[22-26]。李茂等^[27]采用 C2F-PConv 模块对 YOLOv8 模型进行轻量化改进,相比原基础网络 YOLOv8s 模型权重减小 46.22%,将模型部署到嵌入式设备,检测帧率达到 43 帧/s。林华浦等^[28]针对现代化养猪场群猪计数过程中远距离猪只易漏检的问题,在 YOLOv7 目标检测网络中添加小目标尺度检测层,同时使用相似性自适应注意力机制模块(similarity-aware activation module, SimAM)调整权重,并通过引入残差模块和平均池化操作的方式对网络中卷积模块进行改进,模型平均精度提高了 5.1 个百分点,检测速度提升 7.14 个百分点。彭炫等^[29]针对棉花顶芽的识别问题,在 YOLOv5s 模型网络的基础上添加小目标检测层,并引入 CPP-CBAM 注意力机制与 SIOU 回归损失函数,试验结果表明,改进模型对棉花顶芽识别的 mAP_0.5 达到 92.8%,相较于原始模型提高了 2 个百分点。李丽等^[30]提出一种基于改进 YOLOv5 的柑橘识别分类算法 YOLOv5-citrus,在颈部网络插入多通道信息融合网络(multi-channel information fusion network, MCIFNet),可对更深层的柑橘状态特征进行提取,识别精确率和平均精度均值达 92.2% 与 91.1%。这些改进方法为散养鸭蛋的识别与定位提供了借鉴。

散养鸭场面积通常较大,机器人在捡蛋过程中,可能会遇到鸭蛋与机器人之间距离较远的情况,且鸭蛋属于小体积目标,中远距离下检测难度较大,容易造成漏检、误检,影响鸭蛋捡拾机器人的捡蛋工作效率与成功率。针对中远距离下鸭蛋的检测与定位问题,本文以 YOLOv8n 网络模型为基础进行改进,在完成中远距离鸭蛋识别的前提下,通过机器视觉与 ROS 中坐标变换(transform, TF)的方式实现目标鸭蛋位置坐标的获取,拟为鸭蛋捡拾机器人的视觉定位提供技术参考。

1 材料与方法

1.1 数据采集与预处理

试验材料为采购的外观大小正常白壳鸭蛋 300 只,将鸭蛋放置在养鸭场常用的蛋窝内,以及与养鸭场垫料地面类似的试验场地,参照散养鸭场地面蛋分布情况,进行散养鸭蛋图像数据采集。图像采集所用相机为 iPhone14,采集图像分辨率为 1 920×1 080。为保证所训练的鸭蛋检测模型的鲁棒性,所采集数据图像中包括不同角度、距离,存在遮挡、部分掩埋等复杂场景的 JPG 格式照片共 900 张,平均每张图片包含约 10 枚鸭蛋样本。

完成数据集图片采集后,使用 Labelimg 工具进行数据集的标注工作,图像中所有鸭蛋的标注框坐标信息以 txt 文本的方式保存。为防止模型训练过程中出现过拟合现象,同时进一步提高鸭蛋检测模型的泛化性能,需要对数据集进行增强处理。本文采用平移、旋转、镜像、放缩、调整亮度、剪切和通过 skimage 工具添加噪声等方法相互随机组合的方式对数据进行增强处理,原始图片和数据增强图片样本如图 1 所示。

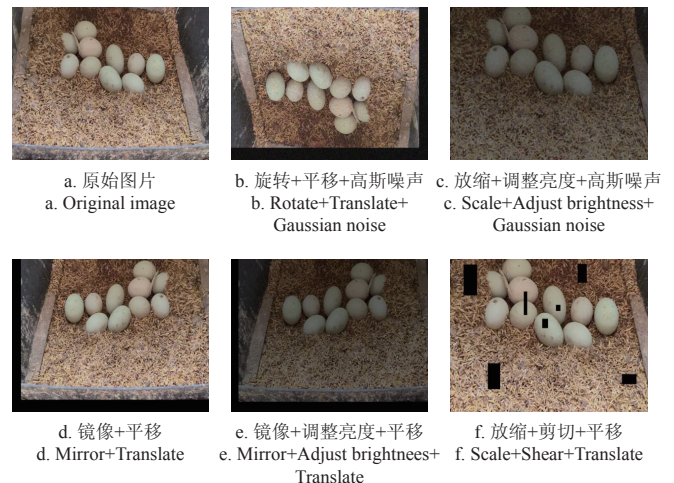


图 1 数据增强样本

Fig.1 Data augmentation samples

1.2 数据集划分

原采集各类图片 900 张,将每张图片通过上述数据增强方法生成 5 张不同图片,得到增强后的图像 4 500 张,原图像与增强后图像合计 5 400 张。把原采集图和增强图像汇总,采用分类随机采样法,将图像按照 60%、20%、20% 的比例划分制作训练集、验证集、测试集 3 种数据集,数量分别为 3 240、1 080、1 080 张,各类样本原图像和数据集图像增强处理情况如表 1 所示。此外,在国家水禽产业体系武汉综合试验站-“农湖 2 号”蛋鸭圈养鸭场,拍摄采集不同距离的散养鸭蛋地面图像 180 张,用于本文改进模型的泛化性能测试。

表 1 数据集分布情况

Table 1 Distribution of datasets

不同图像采集场景 Different image acquisition conditions	距离 Distance/m	增强前图像数量 Number of images before enhancement	增强后图像数量 Number of images after enhancement
0.5 m 近距离,不同光照、羽毛干扰、簇拥遮挡的鸭窝蛋场景	0.5	300	1 500
	1.0	100	500
	2.0	100	500
	3.0	100	500
	4.0	100	500
	5.0	100	500
1.0~6.0 m 中远距离,无遮挡、开阔区域的地面鸭蛋场景	6.0	100	500
	总数量	900	4 500

1.3 试验平台

该研究通过在 PC 端(personal computer)惠普小型计算机工作站完成数据集的处理以及改进模型训练,并将训练所得最优改进模型部署于捡蛋机器人主控板,对地面鸭蛋进行识别与定位,需同时在 PC 端与机器人的主控制器上配置目标鸭蛋图像识别运行环境。其中搭建用于数据集训练和测试的试验平台 PC 端硬件环境为:CPU 处理器 Intel(R) Core(TM) i7-13620H @2.40 GHz, GPU 为 NVIDIA GeForce RTX 4060,内存为 16GB,操作系统为 64 位 Windows 11, CUDA 版本为 11.8, Opencv-Python: 2.4.1, Torch: 4.10.0。机器人底盘采用东莞轮趣科技有限公司生产的阿克曼底盘,主控板采用旷视科技有限公

司生产的旷视 MegBox B1 智能计算单元，搭载了 NVIDIA Tegra X1 芯片，其 CPU 处理器为 Quad-core ARM Cortex-A57，GPU 为：256-core Maxwell，搭载深度相机 Realsense D435i，其底层操作系统为 Ubuntu18.04，及相应软件部署试验环境。

2 目标检测方法

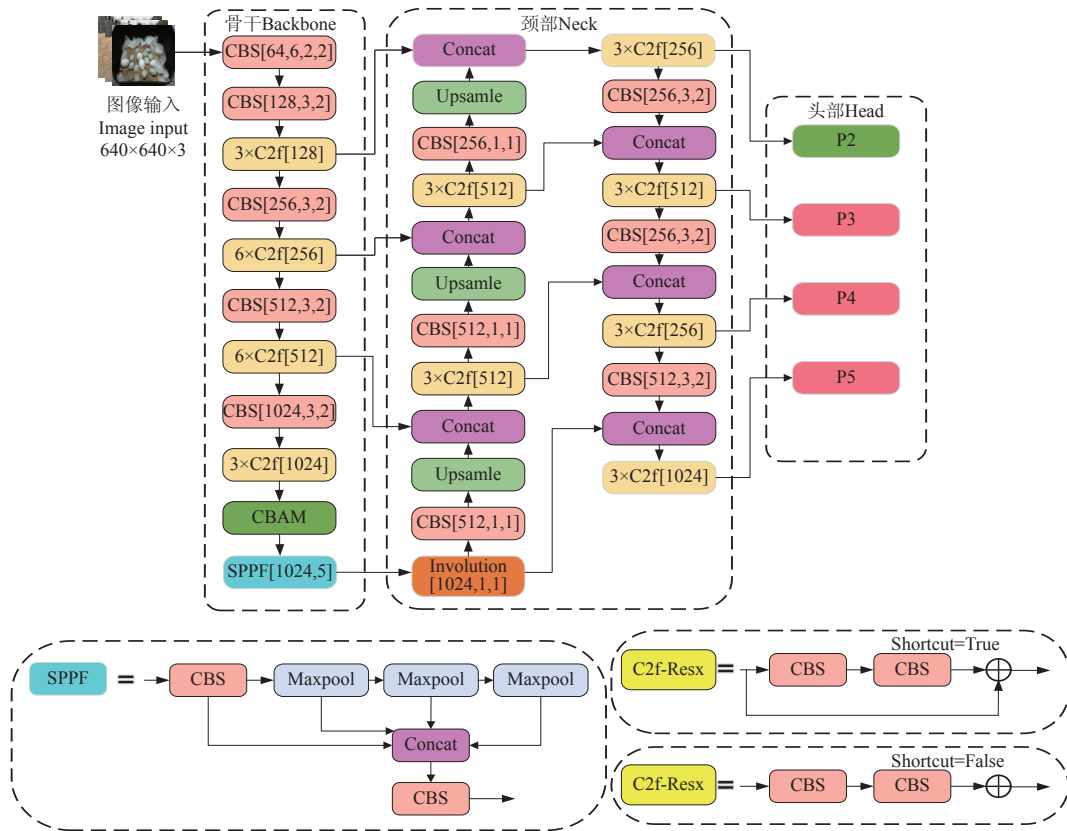
2.1 YOLOv8 检测模型

YOLOv8 网络由输入端、主干网络、颈部网络和头部网络等部分组成。模型训练时，为提高模型网络的鲁棒性并减小 GPU 运算负荷，YOLOv8 在输入端采用 Mosaic 数据增强方法，通过在数据集中抽取 4 张图片进行随机缩放、排布和裁剪并拼接为一张新的图像，以起

到目标数据集增强和提高网络训练的效果。YOLOv8 采用空间金字塔池化模块，可在不同尺度上提取目标特征信息，并选用 C2f 模块替代 C3 模块，进一步降低了模型网络的复杂度。考虑到嵌入式设备的部署问题，本文选用相对较小的 YOLOv8n 网络模型为基础做针对性的模型改进。

2.2 模型改进

因鸭蛋体积较小，中远距离下其在图像中所占的像素面积小，特征难以提取，导致原始的 YOLOv8n 网络对其检测效果不佳，容易出现错检与漏检。针对以上问题，通过在网络模型中添加 CBAM 注意力模块^[31]、Involution 内卷算子^[32]与针对小目标的 P2 检测头对 YOLOv8n 模型进行改进，如图 2 所示。



注：图像输入 Image input 为 3 通道 640×640 像素的图像数据，CBS 为经典模块，C 代表 Conv2d 卷积，B 代表 BatchNorm2d 批量归一化，S 代表 Silu 激活函数；Upsample 表示上采样，Concat 表示张量拼接，Maxpool 表示最大池化，⊕ 表示张量相加，C2f 模块经过 split 后与 concat 组成残差模块，Resx 表示多个残差组件 Res 串联，Shortcut 表示是否采用残差连接；对于 Backbone 中的 C2f，Shortcut = True，对于 Neck 中的 C2f，Shortcut = False，CBAM 为卷积块注意力模块，SPPF 为快速空间金字塔池化模块，Involution 为内卷算子模块，P2~P5 为检测头部对应图像特征金字塔层。

Note: The image input is 3-channel 640×640 pixel image data, CBS is classic module, C is Conv2d convolution, B is BatchNorm2d batch normalization, S is Silu activation function; Upsample is an upsampling module, Concat is tensor concatenation, Maxpool is max-pooling, ⊕ is tensor addition, C2f module forms a residual block in conjunction with concat after being split, Resx is chaining multiple residual components together, Shortcut is whether to use residual connections; In the C2f of Backbone, Shortcut = True, in the C2f of Neck, Shortcut = False. CBAM is the convolutional block attention module, SPPF is the spatial pyramid pooling fast module, Involution is the involution operator module, and P2-P5 are the detection head layers corresponding to the image feature pyramid layer.

图 2 改进网络结构图
Fig.2 Improved network structure diagram

具体改进方法如下：

1) 添加 P2 小目标检测头

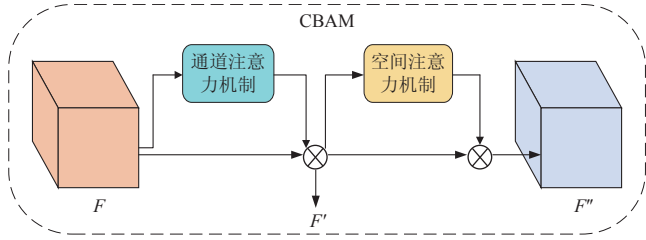
原始的 YOLOv8n 网络包含 P3、P4、P5 有 3 个检测头，可以接收 3 个不同尺寸的特征图，分别对应检测不同尺度的目标。在主干网络中只进行 2 次 CBS 模块操作后，经 C2f 模块与颈部网络融合得到含有底层信息的 P2

检测头，以提高模型对小目标的检测性能。同时，为兼顾模型对中远距离鸭蛋的检测性能，保留 YOLOv8n 原有包含的 3 个目标识别检测头。

2) 添加 CBAM 注意力模块

在 Backbone 主干网络末端的 SPPF 模块前端插入 CBAM 注意力模块，对图像的通道和空间维度进行注意

力加权处理, 使神经网络聚焦于学习图像中的重要特征, 抑制不必要的特征。CBAM 结合了通道注意力机制 (channel attention module) 与空间注意力机制 (spatial attention module), 总体流程如图 3 所示。



注: F 为输入特征图, F' 为中间特征图, 与 F'' 为输出特征图。
Note: F is the input feature map, F' is the intermediate feature map, F'' is the output feature map.

图 3 卷积模块注意力模块结构

Fig.3 Convolutional block attention module (CBAM) structure

在主干网络中加入 CBAM 注意力模块, 可增强卷积神经网络对重要目标特征的提取能力, 抑制网络对无效特征的学习, 提高网络模型对中远距离下的鸭蛋的检测性能。具体计算过程如式 (1) 所示:

$$\begin{cases} F' = M_C(F) \otimes F \\ F'' = M_S(F') \otimes F' \end{cases} \quad (1)$$

式中的 \otimes 表示张量相乘, $M_C(F)$ 表示通道注意力图张量, $M_S(F')$ 表示空间注意力图张量。

通道注意力机制模块对输入的特征图进行通道上的压缩处理, 并同时采用最大池化与平均池化 2 种方法分别对压缩得到的一维矢量进行处理, 得到通道最大池化特征 F_{\max}^C 和通道平均池化特征 F_{avg}^C , 再将 F_{\max}^C 和 F_{avg}^C 传入多层感知器 MLP (multilayer perceptron) 进行处理, 将 MLP 处理所得结果进行张量相加后用 Sigmoid 函数计算权重系数。如式 (2) 所示:

$$\begin{aligned} M_C(F) &= \sigma(\text{MLP}(\text{MaxPool}(F)) + \text{MLP}(\text{AvgPool}(F))) \\ &= \sigma(\text{MLP}(F_{\max}^C) + \text{MLP}(F_{\text{avg}}^C)) \end{aligned} \quad (2)$$

式中 σ 为 Sigmoid 函数, MLP 为多层感知器处理函数, $\text{MaxPool}(F)$ 表示为最大池化运算, $\text{AvgPool}(F)$ 表示为平均池化运算。

经过通道注意力机制模块处理得出的特征图 F' , 会进入空间注意力机制模块进行进一步的处理。在此模块中同样会对特征图 F' 进行最大池化和平均池化处理, 分别得到空间最大池化特征 F_{\max}^S 和空间平均池化特征 F_{avg}^S , 然后在通道维度进行拼接, 并通过 Sigmoid 函数计算其权重。如式 (3) 所示:

$$\begin{aligned} M_S(F') &= \sigma(f^{7 \times 7}(\text{MaxPool}(F'); \text{AvgPool}(F'))) \\ &= \sigma(f^{7 \times 7}(F_{\max}^S; F_{\text{avg}}^S)) \end{aligned} \quad (3)$$

式中 $f^{7 \times 7}$ 表示特征图滤波器大小为 7×7 。

3) 添加 Involution 内卷算子

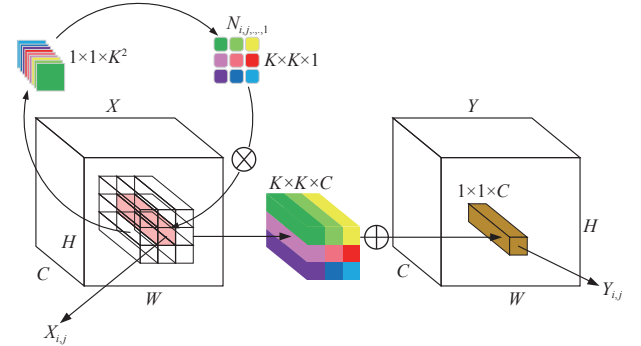
传统卷积具有空间不变性和通道特异性, 在空间维度上卷积核参数共享, 在通道维度上卷积核参数存在特异性, 这种设计使卷积网络能够在不同通道获取更丰富

的特征信息。在实际使用时, 各通道所学习到的特征图往往较为相似, 由此导致大量冗余卷积核出现。同时, 鉴于卷积核数量较多, 为防止计算量过度攀升, 其尺寸会被设计得较小, 这在一定程度上限制了卷积网络的感受野, 不利于小目标与图像背景的区别识别。相对而言, 引入 Involution 内卷具有通道不变性和空间特异性, 一次特征图输入对应的内卷核可以表示为式 (4) 所示:

$$N \in \mathbb{R}^{H \times W \times K \times K \times G} \quad (4)$$

式中 \mathbb{R} 表示内卷核的张量数据空间, N 表示内卷核算子总的张量大小, H 、 W 分别表示输入特征图的高、宽的像素点数, K 表示内卷核的尺寸, 若需提高内卷运算对特征提取的多样性, 可将输入通道分为若干组, 组内通道内卷核共享, G 表示内卷核的分组数。

对于输入特征图中的像素 $X_{i,j} \in \mathbb{R}^C$, 对应的内卷核为 $N_{i,j,\dots,g} \in \mathbb{R}^{K \times K}$, 其中 $g = 1, 2, \dots, G$, 表示组别, 以组别 $g=1$ 为例, 内卷运算过程如图 4 所示。



注: K 为内卷核的空间维度, C 为通道数, $X_{i,j}$ 为输入特征图中的像素, $Y_{i,j}$ 为输出特征图中的像素, $N_{i,j,\dots,g}$ 为内卷核, H 、 W 分别为输入特征图的高、宽的像素点数。

Note: K is the spatial dimension of the involution kernel, C is the number of channels, $X_{i,j}$ is a pixel in the input feature map, and $Y_{i,j}$ is a pixel in the output feature map, $N_{i,j,\dots,g}$ is the kernel volume, and H and W are the pixel points of the height and width of the input feature map, respectively.

图 4 内卷运算过程

Fig.4 Involution operation process

此过程计算式如式 (5) 所示:

$$Y_{i,j,k} = \sum_{(u',v') \in \Delta K} N_{i,j,u'+[K/2],v'+[K/2],k/C} X_{i+u',j+v',k} \quad (5)$$

式中 i, j 为像素坐标, k 为通道编号, $Y_{i,j,k}$ 表示输出特征图中通道 k 上像素 (i, j) 的值, u', v' 表示卷积核中心邻域像素点的偏移量, $N_{i,j,\dots,g}$ 为输入点邻域的像素值, ΔK 表示卷积中心的邻域偏移量集合。

$$\Delta K = \{[-K/2], \dots, [K/2]\} \times \{[-K/2], \dots, [K/2]\} \quad (6)$$

式中 \times 表示为笛卡尔乘法。

由于 Involution 具有与传统卷积互补的特性, 在网络结构中添加 Involution 可以弥补传统卷积通道间无法进行信息交换和感受野受限等缺陷。本文中 Involution 模块的分组 G 取 1, 且输出单像素结果时不需要综合多通道输入, 故其浮点运算量 FLOPs (floating point operations per second) 如式 (7) 所示:

$$\text{FLOPs} = H \times W \times K \times K \times C \quad (7)$$

式中 C 表示通道数，从计算量角度考虑，本文中将 Involution 模块置于 SPPF 与 Neck 部分首个 CBS 模块之间，此处特征图尺寸最小，可以最大程度避免计算量的增加。

2.3 评价指标

本文从平均精度均值 mAP (mean average precision)、置信度损失函数 (obj_loss) 等指标对模型训练及模型的检测效果进行评估。

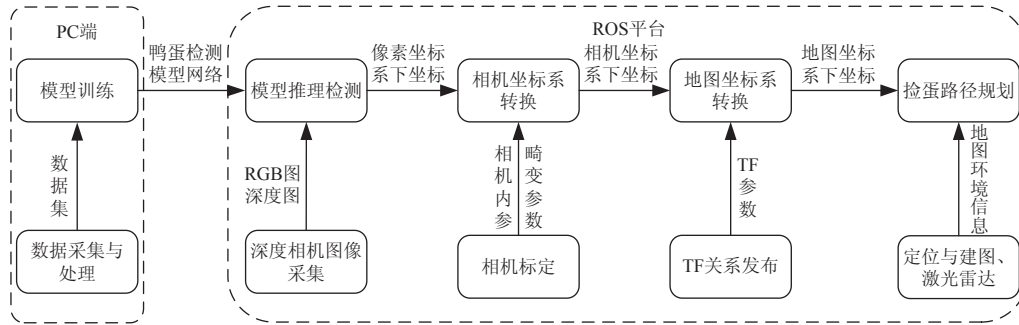
mAP_{0.5} 表示 IoU (intersection over union) 阈值取 0.5 时模型的平均精度均值，对 PR (precision-recall) 曲线进行积分所得，其中的 IoU 表示真实框 A 与预测框 B 的交集与并集之比。 P 与 R 分别表示模型的检测精确率和召回率，mAP、IoU、 P 、 R 的计算式分别如式 (8) ~ (11) 所示。

$$\text{mAP} = \text{AP} = \int_0^1 P \cdot R d(R) \quad (8)$$

$$\text{IoU} = \frac{A \cap B}{A \cup B} \quad (9)$$

$$P = \frac{T_P}{T_N + F_P} \quad (10)$$

$$R = \frac{T_P}{T_P + F_N} \quad (11)$$



注：PC 为电脑端，ROS 为机器人操作系统，RGB 为红绿蓝色彩模式，TF 为坐标变换。

Note: PC is personal computer, ROS is robot operating system, RGB is red, green and blue color mode, TF is transform.

图 5 地面鸭蛋识别定位技术流程图

Fig.5 Flow chart of floor duck egg recognition and positioning technology

3.1 鸭蛋在相机坐标系定位

从像素坐标系到相机坐标系的转换运算过程如式 (12) 所示：

$$z_0 \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & u_0 \\ 0 & f_y & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix} = A \begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix} \quad (12)$$

式中 u 表示像素坐标系下的横坐标， v 表示像素坐标系下的纵坐标， A 表示相机的内参矩阵， f_x 、 f_y 、 u_0 、 v_0 为相机内参，通过相机标定试验获取， z_0 为该点的深度信息，可从深度相机捕获的深度图中直接获取， (x_0, y_0, z_0) 为目标鸭蛋在相机坐标系下的位置坐标。在 MATLAB 平台使用张正友标定法^[35]完成相机标定，图像分辨率为 640×480 ，标定结果如表 2 所示。

3.2 鸭蛋在机器人全局地图坐标系定位

根据相机在机器人底盘上安装的位姿，向 ROS 系统注册相机坐标系，使用 TF 坐标变换库提供的库函数 Transform_Broadcaster()，以话题通信的方式完成相机坐

标系的发布，成功发布相机坐标系之后，TF 坐标系转换节点可通过订阅相关话题，ROS 系统监听从相机坐标系到机器人全局地图坐标系的变换关系，以及目标鸭蛋在相机坐标系下的坐标，并以传感器话题通信的方式，根据相应的坐标系转换矩阵运算，计算出检测到的目标鸭蛋在机器人全局地图坐标系下的位置坐标。

置信度损失函数用于衡量模型预测框的可靠程度，根据预测框与真实标注框的综合交叉比 CIou (complete intersection over union) 作为该预测框的置信度标签，再使用二元交叉熵函数计算置信度损失，置信度损失值越小表示训练出的模型的性能越好。

3 鸭蛋视觉定位方法

地面鸭蛋视觉定位分为相机坐标系定位与地图坐标系定位两个步骤。首先基于深度相机采集目标蛋图像，通过 YOLOv8 的目标检测网络模型可获取得到鸭蛋在图像的像素坐标系下的坐标，通过相机坐标系转换，完成从像素坐标系到相机坐标系的定位；然后，通过 ROS 中的 TF 变换完成从相机坐标系到机器人全局地图坐标系的变换，从而得到鸭蛋在机器人导航地图坐标系下的位置坐标，为机器人捡蛋工作提供导航目标。借鉴移动机器人视觉与标识定位的方法^[33-34]，机器人鸭蛋识别定位的具体技术流程如图 5 所示。

表 2 相机标定试验结果

Table 2 Results of camera calibration experiment

参数类型	Parameter types	标定结果	Calibration results
f_x		594.277	
f_y		596.010	
u_0		319.843	
v_0		245.757	
k_1		0.107	
k_2		-0.173	

注： f_x 为 x 轴等效焦距， f_y 为 y 轴等效焦距， u_0 为图像坐标系原点像素 x 轴坐标， v_0 为图像坐标系原点像素 y 轴坐标， k_1 与 k_2 为径向畸变参数。

Note: f_x is x -axis equivalent focal length, f_y is y -axis equivalent focal length, u_0 is x -axis image coordinate system origin pixel coordinates, v_0 is y -axis image coordinate system origin pixel coordinates, k_1 and k_2 are radial distortion parameters.

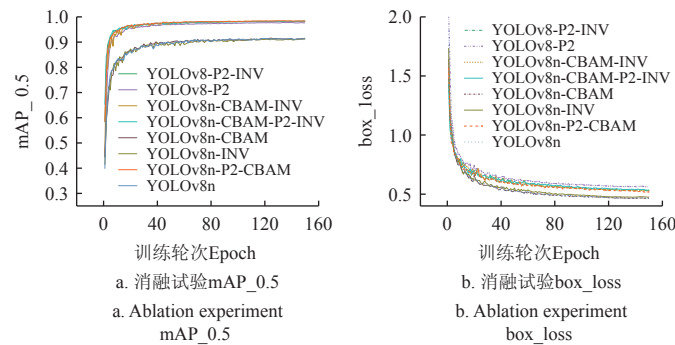
4 结果与分析

4.1 模型训练试验

在使用 YOLOv8n.pt 预训练模型时, 输入图片像素尺寸设为 640×640, 迭代训练轮次设为 150, 学习率设为 0.01, 训练批次设为 16。

4.1.1 模型消融试验

通过搭建的模型训练环境, 设置以上模型训练超参数值, 以 mAP_0.5 及 box_loss 作为模型训练试验的参考指标, 不同改进方法的模型消融训练结果如图 6a~6b。



注: mAP_0.5 为平均精度均值, box_loss 为边界框回归损失。

Note: mAP_0.5 stands for mean Average Precision at IoU threshold of 0.5, box_loss stands for the bounding-box regression loss.

图 6 模型训练消融试验曲线

Fig.6 Curves of model training ablation experiment

表 3 模型消融试验数据分析

Table 3 Model training ablation experiment data analysis

模型种类 Model types	P/%	R/%	mAP_0.5/ %	平均用时/秒 Average time/ms	权重文件 Weight file size/MB	FLOPs/G
YOLOv8n	96.5083.10	91.50	3.6	5.36	6.8	
YOLOv8n-CBAM	96.5083.30	91.70	3.3	5.56	6.9	
YOLOv8n-P2	97.6095.70	97.80	4.0	5.68	10.9	
YOLOv8n-INV	96.4082.60	91.10	3.4	6.00	7.1	
YOLOv8n-P2-CBAM	98.4097.40	98.80	10.3	18.20	41.8	
YOLOv8n-CBAM-INV	98.4097.10	98.70	11.0	18.90	44.3	
YOLOv8n-P2-INV	98.4096.50	98.60	11.5	19.50	48.7	
YOLOv8n-P2-CBAM-INV	98.3096.60	98.60	11.7	20.80	51.4	

注: P 为精确度, R 为召回率, FLOPs 为每秒浮点运算数。

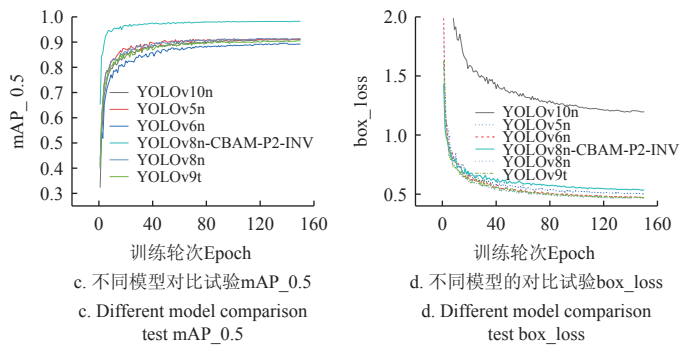
Note: P stands for precision, R stands for recall, and FLOPs stands for floating point operations per second.

由图 6 与表 3 可知, 将 CBAM 注意力模块与 Involution 内卷算子结合使用能为鸭蛋检测模型网络性能带来显著的提升, 并有效抑制 mAP_0.5 和置信度损失数值震荡, 改进模型 YOLOv8n-P2-CBAM-INV 在测试集中的平均精度均值为 98.60%, 召回率为 96.60%, 相较于原 YOLOv8n 模型分别提高了 7.1 和 13.5 个百分点。改进模型的权重文件大小为 20.80 MB, 每秒浮点运算量为 51.4 G, 此模型改进效果优于单独添加 CBAM 模块与 INV 模块的任一改进模型。此外, 将基于此模型改进方法与其他主流模型 YOLOv5n, YOLOv6n, YOLOv9t, YOLOv10n 等模型检测等结果对比如表 4 所示, 最优改进模型 YOLOv8n-P2-CBAM-INV 在对鸭蛋目标识别上, 比 YOLOv9t, YOLOv10n 模型的平均精度均值相对高出 8.04, 7.48 个百分点, 其召回率分别提高 14.90, 14.09 个百分点。YOLOv5n 的权重文件和浮点运算量较小, 但

4.1.2 改进模型与其他先进模型的对比试验

将基于 YOLOv8 改进的模型 v8-P2-CBAM-INV 与 YOLOv5n, YOLOv6n, YOLOv9t, YOLOv10n 等先进开源模型进行对比, 同样以 mAP_0.5 及 box_loss 作为模型训练试验的指标得到模型训练曲线如图 6c~6d。

根据图 6 可知, 模型网络在训练过程中经过 100 训练轮次后, 其 mAP_0.5 曲线及 box_loss 曲线均已收敛趋于稳定, 完整消融试验的模型训练轮次 Epoch 选取 150, 试验所得主要评价指标的数据如表 3 所示。



召回率偏低为 82.90%, 其中 YOLOv6n 模型平均精度均值最低, 仅为 89.60%, 且召回率也比较低。根据该研究在目标鸭蛋测试集上试验数据综合分析, 表 4 中的 4 个主流模型的精确度、召回率和平均精度均值等指标参数, 都与改进模型 YOLOv8n-P2-CBAM-INV 有较大差距。试验结果表明, 选用改进模型 YOLOv8n-P2-CBAM-INV 相比其他模型的效果更好。

表 4 主流模型检测结果对比

Table 4 Model training data analysis

模型种类 Model types	P/%	R/%	mAP_0.5/%	权重 Weight/MB	FLOPs/G
YOLOv5n	96.70	82.90	91.40	4.42	5.8
YOLOv6n	95.50	81.80	89.60	8.15	11.5
YOLOv9t	95.65	81.70	90.56	3.97	6.4
YOLOv10n	94.74	82.51	91.12	5.80	8.2
YOLOv8n-P2-CBAM-INV	98.30	96.60	98.60	20.80	51.4

4.2 不同远距离鸭蛋识别精度试验

数据集图片采集时, 为获取 1.0~6.0 m 的中远距离目标蛋图像, 相机的拍摄角度根据地面目标蛋距离的不同而在垂直地面有较小的前倾角, 相机的拍摄角度的差异, 对检测精度没有影响。

在测试集上对训练所得模型分段距离做测试, 测试集共有 1 080 张图像, 分别考查改进模型对近距离 0.5 m 窝内蛋及中远距离 1.0~6.0 m 窝外鸭蛋的检测性能。该研究将测试集根据不同距离的种类分别进行模型试验, 对添加 P2 检测头、CBAM 注意力机制和 Involution 内卷算子等改进方法进行消融试验, 使用不同改进方法组合训练所得模型, 分别对测试图像进行测试, 观察各模型在测试集上的 mAP_0.5 与 R 指标, 对改进模型性能进行

评估，并将各种改进模型与原始模型进行对比，分析各改进模块对 YOLOv8n 的改进效果，模型测试消融试验结果如表 5 所示。根据其消融试验数据选取具有较好改进效果的模型、原始模型等 5 种模型，分别对鸭场实地采集的分距离测试图像进行进一步检测，检测结果对比如图 7 所示。

表 5 不同距离下模型测试消融试验结果
Table 5 Results of model ablation test under different distances

模型种类 Model types	mAP 0.5/%								R/%								模型大小	单图平均检测用时	FLOPs/ G
	0.5 m	1.0 m	2.0 m	3.0 m	4.0 m	5.0 m	6.0 m	0.5 m	1.0 m	2.0 m	3.0 m	4.0 m	5.0 m	6.0 m	Model size/ MB	Average detection time per image/ms			
YOLOv8n	99.5	99.5	99.5	98.7	93.5	81.8	68.6	99.8	100	99.8	95.5	79.6	74.9	56.4	5.36	8.3	6.8		
YOLOv8n-P2	99.5	99.5	99.5	99.5	99.5	97.9	97.1	99.8	100	99.8	98.7	98.9	95.5	94.6	5.68	10.1	10.9		
YOLOv8n-CBAM	99.5	99.5	99.5	98.7	94.1	80.4	66.6	99.8	100	99.4	95.7	79.9	60.9	55.6	5.56	8.2	6.9		
YOLOv8n-INV	99.5	99.5	99.5	98.9	93.1	78.8	68.0	99.8	100	99.5	96.3	76.7	60.6	57.5	6.00	8.4	7.1		
YOLOv8n-P2-CBAM	99.5	99.5	99.5	99.5	99.5	99.0	98.1	99.9	100	100	99.5	99.8	96.9	94.6	18.20	12.6	41.8		
YOLOv8n-P2-INV	99.5	99.5	99.5	99.5	99.5	98.8	98.1	99.9	100	99.9	99.1	99.5	97.0	93.0	19.50	13.3	48.7		
YOLOv8n-CBAM-INV	99.5	99.5	99.5	99.5	99.5	98.8	97.8	99.8	100	100	99.4	99.5	96.4	94.2	18.90	13.0	44.3		
YOLOv8n-P2-CBAM-INV	99.5	99.5	99.5	99.5	99.5	98.9	98.0	99.8	100	100	99.4	99.6	96.5	95.0	20.80	14.5	51.4		

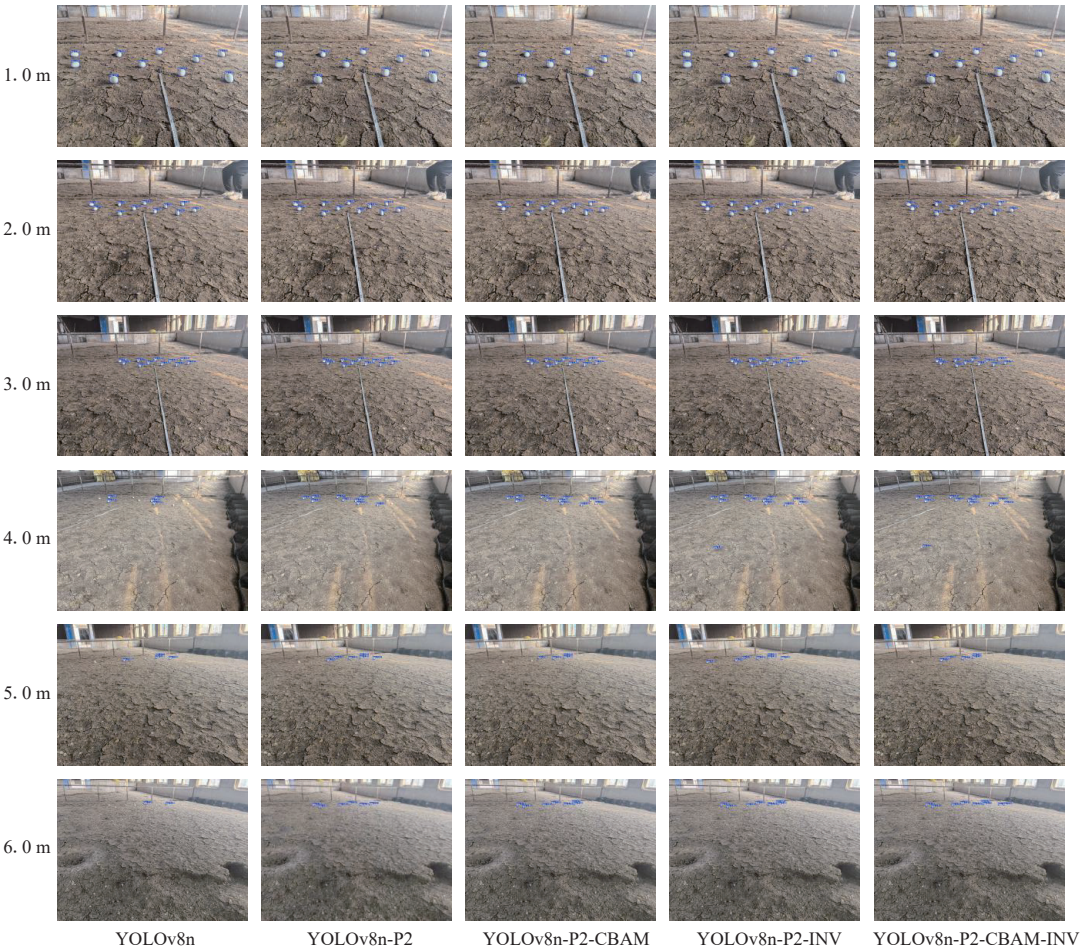


图 7 模型检测效果对比
Fig.7 Comparison of improved model detection effects

由表 5 可知，在 0.5~1.0 m 距离下，YOLOv8n 原模型与各改进模型测试所得平均精度均值和召回率均可达到较好的检测结果。当目标鸭蛋检测距离达到 4.0 m 时，相较于原 YOLOv8n 模型，YOLOv8n-P2-CBAM-INV 在测试集上召回率提升 20.0 个百分点；距离 5.0 m 时，召回率提升 21.6 个百分点；距离 6 m 时，召回率提升 38.6 个百分点，平均精度均值提升 29.4 个百分点。由此结果分析可得，YOLOv8n-P2-CBAM-INV 模型针对中远距离下鸭蛋目标检测有较大优势，漏检率明显降低。改进模型 YOLOv8n-P2-CBAM-INV 的权重文件大小为 20.80 MB，

相较于原始模型 YOLOv8n 增大 15.44 MB，不影响其在 Jetsonx1 边缘计算设备的部署。试验数据显示，YOLOv8n-P2-CBAM-INV 对单张图像平均检测用时为 14.5 ms，可推算出目标图像检测帧率（frames per second, FPS）约达 69 帧/s。

如图 7 所示，在某圈养种鸭养殖基地拍摄采集 1.0~6.0 m 不同距离的散养蛋地面图像数据集，每张图中均有 11 枚鸭蛋，在鸭蛋距离为 1.0、2.0 m 时，所有改进模型均无漏检错检；在鸭蛋距离为 3.0 m 时，原始模型 YOLOv8n 漏检 2 枚鸭蛋，YOLOv8n-P2 漏检 1 枚，无错

检, 其余 3 个复合改进模型均无漏检错检; 在距离为 4.0 m 时, YOLOv8n 漏检 6 枚, YOLOv8n-P2 漏检 2 枚, 其余 3 个复合改进模型均无漏检错检; 在距离为 5.0 m 时, YOLOv8n 漏检 7 枚, YOLOv8n-P2、YOLOv8n-P2-CBAM、YOLOv8n-P2-INV 等模型分别漏检 5、6、7 枚, YOLOv8n-P2-CBAM-INV 漏检 4 枚, 均无错检; 在距离为 6.0 m 时, YOLOv8n 漏检 9 枚, YOLOv8n-P2 漏检 5 枚, YOLOv8n-P2-CBAM、YOLOv8n-P2-INV 各漏检 4 枚, YOLOv8n-P2-CBAM-INV 漏检 2 枚, 均无错检。通过以上检测结果分析, 改进模型 YOLOv8n-P2-CBAM-INV 对目标鸭蛋不同距离下的平均检测效果相对较好。

4.3 鸭场实地拍摄图片的测试试验

为进一步测试改进模型 YOLOv8n-P2-CBAM-INV 的泛化性能, 在湖北省农科院圈养种鸭基地拍摄测试照片, 分鸭窝内、外地面蛋, 以及复杂场景中的地面蛋图像, 进行模型泛化性能测试效果如图 8 所示, 均能较好地检测出目标鸭蛋。

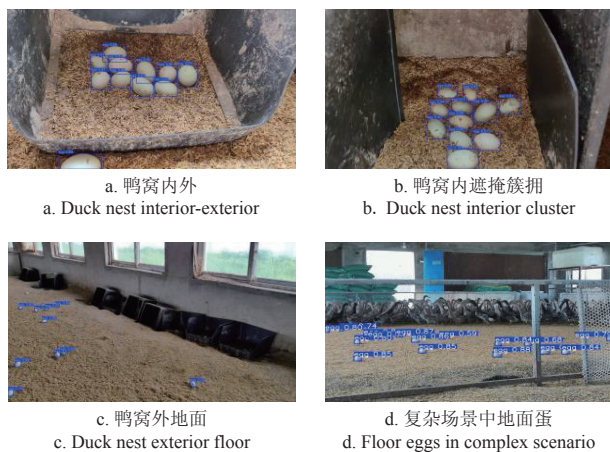


图 8 散养鸭蛋实地场景检测

Fig.8 Field scenario detection of free-caged duck egg

4.4 基于移动平台的鸭蛋视觉定位试验

在室内仿照鸭舍地面蛋场景搭建试验平台, 考虑到嵌入式设备性能有限, 对于高分辨率图像计算耗时过长, 故在实际工作中将相机图像分辨率设为 640×480 。由于分辨率较低, 且鸭蛋体积小, 对于距离较远的鸭蛋特征提取难度大, 故本文仅针对 3.0 m 及以内地面鸭蛋进行试验。

如图 9 所示, 选用 RealSenseD435i 深度相机, 调试安装在捡蛋机器人底座前方支架上。相机的前倾拍摄角为 15° , 是捡蛋机器人上相机坐标系的一个转角参数。通过图像识别及视觉图像坐标系到机器人地图空间坐标系的转换, 可得到机器人视觉对鸭蛋的初定位信息。取 4 个与相机不同距离的鸭蛋集进行识别及坐标距离预测, 距离分别为 0.5、1.0、2.0、3.0 m, 每个距离上分别记录鸭蛋在地图坐标系下的预测坐标与实测坐标, 计算预测点与实测点之间的直线距离, 即定位误差。

表 6 为机器人对鸭蛋的视觉最大定位误差以及每个距离下平均定位误差, 可以看出, 随目标鸭蛋与相机之间的距离增大, 机器人对鸭蛋的平均定位误差整体呈上升趋势, 在距离为 3.0 m 时平均误差为 0.016 9 m, 最大误差为 0.024 4 m。误差来源包括相机自身误差 (深度信

息误差、标定误差等)、相机对于机器人底盘的装配误差、机器人自身定位误差以及计算过程中产生的截断误差等。由以上试验结果可知在 0.5~3.0 m 范围内, 通过本文提出的将图像识别及视觉图像坐标系向机器人地图空间坐标系转换的定位方法, 可实现相机在运动过程中对鸭蛋进行较高精度的实时定位, 3.0 m 距离内测试最大定位误差控制在 0.03 m 以内。

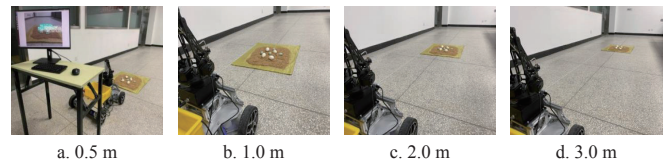


图 9 不同距离鸭蛋识别定位试验

Fig.9 Experiment on identification and localization of duck eggs at different distances

表 6 不同距离下定位误差

Table 6 Positioning error

误差 Error	0.5	1.0	2.0	3.0
最大误差 Maximum error	0.012 7	0.012 7	0.020 6	0.024 4
平均误差 Average error	0.008 4	0.008 4	0.012 6	0.016 9

5 结 论

为实现在中远距离下对鸭蛋的准确识别与定位, 本研究模拟鸭场环境, 采集了多种不同情景以及不同距离的鸭蛋图像, 以 YOLOv8n 模型网络为基础进行改进, 通过添加小目标检测头实现在目标特征更为完整的大尺寸特征图上完成检测, 并引入卷积模块注意力机制模块 (convolutional block attention module, CBAM) 以及 Involution 内卷算子, 增强了特征图不同通道之间的联系, 提高模型网络在通道维度上的感知能力, 从而改善模型网络对中远距离鸭蛋的检测能力。并在此基础上提出一种基于 ROS (robot operating system) 的视觉定位方法, 通过机器视觉与动态坐标变换实现了相机在随机器人运动状态下对目标的高精度实时视觉定位。通过研究试验得出以下结论:

1) 对于中远距离下的鸭蛋检测, 在 YOLOv8n 模型网络中添加小目标检测头、CBAM 注意力机制以及 Involution 内卷算子可在提升模型精度的同时有效抑制 mAP_{0.5} 和置信度损失数值震荡, 加快模型收敛。改进模型 YOLOv8n-P2-CBAM-INV 相较于原始模型, 在测试集上平均精度均值提高了 7.1 个百分点, 改进模型的精确率和召回率分别达 98.30% 和 96.60%。

2) 本研究所用模型改进方法在不影响模型对近距离复杂环境下目标检测性能的前提下, 为中远距离下的目标识别性能带来较大提升, 且鸭蛋与相机距离越远, 提升幅度越明显。在距离为 6.0 m 时, 改进模型 YOLOv8n-P2-CBAM-INV 平均精度均值 mAP_{0.5} 达到 98.0%, 相较于原始模型 YOLOv8n 提高了 29.4 个百分点。

3) 随鸭蛋与相机之间的距离增大, 机器人对鸭蛋的平均定位误差整体呈上升趋势。通过本研究提出的基于 ROS 的视觉定位方法, 在相机与鸭蛋距离 0.5~3.0 m 范围内, 定位误差控制在 0.03 m 以内; 其中, 在距离为

3.0 m 时, 平均误差为 0.016 9 m, 最大误差为 0.024 4 m。结果表明, 该定位方法可实现相机在运动过程中不同距离下对鸭蛋进行较高精度的实时定位。

[参 考 文 献]

- [1] 侯水生, 刘灵芝. 2021 年水禽产业现状、未来发展趋势与建议[J]. 中国畜牧杂志, 2022, 58(3): 227-231, 238.
- HOU Shuisheng, LIU Lingzhi. Current situation, future development trend and suggestion of waterfowl industry in 2021[J]. Chinese Journal of Animal Husbandry, 2022, 58(3): 227-231, 238. (in Chinese with English abstract)
- [2] 肖长峰, 吕文伟, 朱丽慧, 等. 我国蛋鸭养殖的现状、问题及对策分析[J]. 上海畜牧兽医通讯, 2020(4): 56-57.
- [3] 马美湖. 我国禽蛋产业发展现状及需解决的重大科技问题[J]. 华中农业大学学报(社会科学版), 2010(5): 12-18.
- MA Meihu. China's egg industry: Status quo and issues of science and technology[J]. Journal of Huazhong Agricultural University(Social Sciences Edition), 2010(5): 12-18. (in Chinese with English abstract)
- [4] 苑进, 李扬, 刘雪美, 等. 禽蛋自动捡拾系统结构设计及机械手运动规划[J]. 农业工程学报, 2016, 32(8): 48-55.
- YUAN Jin, LI Yang, LIU Xuemei, et al. Structure design of egg auto-picking system and manipulator motion planning[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2016, 32(8): 48-55. (in Chinese with English abstract)
- [5] 刘成良, 贡亮, 苑进, 等. 农业机器人关键技术研究现状与发展趋势[J]. 农业机械学报, 2022, 53(7): 1-22, 55.
- LIU Chengliang, GONG Liang, YUAN Jin, et al. Current status and development trends of agricultural robots[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(7): 1-22, 55. (in Chinese with English abstract)
- [6] 阳康, 张静, 蔡文涛, 等. 基于 ROS 的轻量化机械臂系统设计[J]. 传感器与微系统, 2020, 39(12): 84-87.
- YANG Kang, ZHANG Jing, CAI Wentao, et al. Design of ROS-based lightweight manipulator system[J]. Transducer and Microsystem Technologies, 2020, 39(12): 84-87. (in Chinese with English abstract)
- [7] 鲁敬敬, 秦云川, 刘志中, 等. 机器人操作系统 ROS 安全性研究综述[J]. 软件学报, 2024, 35(2): 1010-1027.
- LU Jingjing, QIN Yunchuan, LIU Zhizhong, et al. Survey on security of robot operating system ROS[J]. Journal of Software, 2024, 35(2): 1010-1027. (in Chinese with English abstract)
- [8] 寿佳鑫, 张志明, 苏永清, 等. 基于 ROS 和激光雷达的室内移动机器人定位和导航系统设计与实现[J]. 机械与电子, 2018, 36(11): 76-80.
- SHOU Jiaxin, ZHANG Zhiming, SHU Yongqing, et al. Design and implementation of indoor positioning and navigation system of mobile robot based on ROS and LiDAR[J]. Machinery & Electronics, 2018, 36(11): 76-80. (in Chinese with English abstract)
- [9] VROEGINDEWEIJ B A, BLAAUW S K, IJSSELMUIDEN J M M, et al. Evaluation of the performance of PoultryBot, an autonomous mobile robotic platform for poultry houses[J]. Biosystems Engineering, 2018, 174: 295-315.
- [10] REN G Q, LIN T, YING Y B, et al. Agricultural robotics research applicable to poultry production: A review[J]. Computers and Electronics in Agriculture, 2020, 169: 105216.
- [11] 吕惠芳. 基于单目视觉的散养鸭蛋识别与定位研究[D]. 福州: 福建农林大学, 2022.
- LV Huifang. Study on the Identification and Localization of Free-range Duck Eggs Based on Monocular Vision[D]. Fuzhou: Fujian Agriculture and Forestry University, 2022. (in Chinese with English abstract)
- [12] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot multibox detector[C]//European Conference on Computer Vision. Berlin, Germany: Springer, 2016: 21-37.
- [13] FU C Y, LIU W, RANGA A, et al. DSSD: Deconvolutional single shot detector[J]. Computer Science, 2017, 1(1): 10-21.
- [14] 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. Seattle, WA, 2016: 779-788.
- [15] REDMON J, FARHADI A. YOLO 9000: better, faster, stronger[C]//Proc of IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE Press, 2017: 6517-6525.
- [16] REDMON J, FARHADI A. Yolov3: An incremental improvement[EB/OL]. (2018-04-08)[2024-05-07]. <https://arxiv.org/abs/1804.02767>
- [17] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2020, 42(2): 318-327.
- [18] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. Yolov4: Optimal speed and accuracy of object detection[EB/OL]. (2020-04-23)[2024-05-07]. <https://arxiv.org/abs/2004.10934>
- [19] TANG S, ZHANG S, FANG Y. HIC-YOLOv5: Improved YOLOv5 for small object detection[C]//International Conference on Robotics and Automation (ICRA), Yokohama, Japan: IEEE, 2024: 10610273.
- [20] 王琳毅, 白静, 李文静, 等. YOLO 系列目标检测算法研究进展[J]. 计算机工程与应用, 2023, 59(14): 15-29.
- WANG Linyi, BAI Jing, LI Wenjing, et al. Research progress of YOLO series target detection algorithms[J]. Computer Engineering and Applications, 2023, 59(14): 15-29. (in Chinese with English abstract)
- [21] 祝志慧, 何昱廷, 李沃霖, 等. 基于改进 YOLOv7 模型的复杂环境下鸭蛋识别定位[J]. 农业工程学报, 2023, 39(11): 274-285.
- ZHU Zhihui, HE Yuting, LI Wolin, et al. Improved YOLOv7 model for duck egg recognition and localization in complex environments[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2023, 39(11): 274-285. (in Chinese with English abstract)
- [22] 赵祚喜, 魏洪飞, 黄渊, 等. 基于改进 YOLOv7 的破壳鸡蛋在线实时检测系统[J]. 农业工程学报, 2023, 39(20): 255-265.
- ZHAO Zuoxi, WEI Hongfei, HUANG Yuan, et al. Online real-time detection system for cracked eggs using improved YOLOv7[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2023, 39(20): 255-265. (in Chinese with English abstract)
- [23] 周宏平, 金寿祥, 周磊, 等. 基于迁移学习与 YOLOv8n 的田间油茶果分类识别[J]. 农业工程学报, 2023, 39(20): 159-166.
- ZHOU Hongping, JIN Shouxian, ZHOU Lei, et al. Classification and recognition of camellia oleifera fruit in the field based on transfer learning and YOLOv8n[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2023, 39(20): 159-166. (in Chinese with English abstract)
- [24] 谭厚森, 马文宏, 田原, 等. 基于改进 YOLOv8n 的香梨目标检测方法[J]. 农业工程学报, 2024, 40(11): 178-185.
- TAN Housen, MA Wenhong, TIAN Yuan, et al. Improved YOLOv8n object detection of fragrant pears[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2024, 40(11): 178-185. (in Chinese with English abstract)
- [25] 张新月, 胡广锐, 李浦航, 等. 基于改进 YOLOv8n 的轻量化红花识别方法[J]. 农业工程学报, 2024, 40(13): 163-170.
- ZHANG Xinyue, HU Guangrui, LI Puhang, et al. Recognizing safflower using improved lightweight YOLOv8n[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2024, 40(13): 163-170. (in Chinese with English abstract)
- [26] 杨大勇, 黄正栋, 郑昌贤, 等. 基于改进 YOLOv8n 的茶叶嫩梢检测方法[J]. 农业工程学报, 2024, 40(12): 165-173.
- YANG Dadong, HUANG Zhengyue, ZHENG Changxian, et al.

- Detecting tea shoots using improved YOLOv8n[J]. *Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE)*, 2024, 40(12): 165-173. (in Chinese with English abstract)
- [27] 李茂, 肖洋轶, 宗望远, 等. 基于改进 YOLOv8 模型的轻量化板栗果实识别方法[J]. *农业工程学报*, 2024, 40(1): 201-209. LI Mao, XIAO Yangyi, ZONG Wangyuan, et al. Detecting chestnuts using improved lightweight YOLOv8[J]. *Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE)*, 2024, 40(1): 201-209. (in Chinese with English abstract)
- [28] 林华浦, 张凯, 李浩, 等. 基于多尺度融合注意力机制的群猪检测方法[J]. *农业工程学报*, 2023, 39(21): 188-195. LIN Huapu, ZHANG Kai, LI Hao, et al. Detecting herd pigs using multi-scale fusion attention mechanism[J]. *Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE)*, 2023, 39(21): 188-195. (in Chinese with English abstract)
- [29] 彭炫, 周建平, 许燕, 等. 改进 YOLOv5 识别复杂环境下棉花顶芽[J]. *农业工程学报*, 2023, 39(16): 191-197. PENG Xuan, ZHOU Jianping, XU Yan, et al. Cotton top bud recognition method based on YOLOv5-CPP in complex environment[J]. *Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE)*, 2023, 39(16): 191-197. (in Chinese with English abstract)
- [30] 李丽, 梁继元, 张云峰, 等. 基于改进 YOLO v5 的复杂环境下柑橘目标精准检测与定位方法[J]. *农业机械学报*, 2024, 55(8): 280-290. LI Li, LIANG Jiyuan, ZHANG Yunfeng, et al. Accurate detection and localization method of citrus targets in complex environments based on improved YOLO v5[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2024, 55(8): 280-290. (in Chinese with English abstract)
- [31] WOO S, PARK J, LEE J Y, et al. CBAM: Convolutional block attention module[C]//*Proceedings of the European Conference on Computer Vision*. Munich, Germany: Springer, 2018: 3-19.
- [32] LI D, HU J, WANG C H, et al. Involution: Inverting the inheritance of convolution for visual recognition[J]. *arXiv*, 2021(1):2103.06255.
- [33] 丁文东, 徐德, 刘希龙, 等. 移动机器人视觉里程计综述[J]. *自动化学报*, 2018, 44(3): 385-400. DING Wendong, XU De, LIU Xilong, et al. Review on visual odometry for mobile Robots[J]. *Acta Automatica Sinica*, 2018, 44(3): 385-400. (in Chinese with English abstract)
- [34] 王志, 朱世强, 李月华, 等. 基于视觉里程计和自然信标融合的移动机器人定位算法[J]. *农业工程学报*, 2017, 33(10): 70-77. WANG Zhi, ZHU Shiqiang, LI Yuehua, et al. Mobile robot localization algorithm by combining visual odometry with natural landmarks[J]. *Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE)*, 2017, 33(10): 70-77. (in Chinese with English abstract)
- [35] ZHANG Z Y. A flexible new technique for camera calibration[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2000, 22(11): 1330-1334.

Recognition and localization of free-range duck eggs based on improved YOLOv8n

GONG Dongjun^{1,2}, LIN Weiguo^{1,3*}, YANG Hao^{1,3}, CHEN Rui^{1,3}, WANG Qiaohua^{1,3}, WANG Shucui^{1,3}

(1. College of Engineering, Huazhong Agricultural University, Wuhan 430070, China; 2. Wuhan Vocational College of Software and Engineering/(Wuhan Open University), Wuhan 430205 China; 3. Key Laboratory of Agricultural Equipment in Mid-lower Yangtze River, Ministry of Agriculture and Rural Affairs, Wuhan 430070, China)

Abstract: The duck egg industry has always dominated an important and indispensable position in the agricultural field. However, the huge amount of egg duck breeding has also brought some challenges to the current industry in recent years. Among them, the low-level automation is ever lagging behind to limit the production efficiency and product quality in duck and egg breeding. Furthermore, the duck eggs can be collected one by one by hand, leading to low work efficiency and high labor intensity. Manual operation cannot fully meet the large-scale production in modern agriculture. At the same time, the frequent entry and exit of breeders into duck farms has inevitably brought some hidden dangers to the health of ducks, such as various pathogens and pollutants from the external environment. This research aims to accurately identify and position the free-range duck eggs at medium and long distances in the key technology of egg collection robots. The automation and intelligence of duck egg collection were then realized using the YOLOv8n network model. A P2 detection head was also added to detect the small target. Accurate detection was achieved in the large-size feature maps that retained the more target features. The CBAM attention module was inserted at the end of the backbone network. The attention was weighted in the channel and spatial dimensions. The neural network was effectively enhanced to focus on the important features, and then suppress the interference of unnecessary features. At the same time, the Involution operator was introduced to significantly improve the detection performance for the duck eggs at medium and long distances. The traditional convolution was then modified in the channel-to-channel information exchange and receptive field. A visual positioning was proposed using ROS. The high-precision and real-time visual positioning of target duck eggs was achieved using machine vision and dynamic TF transformation when the camera was used in the moving state with the robot. The test results show that the weight of the improved model (YOLOv8n-P2-CBAM-INV) was 20.8 MB, and the floating-point operations per second were 51.4 G. The accuracy and recall rates of the improved model were 98.30% and 96.60%, respectively. The mean average precision on the test set increase 7.1 percentage points compared with the original YOLOv8n. The mean average precision of this improved model reached 98% at a distance of 6.0 m, which was increased 29.4 percentage points compared with the original model. The visual positioning test of the mobile robot was carried out to verify the range of 0.5 to 3.0 m. The high-precision and real-time motion positioning was realized at the medium and long-distance duck eggs by the mobile egg-picking robot in the laboratory. The positioning error was controlled within 0.03 m. This improved model can be deployed to the mobile edge platform of image processing. The finding can also provide support to the mobile egg-picking robots, together with the similar autonomous mobile harvesting robots.

Keywords: object detection; visual positioning; duck eggs; recognition; YOLOv8n; robot operating system