

基于改进 YOLOv8 的野生菌分类方法

闵 瑞, 嵇小辅

(江南大学 物联网工程学院, 江苏 无锡 214000)

摘要: 针对野生菌种类繁多, 传统的人工识别准确率低、效率低下的难题, 提出了一种基于 YOLOv8 的野生菌分类新方法; 引入 C2f-MSBlock 卷积, 更好地获取野生菌的多尺度特征, 减少了计算成本; 引入多头自注意力机制模块, 避免模型陷入局部最优从而提升分类模型的精度; 针对传统检测头难以有效捕捉野生菌在菌伞和斑点等细节上差异的问题, 提出了一种不对称双检测头的方案; 针对边界框回归中出现的损失问题, 引入了 Inner-CIoU 损失函数, 使得模型可以灵活调整边界框的尺度大小, 减少了边界框损失; 实验结果表明, 在野生菌分类任务中, 相较于野生菌领域最优的 3DRe-OLO 算法, 改进后方法使 $mAP@0.5$ 提升了 0.3%, 精确率提升了 1.2%, 召回率提升了 0.3%, 验证了改进的有效性。

关键词: 野生菌; 目标检测; YOLOv8; 多头自注意力机制; Ciou; C2f

Wild Fungi Classification Method Based on Improved YOLOv8

MIN Rui, JI Xiaofu

(School of Internet of Things, JiangNan University, Wuxi 214000, China)

Abstract: A novel method for wild fungi classification based on YOLOv8 is proposed to address the challenges posed by the diverse range of wild fungi species and the low accuracy and efficiency of traditional manual identification methods. The approach introduces the C2f-MSBlock convolution to better extract the multi-scale features of wild fungi while reducing computational costs. Additionally, a multi-head self-attention mechanism is incorporated to prevent the model from converging to local optima, thereby improving classification accuracy. Regarding the difficulty of traditional detection heads in effectively capturing the differences in details such as fimbriae and spots of wild fungi, an asymmetric dual detection head scheme is proposed. To address the issue of loss in bounding box regression, an Inner-CIoU loss function is introduced to enable the model to flexibly adjust the size of the bounding box, reducing the bounding box loss. Experimental results demonstrate that, compared to the best 3DRe-YOLO algorithm in the field of wild fungi, the proposed method increases the $mAP@0.5$ by 0.3%, the accuracy by 1.2%, and the recall by 0.3%, validating the effectiveness of the improved method.

Keywords: wild fungi; object detection; YOLOv8; multi-head self-attention; Ciou; C2f

0 引言

野生菌具有口味独特、蛋白量高、脂肪量低、富含维生素和矿物质等特点, 深受广大消费者的喜爱和推崇。野生菌种类繁多、形态各异, 一些有毒菌和无毒菌在形态、大小上都非常相似, 这使得人们难以肉眼分辨野生菌是否可食。因此, 近年来我国误食野生菌而中毒的患者日益增多^[1]。高精度、高效率、简单方便的野生菌分类方法对于提高野生菌的食用安全性非常重要。传统的识别方法主要有仪器化学检测法^[2], 真菌分类学鉴定方法^[3]等, 这些方法操作复杂、检测成本高, 难以实

现在线实时检测。

目前主流目标检测方法主要可以分成两类: 传统窗口扫描算法和深度学习目标检测算法。传统窗口扫描算法利用窗口来提取图像中的像素和信息进而检测图像中是否包含目标对象, 其中主要有基于梯度方向直方图的特征描述算法^[4]、结合积分图技术的特征描述算法^[5]、尺度不变特征变换算法^[6]等。基于窗口扫描的目标检测算法需要在图像的各个位置 and 不同尺度上滑动窗口, 并对图像进行多次评估, 会导致窗口重叠而产生冗余, 从而降低精确率。深度学习目标检测算法主要分为单阶段目标检测方法和多阶段目标检测方法。单阶段目标检测

收稿日期: 2024-11-08; 修回日期: 2024-12-17。

作者简介: 闵 瑞(2000-), 男, 硕士研究生。

嵇小辅(1979-), 男, 博士, 教授。

引用格式: 闵 瑞, 嵇小辅. 基于改进 YOLOv8 的野生菌分类方法[J]. 计算机测量与控制, 2025, 33(11): 284-291.

方法只需要提取一次特征就能完成检测任务, 其典型网络为 YOLO^[7] 系列检测算法。多阶段目标检测方法是基于 CNN 的目标检测算法, 其思想是先提取预选框, 缩放至统一尺度进行 CNN 特征提取, 然后在全连接层进行分类, 其典型网络有 R-CNN^[8], Fast R-CNN, Faster R-CNN 等。传统窗口扫描算法和基于深度学习的目标检测算法相比, 它难以捕捉上下文信息, 缺乏适应性和泛化能力。

基于深度学习的目标检测在果蔬类、农作物、野生菌等食用物检测领域^[9-14] 都有广泛应用。文献 [15] 使用 Xception 与 ResNet50 模型对野生菌进行识别, 有效提高了分类的准确性和泛化性; 文献 [16] 提出了结合多级区域选择和跨层特征融合的 ViT 架构算法, 解决了现有野生菌分类算法在背景噪声较大时识别率较低的难题; 文献 [17] 提出了野生菌检测的改进 YOLO 算法, 降低了模型复杂度, 提高了模型训练速度; 文献 [18] 采用随机翻转、添加高斯噪声等方法来扩充毒蕈数据集; 文献 [19] 提出 MSFResNet 网络模型来抓取深层网络的特征, 提升了模型的识别效果。上述模型的网络参数量相较于基础网络明显增加, 导致计算资源消耗较大。

本文针对野生菌种类繁多、形态各异且难以区分的难题, 提出一种基于改进 YOLOv8 模型的野生菌分类方法, 提升了模型的正例个数, 保证了模型的泛化性并且降低了模型复杂度。本文研究对于提高野生菌识别的准确性、促进野生菌学研究和生态环境保护具有重要意义。具体而言, 改进后的模型具有以下优点:

- 1) 引入 C2f-MSBlock 来代替传统的卷积结构, 采用多尺度特征提取的方法, 在维持模型精度的同时, 有效地降低了计算复杂度;
- 2) 通过引入多头自注意力机制, 实现对序列在不同方面的信息捕捉, 使得模型可以获得全局特征从而提升模型精确度;
- 3) 提出包含两个检测分支的一种不对称检测头, 有效避免较大的感受野整合更多的背景信息从而忽略小目标的特征, 有助于模型更好地适应目标大小的变化, 进而提升检测的精确性和效率;
- 4) 引入了 Inner-CIoU 损失函数, 避免了边界框回归中出现的 IoU 损失的问题, 提高了样本的回归效果。

1 YOLOv8 网络模型

YOLOv8^[20] 是 Ultralytics 公司于 2023 年提出的 YOLO 系列基本模型之一, 是一种基于 PyTorch 框架的单阶段目标检测算法, 具有模型复杂度低、检测精确率较高等优点。YOLOv8 算法模型分为输入网络、主干网络、Neck 网络以及输出端四部分。输入网络通过将

输入的图像自适应缩放, 然后通过 Mosaic 数据增强、灰度填充等方式对输入的数据集进行数据增强; 主干网络为 CSPDarkNet, 采用 C2f 卷积结构以及具有金字塔结构的 SPPF 来提取图片特征; Neck 网络保证了信息的高效传递, 从而增强了目标检测的特征表达能力; 在输出端中, YOLOv8 采用了不依赖于预设锚框的目标检测方法, 实现了正负样本匹配和损失计算以及分类和检测头的分离。

2 野生菌的改进 YOLOv8 目标检测算法

2.1 改进算法的整体结构

针对野生菌目标小, 数量多的特点, 本文在 YOLOv8 算法基础上对检测头、卷积函数、损失函数等进行改进, 并添加多头自注意力机制。整体的网络结构如图 1 所示。

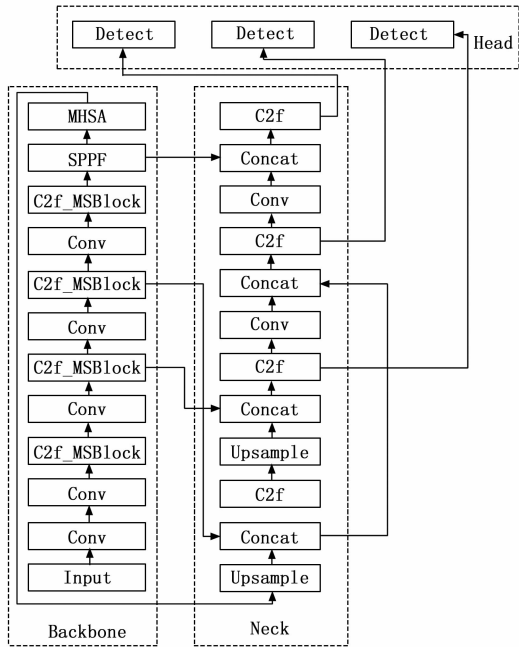


图 1 YOLOv8 改进后结构图

YOLOv8 的卷积层虽然可以通过特征融合来提高网络的表达能力, 但是传统卷积的计算复杂度较大, 因此本文引入 C2f-MSBlock 卷积, 通过多尺度特征提取, 在保证精度的同时降低计算复杂度。为了进一步提升模型对野生菌特征的捕获能力, 本文引入 MHSA 多头自注意力机制, 充分利用其整合全局特征的优势来获得更全面的特征表示。同时, 提出了一种具有两个不同分支的检测头结构, 分别专注于大尺寸目标和小尺寸目标, 确保模型能够准确捕捉野生菌的细节特征, 提升检测精度。此外, 引入了 Inner-CIoU 损失函数, 避免了边界框回归中的 IoU 损失, 通过微调损失函数中的比例因子来调整边界框大小, 可以更好地适应目标, 进一步提升模型的精度。

2.2 C2f-MSBlock 卷积

2.2.1 C2f-MSBlock 卷积块设计

C2f 是神经网络中常用的卷积模块，旨在增强表达特征的能力，提高模型推理的精度。但是传统的 C2f 卷积通过多个分支进行特征融合，会导致较高的计算成本。此外，在卷积的计算过程中卷积核的大小不随目标的变化而变化，导致推理模型忽略部分特征，从而降低模型精度。

由于野生菌多以菌落形式出现，传统的 C2f 卷积难以处理该目标。因此，本文采用 C2f-MSBlock^[21]将输入特征分成几个分支，每个分支处理不同尺度的信息特征，通过多尺度特征的提取和融合来提高检测模型的精度。在网络结构上，C2f-MSBlock 在浅层使用小卷积核来获取高分辨率特征，在深层采用大卷积核来获取大尺度特征，这种结构通过逐步融合相邻分支之间的特征，较好地提高模型精确度。C2f-MSBlock 的结构图如图 2 所示。

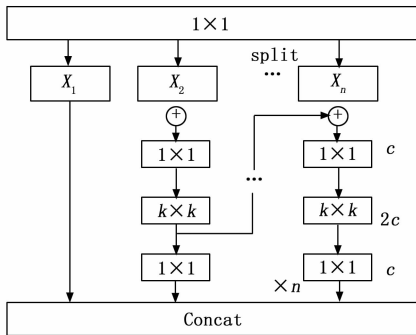


图 2 C2f-MSBlock 卷积

设为输入特征。经过 1×1 的卷积转换后， X 的通道维数增加到 $n \times c$ ，然后将输入特征 X 分成 n 个不同的组 $\{X_i\}$ ，其中。为了降低计算成本，我们选择 $n = 3$ 。除了 X_1 之外，每一组都要经过一个反向瓶颈层，表示为 $IB_{k \times k}()$ ，其中 c 表示通道数， k 表示内核大小。 Y_i 的数学表示可以描述如下：

$$Y_i = \begin{cases} X_i & i = 1 \\ IB_{k \times k}(Y_{i-1} + X_i) & i > 1 \end{cases} \quad (1)$$

根据式 (1)，反向瓶颈层充当跨级连接并保留来自前一层的信息。完成了所有部分的连接之后，应用 1×1 卷积进行交互，当网络深入时，这个 1×1 卷积也用于调整通道数。

2.2.2 异构卷积核选择方法

在目标检测中，编码器是指特征提取网络，它用于在图像上进行卷积操作，以提取图像特征。为了使用不同阶段的卷积来获取多尺度特征，本文采用异构卷积核选择方法来获取更丰富的特征信息。在浅层使用内核较小的卷积，在深层使用内核较大的卷积，从浅层到深层

逐步增加中间阶段的内核大小。这种策略将不同大小的卷积核放在最合适的位置，最大程度地提高了卷积的利用效率，有效保持特征分辨率的一致性。

如图 3 所示，异构卷积核的主干由 4 个阶段组成，每个阶段之后紧接一个步幅为 2 的 3×3 卷积，用于下采样。

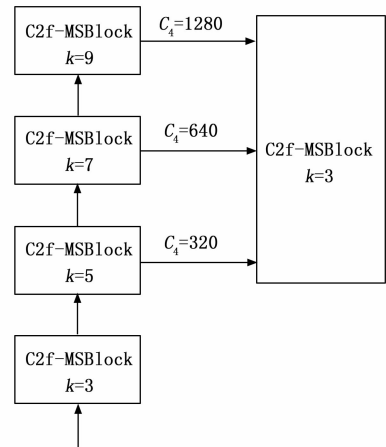


图 3 异构卷积核选择方法

为了融合不同阶段的多尺度特征，将 PAFPN^[22] 作为颈部，在编码器上构建特征金字塔。在金字塔结构中，从浅层获取的特征用于检测小物体，从深层获取的特征用于检测大型对象。根据不同的 k 选择不同的内核大小，从而扩大深层的感受野的同时不会对浅层的感受野造成影响。

2.3 MHSA 自注意力模块

注意力机制可以让模型在处理目标时重点聚焦于最相关的信息，选择性关注输入序列中的重点部分。传统注意力机制只使用单一注意力头，无法从多个角度获取输入特征，可能会使得模型陷入局部最优解，降低模型精度。

多头自注意力 (MHSA, multi-head self-attention)^[23] 是一种改进的注意力机制，这种机制允许模型关注重要特征，增强特征权值，从而提供更全面的特征表示。在训练过程中，MHSA 通过并行计算，独立更新每个注意力头的梯度，缓解了梯度消失以及梯度爆炸的问题，提高了训练稳定性。

MHSA 通过将输入数据 X 映射为查询 Q 、键 K 和值 V ，同时生成多个不同的头，每个头关注数据的不同子空间，实现对序列在不同尺度的信息捕捉，MHSA 模块如图 4 所示。

在 MHSA 中，输入为 $x = [h, w, d]$ ，经过三次线性变换，生成大小均为 $x = [h, w, d]$ 的查询矩阵、键矩阵和值矩阵。 q 和 k 通过矩阵乘法得到内容矩阵，用于表示输入序列中的特征向量。 R_h 和 R_w 分别表示垂

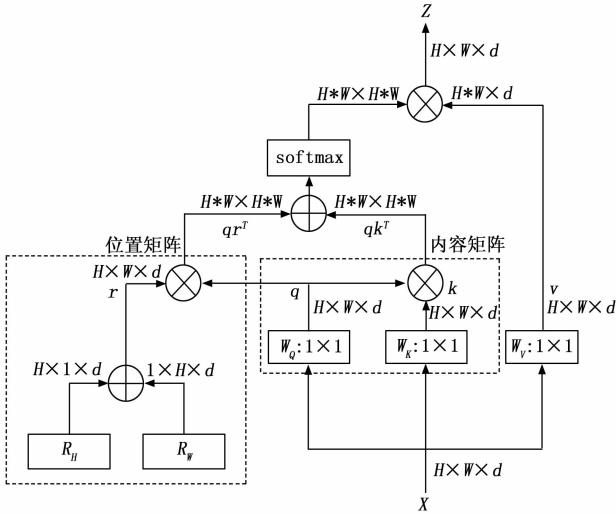


图 4 MHSA 结构图

直和水平方向上的相对信息, 将两者相加后得到相对位置矩阵, 再将 q 和相对位置编码矩阵通过矩阵乘法后得到位置矩阵, 位置矩阵可以表示相对位置信息。最后, 将内容矩阵和位置矩阵相加得到整个 MHSA, 表示为:

$$\text{softmax}(qr^T + qk^T)v \quad (2)$$

MHSA 可以同时考虑输入中特征的所有位置信息, 从而捕获图像中长距离的依赖关系。此外, 野生菌常以菌群和菌落的形式出现, 存在部分遮挡的现象。MHSA 可以使得目标受到遮挡时, 依然可以保持较好的性能, 增加了模型的鲁棒性, 大幅度提升模型精确度。

2.4 不对称双检测头

检测头负责将从编码器中提取到的特征生成边界框, 再根据不同的边界框输出一个类别概率分布和置信度评分, 完成对目标的检测和分类, 是检测模型的重要组成部分。

传统的检测头虽然设计简洁, 能够进行端对端的训练, 但是由于其较大的感受野, 会导致在检测小目标时目标图像的边界框不够精确, 降低模型的精确度。本文提出一种不对称双检测头来替代传统的检测头, 在此结构中通过非对称的双头结构来提升检测性能, 如图 5 所示。

这种设计包含两个不同的检测分支, 一个用于检测大尺寸目标, 另一个则专注于小尺寸目标。这样的非对称布局有以下优点:

1) 由于野生菌类别较多, 使用不对称检测头能够更注重物体的类别, 根据每个特征层的特点分配不同的计算资源, 从而减少冗余计算提升整体计算速度;

2) 在复杂场景下, 模型可以使用不同大小和数量的卷积核来提取特征, 提高对不同尺度目标的检测

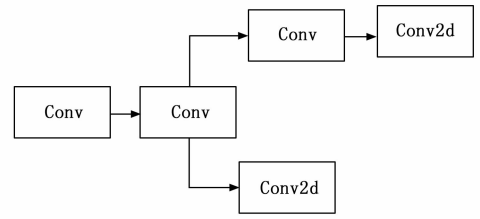


图 5 不对称双检测结构图

精度;

3) 针对不同的任务类型, 网络被设计成独立的路径, 每条路径分别负责特定的任务, 使模型能够并行地处理不同特征层的数据, 从而提升模型精度。

2.5 Inner-CIoU 损失函数

损失函数在模型训练中具有至关重要的作用。在模型训练过程中, 损失函数能够衡量预测结果与真实结果之间的差异。通过最小化损失函数来调整模型的参数, 能够使得模型可以更好地拟合数据。

IoU Loss 是一种用于衡量两个边界框之间相似度的损失函数, 该损失函数可以精确描述预测框和真实框的匹配程度, 以 IoU 作为回归损失函数其定义如下:

$$\text{IoU} = \left| \frac{B \cap B^{\text{gt}}}{B \cup B^{\text{gt}}} \right| \quad (3)$$

其中: B 代表预测框, B^{gt} 代表真实框。IoU 的值在 0 到 1 之间, 值越大表示两个框的重合程度越高, 预测效果越好。在定义了 IoU 之后, 其损失函数可以被定义为:

$$L_{\text{IoU}} = 1 - \text{IoU} \quad (4)$$

IoU 损失函数能直观反映预测框与真实框之间的差异, 但是 IoU 损失函数也有缺点, 例如, 对于在边界框重叠较小的情况会导致训练缓慢; 当 IoU 值接近于 0 时, 会由于梯度较小导致参数更新不明显。为了克服以上缺点, 衍生出了基于 IoU 的各类变种损失函数, 其中 CIoU 损失函数通过综合考虑重叠区域、中心点距离和长宽比, 有助于模型学习到更加鲁棒的特征表示, 从而增强模型的准确性的同时提升模型的泛化能力, CIoU 可以表示为:

$$\nu = \frac{4}{\pi} \left(\arctan \frac{\omega^{\text{gt}}}{h^{\text{gt}}} - \arctan \frac{\omega}{h} \right)^2 \quad (5)$$

$$\alpha = \frac{\nu}{(1 - \text{IoU}) + \nu} \quad (6)$$

$$L_{\text{CIoU}} = 1 - \text{IoU} + \frac{\rho^2(b, b^{\text{gt}})}{c^2} + \alpha \omega \quad (7)$$

其中: ρ 为欧氏距离, b 和 b^{gt} 分别为 B 和 B^{gt} 的中心点, c 为最小边界框的对角线, ν 表示长宽比的一致性, α 表示正权衡参数, ω^{gt} 和 h^{gt} 表示目标框的宽度和高度, ω 和 h 表示预测框的宽度和高度。

针对边界框回归中出现的 IoU 损失的问题, 本文

采用了 Inner-CIoU^[24] 损失的解决方案。该损失函数使用辅助边界框来加速回归过程，而不引入新的损失项，通过在训练过程中使用不同尺度的辅助边界框计算损失，来提高样本的回归效果；控制比例因子 r 来调整辅助边界框的尺度大小，以适应不同的数据集和检测器。本文采用的 Inner-CIoU 的损失函数可以由以下得到：

$$b_l^g = x_c^g - \frac{\omega^g \cdot r}{2} \tag{8}$$

$$b_t^g = y_c^g - \frac{h^g \cdot r}{2} \tag{9}$$

$$b_r^g = x_c^g + \frac{\omega^g \cdot r}{2} \tag{10}$$

$$b_b^g = y_c^g + \frac{h^g \cdot r}{2} \tag{11}$$

$$b_l = x_c - \frac{\omega \cdot r}{2} \tag{12}$$

$$b_t = y_c - \frac{h \cdot r}{2} \tag{13}$$

$$b_r = x_c + \frac{\omega \cdot r}{2} \tag{14}$$

$$b_b = y_c + \frac{h \cdot r}{2} \tag{15}$$

$$\text{inter} = [\min(b_r^g, b_r) - \max(b_l^g, b_l)] \times [\min(b_b^g, b_b) - \max(b_t^g, b_t)] \tag{16}$$

$$\text{union} = \omega^g h^g r^2 + \omega h r^2 - \text{inter} \tag{17}$$

$$\text{IoU}^{\text{inner}} = \frac{\text{inter}}{\text{union}} \tag{18}$$

$$L_{\text{inner-CIoU}} = L_{\text{CIoU}} + \text{IoU} - \text{IoU}^{\text{inner}} \tag{19}$$

Inner-CIoU 损失函数通过引入辅助边界框来评估边界框中的重叠部分。由于小目标的相对位置误差更容易被放大，而 Inner-CIoU 可以减少位置偏差的敏感性，因此对于野生菌而言具有较好的适应性。

3 实验结果与分析

3.1 数据集与实验设备

本文使用的数据集为 Roboflow 提供的 Mushroom 数据集。数据集中包括巴西蘑菇，鸡枞菌，猴头菇，茶树菇等 21 类，其中训练集 7281 张、验证集和测试集各 838 张图片。

实验使用的计算机设备为：12th Gen Intel (R) Core (TM) i5-12400F 2.50 GHz；内存为 16G；显卡为 NVIDIA GeForce RTX 3060 ti；python 版本为：3.9；系统为 Windows 11（64 位）；cuda 版本为 12.2；torch 版本为 2.2.1。且在相同环境下训练 300 轮，优化器为 SGD。

3.2 数据增强及超参数

本文使用了 3 种数据增强，分别为 Mixup^[25]，Flipud 以及 Copy-paste^[26]，相关的超参数设置如表 1 所示，其余超参数均为 YOLOv8 默认设置。

表 1 数据增强相关超参数设置

参数名称	超参数设置
Mixup	0.25
Flipud	0.2
Copy-paste	0.3

3.3 模型评价指标

评价模型性能的主要指标通常包括精确度（P，precision）、召回率（R，recall）、平均精度（AP，average precision）、平均精确度均值（mAP，mean average precision）、每秒浮点运算次数（FLOPs）以及权重文件大小（Weights）：

$$Precision = \frac{TP}{TP + FP} \tag{20}$$

$$Recall = \frac{TP}{TP + FN} \tag{21}$$

$$AP = \int_0^1 P(r) dr \tag{22}$$

$$mAP = \frac{\sum_{i=1}^c AP_i}{C} \tag{23}$$

精确度是指模型在检测出的目标中真正属于目标类别的比例。TP（True Positive）表示模型预测正确的正样本个数，FP（False Positive）表示模型将负样本预测为正样本的个数，FN（False Negative）表示模型将正样本预测为负样本的个数。 $mAP@0.5$ 表示在 0.5 的 IoU 阈值下各个类别的平均精度， $mAP@0.5$ 越高表示模型的表现越好。每秒浮点运算次数即计算量，用于衡量模型训练时的计算复杂度。权重文件用于加载预训练模型、保存训练结果等，权重文件的大小能反映训练的数据量以及模型的复杂程度。

3.4 对比实验

为了验证改进模型在检测野生菌方面的有效性，将其与 4 个主流模型进行对比，对比结果如表 2 所示。

表 2 与其他模型的对比

模型	$mAP@0.5$	Precision	Recall
APHS-YOLO ^[27]	0.550	0.695	0.513
BiFormer ^[28]	0.567	0.694	0.523
3DRe-YOLO ^[29]	0.608	0.711	0.527
文献 ^[30]	0.527	0.708	0.512
Ours	0.611	0.723	0.530

通过对比实验结果可以看出，本文提出的改进模型在关键性能指标 $mAP@0.5$ 、Precision 和 Recall 上均表现出色，取得了最优的性能表现。具体来说，与 3DRe-YOLO 模型相比，改进后的模型的 $mAP@0.5$ 提高了 0.3%，显示出更高的整体检测精度；Precision 指标则提升了 1.2%，进一步证明了模型在准确识别目标方面

的优势;同时,Recall 指标提高了 0.3%,表明模型在捕获目标能力上的显著改进。这些实验结果充分表明了改进后的模型在性能上的显著优势,能够满足实际应用的需求。

3.5 消融实验

为了验证不同因素对模型改进的影响,将 YOLOv8n 作为基准模型,在其基础上分别添加改进,并对每个改进功能进行分析。在相同的训练环境下得到表 3,单独加入不对称检测头之后,模型的 $mAP@0.5$ 维持不变、权重文件略微增加,但是计算量降低了 0.3 G,这说明检测头可以维持精度的同时大幅减少计算量。单独加入 MHSA 之后模型的 $mAP@0.5$ 大幅度提升,提升了 2.5%,同时计算量和原模型相比增大了 0.2 G,权重文件增加了 0.37 M。单独加入 C2f-MSBlock 后模型的 $mAP@0.5$ 略微提升,但与此同时模型的计算量也减少了 0.2 G,权重文件增加 0.38 M。当同时加入不对称检测头和 C2f-MSBlock 之后,模型的 $mAP@0.5$ 提升了 1.6%,但是计算量大幅降低,和原模型相比,减少了 0.6 G,证明了不对称检测头和 C2f-MSBlock 可以保证模型精确度的同时减少模型计算量。当同时加入不对称检测头和 MHSA 时 $mAP@0.5$ 提升了 3.3%,模型计算量减少了 0.3 G。加入 MHSA 和 C2f-MSBlock 时, $mAP@0.5$ 提升了 3.6%。当加入不对称双检测头, MHSA 和 C2f-MSBlock 时模型的 $mAP@0.5$ 达到最大,较基准模型分别提升 4.2%和 3.8%,此时虽然计算量不是最低,但是比较基准模型降低了 0.3 G,权重文件增加了 0.38 M,说明了这 3 个改进可以在提高精度同时,可以降低计算量。因此本文改进的模型更适合野生菌的识别。

YOLOv8n 模型和改进后模型的平均精度增长对比如图 6 所示。由图 6 可知,改进后的模型在平均精度方面有显著提升,因此改进后的模型更适合野生菌的检测。

3.6 不同比例因子的损失函数

在 Inner-CIoU 损失函数中,比例因子 r 的大小

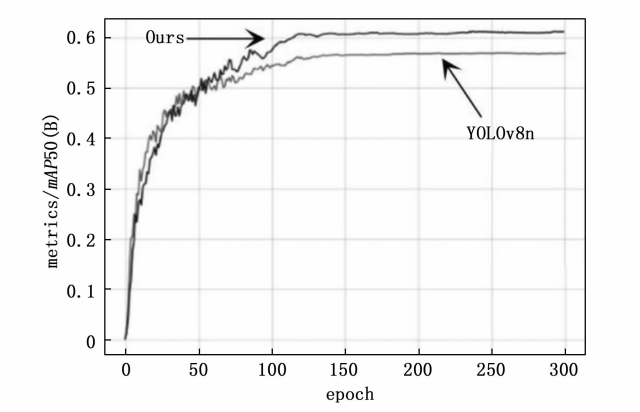


图 6 消融实验对比图

能够调整辅助边界框的大小,为了选取损失函数中最佳比例因子,将不同比例因子对模型的影响进行对比实验,结果如表 4 所示。由表 4 可以得出 r 为 0.7 时模型的精度最佳。当 $r=1$ 时,损失函数即为 Ciou,当 $r<1$ 时实际框大于辅助框, $r>1$ 时辅助框大于实际框,在 $r=0.7$ 时效果最好,由此可以看出较小的辅助框更适合野生菌的检测。

表 4 比例因子的影响

r	$mAP@0.5$
0.2	0.584
0.7	0.611
1	0.601
1.2	0.596

3.7 检测效果可视化

为了验证模型的实际效果,将 3DRe-YOLO 的模型和改进后的模型进行了对比。可以从图 7 中看出,改进后的模型能够更精确地框选目标对象,避免了原模型中边界框偏移或未覆盖完整目标的情况;在多目标重叠区域,改进模型能够正确地区分出相邻目标;在较为复杂环境中也能够精准地识别小型对象;对于一幅图中有多个野生菌类别的情况,改进后的模型减少了误检,检测的精确度也有了显著提升,证明了改进后模型的有效性。

表 3 消融实验结果

YOLOv8n	不对称双检测头	MHSA	C2f-MSBlock	$mAP@0.5$	Precision	Recall	FLOPs	Weights
✓				0.569	0.696	0.492	8.2G	6.00M
✓	✓			0.569	0.679	0.497	7.9G	6.10M
✓		✓		0.594	0.666	0.506	8.4G	6.37M
✓			✓	0.587	0.626	0.547	8.0G	6.38M
✓	✓		✓	0.585	0.664	0.520	7.6G	6.37M
✓	✓	✓		0.602	0.687	0.515	7.9G	6.48M
✓		✓	✓	0.605	0.703	0.521	8.2G	6.38M
✓	✓	✓	✓	0.611	0.723	0.530	7.9G	6.38M

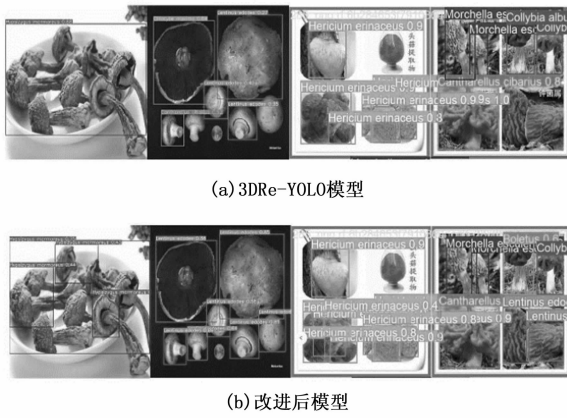


图 7 改进前后检测效果对比

为了更直观地展示改进后模型的效果提升,采用黑白热力图来直观展示检测的效果,其中白色表示置信度较高的部分,黑色表示置信度较低的部分。HiResCAM通过不平均梯度,保证了解释的精确性,可以将其用于提高图像质量和清晰度,有利于图像的分析。

通过热力图中白色区域的覆盖情况可以看出改进后的模型对于目标区域的识别更全面。此外,改进后的模型能够更好地排除干扰信息。在 3DRe-YOLO 的模型中,目标区域的白色部分出现断裂以及明显的白色噪点,而改进后的模型背景区域保持纯黑。证明了改进后的模型不仅可以精准地检测出目标物体,且对于不同环境也有一定的适应能力,可以较好地排除错误信息。

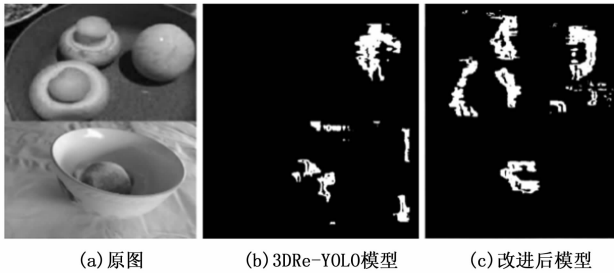


图 8 改进前后热力图对比

4 结束语

为了解决野生菌识别时精确度低,成本高昂等问题,本文提出了一种改进 YOLOv8 模型。通过将主干结构的 C2f 卷积替换为 C2f-MSBlock 卷积来获取多尺度特征;加入了 MHSA 注意力机制来更好地获取全局特征,避免了模型陷入局部最优;为了减少冗余计算、获取细节特征,提出了不对称双检测头;最后将 CIoU 损失函数改为 Inner-CIoU 损失函数,使得模型可以根据野生菌的特点灵活调整边界框大小。由实验可知,本文采用的改进模型在可以改善模型精度的同时减少模型的计算量,有助于将模型部署于实际场合。

参考文献:

- [1] 李林静,李高阳,谢秋涛. 毒蘑菇毒素的分类与识别研究进展 [J]. 中国食品卫生杂志, 2013, 25 (4): 383 - 387.
- [2] GRANDSTAFF D, TERRY JR D. Rare earth element composition of paleogene vertebrate fossils from toadstool geologic park, nebraska, USA [J]. Applied Geochemistry, 2009, 24 (4): 733 - 745.
- [3] 桂明英,何容,郭永红,等. 基于形态特征和 ITS 序列对新疆芦苇根蘑菇的分类鉴定 [J]. 食用菌, 2014, 36 (4): 14 - 16.
- [4] BHATTARAI B, SUBEDI R, GAIRE R R, et al. Histogram of oriented gradients meet deep learning: a novel multi-task deep network for 2D surgical image semantic segmentation [J]. Medical Image Analysis, 2023, 85: 102747.
- [5] FELZENSZWALB P F, GIRSHICK R B, MCALLESTER D, et al. Object detection with discriminatively trained part-based models [J]. IEEE transactions on pattern analysis and machine intelligence, 2009, 32 (9): 1627 - 1645.
- [6] LOWE D G. Distinctive image features from scale-invariant keypoints [J]. International Journal of Computer Vision, 2004, 60 (2): 91 - 110.
- [7] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection [C] // Las Vegas: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 779 - 788.
- [8] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation [C] // Ohio: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2014: 580 - 587.
- [9] FENG X, LI D D, WANG W J, et al. Image recognition of wheat leaf diseases based on lightweight convolutional neural network and transfer learning [J]. Life, 2021, 13 (10): 2125.
- [10] KISS N, CZÜNI L. Mushroom image classification with CNNs: a case-study of different learning strategies [C] // Beijing: 2021 12th International Symposium on Image and Signal Processing and Analysis (ISPA), 2021: 165 - 170.
- [11] LIU B, DING Z, TIAN L, et al. Grape leaf disease identification using improved deep convolutional neural networks [J]. Frontiers in Plant Science, 2020, 11: 1082.
- [12] WANG J, LIU X H. Pathological recognition of apple leaves based on deeply separable convolution [J]. Computer Systems & Applications, 2020, 29 (11): 190 - 195.
- [13] ZHANG L, MA R, ZHANG Y X. Application of improved LeNet-5 model in apple image recognition [J]. Computer Engineering and Design, 2018, 39 (11): 3570 - 3575.
- [14] ZHOU Z, HE D, ZHANG H, et al. Non-destructive de-

- tection of moldy core in apple fruit based on deep belief network [J]. Food Sci, 2017, 38 (14): 297–303.
- [15] 沈若兰, 黄英来, 温 馨, 等. 基于 Xception 与 ResNet50 模型的蘑菇分类方法 [J]. 黑河学院学报, 2020, 11 (7): 4.
- [16] 李俊仪, 李向阳, 龙朝勋, 等. 基于多级区域选择与跨层特征融合的野生菌分类 [J]. 计算机工程, 2024, 50 (9): 179–188.
- [17] 张泽冰, 张冬妍, 姜蕴伟, 等. 改进 YOLOv4 的野生菌视觉检测方法 [J]. 计算机工程与应用, 2023, 59 (20): 228–236.
- [18] 樊帅昌. 基于深度残差网络与迁移学习的毒蕈图像识别 [D]. 杭州: 浙江农林大学, 2019.
- [19] 杨 阳. 基于深度学习的野生菌识别方法研究 [D]. 兰州: 兰州交通大学, 2023.
- [20] LOU H, DUAN X, GUO J, et al. DC-YOLOv8: small-size object detection algorithm based on camera sensor [J]. Electronics, 2023, 12 (10): 2323.
- [21] CHEN Y, YUAN X, WU R, et al. YOLO-MS: rethinking multi-scale representation learning for real-time object detection [J/OL]. <https://arxiv.org/abs/2308.05480>.
- [22] LIN T Y, DOLLÁR P, GIRSHICK R, et al. Feature pyramid networks for object detection [C] // Hawaii: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017: 2117–2125.
- [23] SRINIVAS A, LIN T Y, PARMAR N, et al. Bottleneck transformers for visual recognition [C] // Kuala Lumpur: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021: 16519–16529.
- [24] ZHANG H, XU C, ZHANG S. Inner-iou: more effective intersection over union loss with auxiliary bounding box [J/OL]. <https://arxiv.org/abs/2311.02877>
- [25] ZHANG H, CISSE M, DAUPHIN Y N, et al. mixup: Beyond empirical risk minimization [J]. <https://arxiv.org/html/2405.14458v1>.
- [26] GHIASI G, CUI Y, SRINIVAS A, et al. Simple copy-paste is a strong data augmentation method for instance segmentation [C] // Kuala Lumpur: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021: 2918–2928.
- [27] LIU R M, SU W H. APhS-YOLO: a lightweight model for real-time detection and classification of stropharia rugoso-annulata [J]. Foods, 2024, 13 (11): 1710.
- [28] ZHU L, WANG X, KE Z, et al. BiFormer: vision transformer with Bi-Level routing attention [J]. 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2023: 10323–10333.
- [29] JIANG R, CAI H, YU T. Bisporus mushroom 3D reconstruction for digital twin of smart factory [C] // Singapore, Springer, 2024: 47–62.
- [30] FENG W, LIU M, SUN Y, et al. The use of a blueberry ripeness detection model in dense occlusion scenarios based on the improved YOLOv9 [J]. Agronomy, 2024, 14 (8): 1860.
- 834–837.
- [16] WEISS A J. Direct position determination of narrowband radio frequency transmitters [J]. IEEE Signal Processing Letters, 2004, 11 (5): 513–516.
- [17] YIN J X, et al. Direct localization of multiple stationary narrowband sources based on angle and doppler [J]. IEEE Communications Letters, 2017, 21 (12): 2630–2633.
- [18] GARCIA NIL, et al. Direct localization for massive MIMO [J]. IEEE Transactions on Signal Processing, 2017, 65 (10): 2475–2487.
- [19] FAN W, LIU S, LI C, et al. Fast direct localization for millimeter wave MIMO systems via deep ADMM unfolding [J]. in IEEE Wireless Communications Letters, 2023, 12 (4): 748–752.
- [20] 李廷蔚. 基于 GAA 模型的分布源直接定位方法 [D]. 成都: 电子科技大学, 2017.
- [21] 吴癸周, 郭福成, 张 敏. 信号直接定位技术综述 [J]. 雷达学报, 2020, 9 (6): 998–1013.
- [22] 吴癸周. 单个运动接收站对辐射源信号的直接定位方法研究 [D]. 长沙: 国防科技大学, 2019.