

一种基于 FR-YOLO 注意力增强的轻量化中药饮片识别检测方法

于喆源¹, 王晓霞^{1*}, 李晓芳¹, 樊东琴²

1. 张掖市质量检验检测研究院, 甘肃 张掖 734000

2. 张掖市广泰药材有限责任公司, 甘肃 张掖 734000

摘要: **目的** 改进 YOLOv8 模型, 为中药生产、调剂与教学等实际应用场景提供一种高精度、高效率的中药饮片自动化检测方案。**方法** 以 YOLOv8 为基线模型, 提出轻量化残差深度注意力瓶颈模块 (residual depthwise-attention Bottleneck, RDA-Bottleneck), 替换 C2f 中的 Bottleneck 模块, 以降低冗余计算并增强通道特征表达; 提出频域感知空间注意力模块 (frequency-aware spatial attention, FASA), 替换 YOLOv8 的 Backbone 网络与 Neck 网络中的 Conv 模块, 以增强检测网络在多尺度特征学习过程中的判别能力; 构建包含板蓝根、甘草等 10 种中药饮片数据集, 共 8 281 张图像, 用于评价改进模型的检测性能。**结果** 相较于 YOLOv8 模型, 改进后的模型参数量降低了 39.5%, 浮点运算量 (floating point operations, FLOPs) 降低了 34.1%, 单阈值平均精度 mAP50 提升了 0.2%, 多阈值平均精度 mAP50-95 提升了 0.2%。**结论** 改进后的模型在复杂背景与形态多变的中药饮片检测任务中展现出更高的检测精度与推理效率, 为中药饮片自动化检测提供了一种高效的新方法。

关键词: 中药饮片检测; YOLOv8; 轻量化目标检测; 注意力机制; 残差深度注意瓶颈模块; 频域感知空间注意力模块

中图分类号: TP18; R283

文献标志码: A

文章编号: 0253-2670(2026)10-3902-10

DOI: 10.7501/j.issn.0253-2670.2026.10.019

A lightweight recognition and detection method for Chinese herbal slices based on FR-YOLO attention enhancement

YU Zheyuan¹, WANG Xiaoxia¹, LI Xiaofang¹, FAN Dongqin²

1. Zhangye Institute of Quality Inspection and Testing, Zhangye 734000, China

2. Zhangye Guangtai Medicinal Materials Co., Ltd., Zhangye 734000, China

Abstract: Objective To improve the YOLOv8 model and provide a high-accuracy and high-efficiency automated detection solution for Chinese herbal slices in practical scenarios such as production, dispensing, and teaching. **Methods** Taking YOLOv8 as the baseline model, a lightweight residual depthwise-attention Bottleneck (RDA-Bottleneck) is proposed to replace the Bottleneck blocks in C2f, reduce redundant computations and enhance the expression of channel features. To strengthen the discriminative capability of the detector during multi-scale feature learning, a frequency-aware spatial attention (FASA) module is introduced to replace the Conv blocks in the Backbone and Neck of YOLOv8. To evaluate the performance of the improved model, a dedicated dataset containing 10 categories of Chinese herbal slices [e.g., Banlangen (*Isatidis Radix*) and Gancao (*Glycyrrhizae Radix et Rhizoma*)] is constructed, comprising 8 281 images. **Results** Compared with YOLOv8, the proposed model reduces parameters by 39.5% and floating point operations (FLOPs) by 34.1%, while improving single-threshold average accuracy mAP50 by 0.2% and multi-threshold average accuracy mAP50-95 by 0.2%. **Conclusion** The improved model achieves higher detection accuracy and inference efficiency for Chinese herbal slice detection under complex backgrounds and diverse appearances, providing an effective method for automatic detection of Chinese herbal slices.

Key words: Chinese herbal slices detection; YOLOv8; lightweight object detection; attention mechanism; residual depthwise-attention Bottleneck; frequency-aware spatial attention

中药饮片是中医药产业链中连接种植、加工与临床应用的关键环节, 其质量直接影响中药处方的

疗效与医药安全性^[1]。然而, 传统的中药饮片鉴别与分拣工作主要依赖人工经验, 不仅效率低下, 而

收稿日期: 2026-01-02

基金项目: 甘肃省药品监督管理局青年科技创新项目 (2024GSMPA034)

作者简介: 于喆源, 工程师, 从事食品药品检验检测及品质评价。E-mail: 305117349@qq.com

*通信作者: 王晓霞, 工程师, 从事食品药品检验检测及品质评价。E-mail: 429416455@qq.com

且受主观因素影响较大,容易产生误判^[2]。随着中药产业智能化与标准化建设的推进,中药饮片的自动化、高精度识别检测已成为保障药品质量与实现现代化管理的重要环节^[3]。

近年来,深度学习技术广泛应用于计算机视觉领域,在目标检测任务中展现出卓越的特征学习与泛化能力^[4]。在中药饮片检测识别的研究工作中,许多学者已经开展了大量探索性工作。例如,刘晶等^[5]基于 EfficientDet 提出结合注意力机制的中药饮片识别方法,在易混淆饮片(川贝母、平贝母)检测中取得 99.34%的识别准确率。李云阳等^[6]提出改进 YOLOv5 的中药饮片缺陷检测算法,通过引入 FasterNet 结构、SimAM 注意力与 SimOTA 标签匹配机制,提高了缺陷识别能力,其 mAP 达到了 87.53%。谷瑞等^[7]构建了 BIM-Res2Net-ICA 中药饮片识别模型,通过双向多尺度特征融合与改进坐标注意力机制以增强特征表达能力,并结合焦点损失(focal loss)提升分类鲁棒性,在 20 类饮片识别任务中取得 97.48%的准确率。在对 YOLOv8 的改进上,孙兴等^[8]提出了一种改进的 YOLOv8 中药饮片识别方法,通过在 YOLOv8n 中引入 GhostC2f、DySnakeC2f、SimSPPF 以及坐标注意力机制,以降低模型参数量并增强对细粒度及小目标特征的表达能力,从而实现中药饮片的高效自动识别。该工作主要通过引入已有轻量化或高效模块(如 GhostBottleneck、DSConv 及 SimSPPF)对 YOLOv8 进行组合式优化,其重点在于借鉴现有成熟模块提升模型性能与效率。

尽管现有研究在提升中药饮片检测精度方面取得一定进展,但多数方法仍依赖复杂网络结构,因此模型体量大、计算开销高,难以满足边缘设备部署与实时检测需求^[8]。此外,中药饮片具有形态多变、纹理差异细微及类别相似度高特征,在复杂背景和遮挡干扰下易出现小目标漏检与类别混淆等问题。为解决上述问题,本研究提出一种基于注意力增强的轻量化中药饮片目标检测模型 FR-YOLO,主要工作内容如下:(1)提出轻量化残差深度注意瓶颈模块(residual depthwise-attention Bottleneck, RDA-Bottleneck),替换 YOLOv8^[9]中 C2f 结构的 Bottleneck。RDA-Bottleneck 采用倒残差结构,以深度可分离卷积取代传统卷积、输出端使用自主改进的注意力机制(residual squeeze-excitation, RSE),在显著降低参数量与浮点运算量(floating

point operations, FLOPs)的同时提升特征提取能力。(2)提出频域感知空间注意力模块(frequency-aware spatial attention, FASA),替换 YOLOv8 的 Backbone 网络与颈部(Neck)网络中的卷积(Conv)模块。FASA 结合空间注意力(spatial attention, SA)和频域空间混合模块(frequency domain spatial mixing module, FSM),采用轻量化双分支建模机制,卷积分支关注语义表征,池化分支关注结构与边缘细节,在降低参数量和 FLOPs 的同时强化语义判别能力与结构细节,从而提升多尺度特征学习能力。

1 设备与材料

1.1 实验环境

本研究在 Linux(Ubuntu)系统环境下完成实验,硬件平台采用 NVIDIA RTX 4090D GPU(显存 24 GB)单卡运行。软件配置包括 Python 3.8.10、PyTorch 2.0.0(CUDA 11.8 支持)框架^[10]。模型训练阶段将输入图像统一缩放至 640×640,训练轮数(epochs)设为 50,批量大小(batch size)设为 8。优化器采用随机梯度下降(stochastic gradient descent, SGD)优化策略,学习率、权重衰减等超参数沿用 YOLOv8n 的默认配置,以保证对比实验的公平性。

1.2 数据集

为了验证所提出 FR-YOLO 模型在中药饮片检测任务中的有效性,本研究构建了 1 个包含 10 类常见中药饮片的专用数据集,共 8 281 张图像。数据来源在张掖市广泰药材有限责任公司及其种植基地拍摄,以及在百度、Kaggle 平台等网络上采集,其中 3 000 张为实地拍摄获得,5 281 张来源于百度、Kaggle 等公开网络平台。涵盖白术、甘草、枸杞等多种临床应用频度较高的饮片类别。数据集中饮片在形态、尺寸、背景光照等方面具有显著差异,更贴近真实场景中存在的检测难点,有利于全面评价模型的泛化能力。所有数据均使用 LabelImg 工具按照 YOLO 标注规范进行边界框标注,并按照 8:2 的比例划分为训练集与测试集,确保类别分布一致性与实验公平性。数据集详细统计信息如表 1 所示,样本展示见图 1。

为减少原始图像分辨率差异对数据标注与模型训练的影响,本研究首先将数据集图像统一整理为 300×300 大小。考虑到不同中药饮片类别在形态特征和个体可分辨性方面存在差异,采用差异化标注策略。其中,对于草果和草豆蔻 2 类饮片,由于其个体边界相对清晰,因此对图像中每个可辨识目标分别进行实

表 1 数据集详细信息

Table 1 Detailed dataset information			
类别	数量	类别	数量
白术	824	茯苓	758
板蓝根	822	甘草	955
草果	831	枸杞	916
草豆蔻	622	荔枝核	816
冬虫夏草	908	人参	829

例级标注；而对于板蓝根、冬虫夏草等类别，由于其常呈现堆叠、交错和边界模糊等特征，难以准确区分单个个体，因此将图像中的饮片区域作为整体目标进行标注。标注的数据集类别整体分布情况如图2所示。

在模型训练阶段，所有图像统一调整为 640×640 输入尺寸，并采用 YOLOv8 默认的预处理与数据增强策略，包括缩放与填充、随机水平翻转、Mosaic 增强、MixUp 增强以及 HSV 颜色扰动等操



图 1 数据集样本的展示

Fig. 1 Display of dataset samples

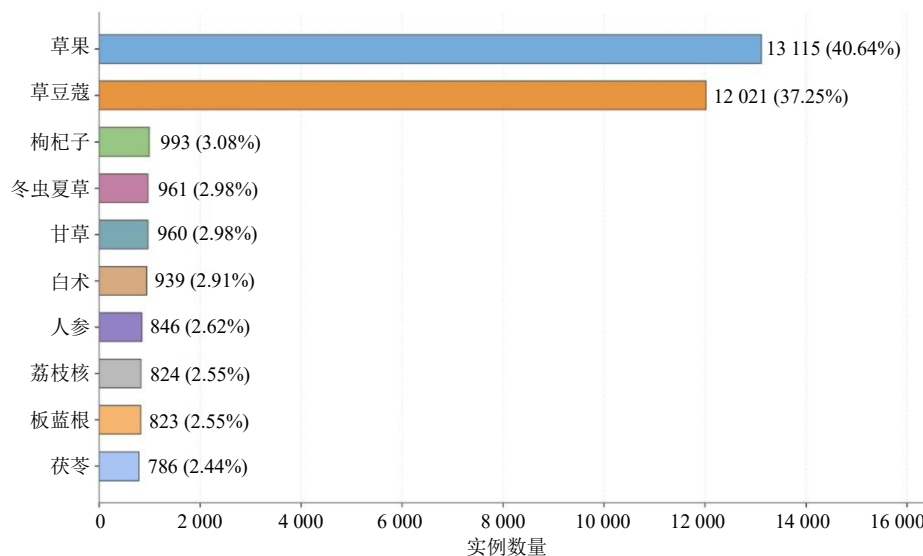


图 2 类别分布情况

Fig. 2 Category distribution

作，以提高模型对尺度变化、复杂背景和光照变化的适应能力。

2 方法

2.1 对比模型与评估指标

为全面验证本研究的 FR-YOLO 模型性能，选

择单阶段 YOLO 系列中近期发布的轻量化模型作为对比，主要包括 YOLOv8n^[9]、YOLOv9t^[11]、YOLOv10n^[12]、YOLOv11n^[13]、YOLOv12n^[14]、YOLOv13n^[15]。此外，还选择了 FCOS^[16]、TOOD^[17]、RetinaNet^[18]3 个非 YOLO 系列的检测模型。上述模

型均基于同一数据集、相同训练策略与硬件环境进行实验，以确保对比结果的公平性与可靠性。

针对检测精度评估，采用目标检测主流指标精度 (precision, P)、召回率 (recall, R)、单阈值平均精度 (mAP50) 和多阈值平均精度 (mAP50-95) 进行性能评估。 P 越高表示误报越少，其数学定义如下：

$$P = TP / (TP + FP)$$

TP 为正确检测的目标数量，FP 为错误检测 (误检) 数量

R 越高表示目标被检测出的比例越高，其数学定义如下：

$$R = TP / (TP + FN)$$

FN 为漏检目标数量。

单 IoU 阈值下的平均精度 (mAP50) 的数学定义如下：

$$mAP50 = \frac{1}{C} \sum_{i=1}^C AP_i^{IoU=0.5}$$

C 为类别数， AP_i 表示第 i 类的平均精度，IoU=0.5 表示预测框与真实框交并比阈值为 0.5

mAP50-95 的数学定义如下：

$$mAP50-95 = \frac{1}{10C} \sum_{t=0.5}^{0.95} \sum_{i=1}^C AP_i^{IoU=t}$$

IoU 阈值从 0.5~0.95 以 0.05 为步长共 10 个区间，更能反映模型在定位精度与分类能力上的综合

性能。

为衡量模型轻量化程度，采用参数量 (params) 和浮点计算量 (floating point operations, FLOPs) 指标进行评估。参数量衡量模型规模，其值越小说明模型越便于实际部署；FLOPs 衡量计算复杂度，其值越低则说明模型的推理速度越高。通过精度指标与轻量化指标的综合对比，可全面验证 FR-YOLO 在检测性能与推理效率之间的协同优化效果，可体现其在实际应用中的优势。

2.2 FR-YOLO 模型

FR-YOLO 整体架构沿用 YOLOv8 网络设计，由骨干特征提取网络 Backbone、多尺度特征融合网络 Neck 和检测头 3 部分构成。在骨干网络 Backbone 中，FR-YOLO 仅保留 YOLOv8 的首层 Conv 模块，其余所有 Conv 模块均替换为所提出的 FASA，以增强浅层至深层特征建模能力并降低卷积计算开销。原 YOLOv8 中 C2f 模块内部所使用的标准 Bottleneck 替换为本研究提出的轻量化 RDA-Bottleneck，替换后的 C2f 模块在 FR-YOLO 中记为 RDAC2f，以提升深层语义表征能力。检测头部分保持不变，继续沿用 YOLOv8 的多尺度检测结构。模型整体结构如图 3 所示。FR-YOLO 首先由骨干网络提取中药饮片图像的多阶段层级特征，其中 FASA 通过双分支建模机制分别聚焦语义与边缘结

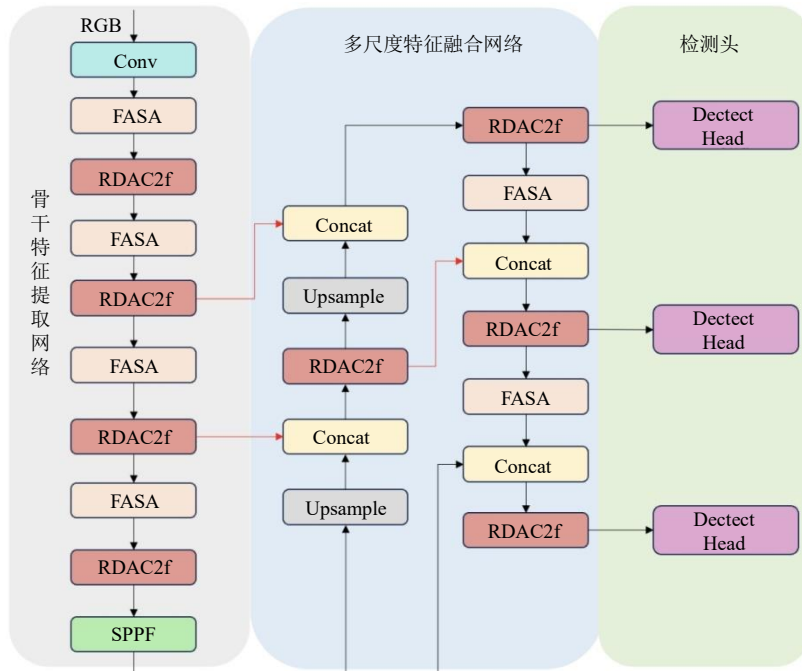


图 3 FR-YOLO 网络结构

Fig. 3 FR-YOLO network structure

构信息,并结合 SA 空间注意力增强显著区域特征,同时利用 FSM 模块对高频细节进行增强,提升目标与边缘特征表达能力。随后, RDAC2f 以倒残差结构结合深度可分离卷积,有效减少通道间冗余计算,并通过 RSE 注意力实现特征增强与冗余抑制,进一步强化语义判别能力。多尺度特征融合网络利用上采样与跨层连接策略聚合深浅层特征,并在融合阶段持续引入 FASA 与 RDAC2f 以提升多尺度特征信息的融合能力。最终,融合后的特征由检测头实现类别预测与定位输出。

2.3 RDA-Bottleneck 模块

YOLOv8 中的 C2f 模块采用标准 Bottleneck 作为

基本特征变换单元,其核心由 2 层常规卷积块构成,并通过残差连接实现特征复用。尽管该结构在通用目标检测任务中具备较好的特征提取能力,但其仍存在以下不足:(1)连续使用标准卷积会引入较高的通道间耦合计算,导致参数量与 FLOPs 增加,不利于模型轻量化与边缘端部署;(2)原始 Bottleneck 缺乏显式的特征选择与通道重标定机制,难以针对中药饮片检测中常见的纹理细微、类别相似及背景干扰等问题进行有效抑制。为此,本研究提出一种轻量化 RDA-Bottleneck,用于替换原 C2f 模块中的 Bottleneck 单元,并将改进后的 C2f 结构记为 RDAC2f。RDA-Bottleneck 和 Bottleneck 的结构如图 4 所示。

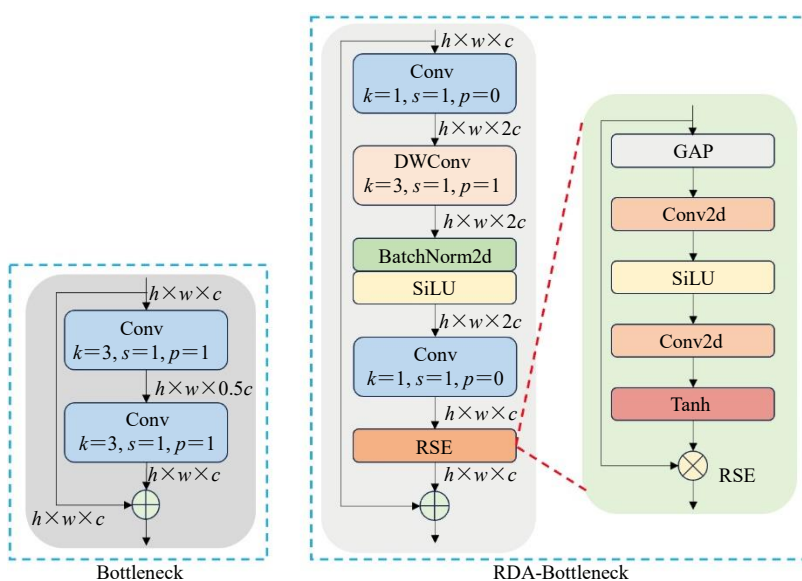


图 4 RDA-Bottleneck 和 Bottleneck 的网络结构

Fig. 4 Network structures of RDA-Bottleneck and Bottleneck

如图 4 所示,RDA-Bottleneck 采用倒残差设计,其整体结构由 1×1 升维 Conv、 3×3 深度可分离卷积 (DWConv)、 1×1 降维 Conv 以及通道 RSE 模块构成。RDA-Bottleneck 首先通过 1×1 Conv 对输入特征进行通道升维,以增强特征表达能力;随后引入 3×3 深度 DWConv 进行空间建模,在保持感受野不变的情况下显著降低计算复杂度;再通过 1×1 Conv 将特征映射回原通道维度,实现高效的信息压缩与融合。在此基础上,本研究原创性地提出残差挤压激励注意力机制 (residual squeeze-excitation attention, RSE-Attention) 并嵌入于模块输出端。RSE 在标准 SE 注意力结构^[19]的基础上引入残差缩放策略,将 Sigmoid 门控替换为 Tanh+残差连接形式,使通道权重被约束在 $(0, 2)$ 区间内,从而在避免

过度抑制特征的同时实现对显著通道的自适应增强,有利于保留中药饮片的基础纹理与边缘信息。整个模块中结合批量归一化 (batch normalization) 与 SiLU 激活函数以提升训练稳定性和非线性表达能力。通过在 C2f 模块中以 RDA-Bottleneck 替换原始 Bottleneck,所构建的 RDAC2f 在提升特征表达能力的同时,有效降低了冗余计算开销,还增强了通道级特征判别性。该设计使网络在复杂背景与细粒度差异显著的中药饮片检测任务中具备更强的特征提取效率与鲁棒性。

2.4 FASA 模块

在原始 YOLOv8 网络中,Conv 模块作为最基本的特征变换单元,被广泛用于骨干网络 Backbone 与 Neck 网络中的多处位置。该模块由标准二维卷

积、批归一化及非线性激活函数构成，能够在较低设计复杂度下完成局部特征提取。然而，这种“卷积、归一化、激活”的串联结构在复杂细粒度目标检测任务中仍存在以下局限性：（1）标准卷积在通道与空间维度上进行均匀建模，缺乏对关键通道与显著空间区域的自适应选择能力，难以提取中药饮片中具有判别意义的纹理、边缘与局部结构信息；

（2）原始 Conv 模块未显式建模频域信息，对高频细节与低频背景的区别能力有限，在复杂背景与类别相似度较高的场景下易引入冗余响应。针对上述问题，本研究提出一种融合频域感知与空间注意力的特征增强模块 FASA，用于替换 YOLOv8 主体结构中独立使用的 Conv 模块（骨干网络首层 Conv 仍予以保留），替换的具体位置见图 5。

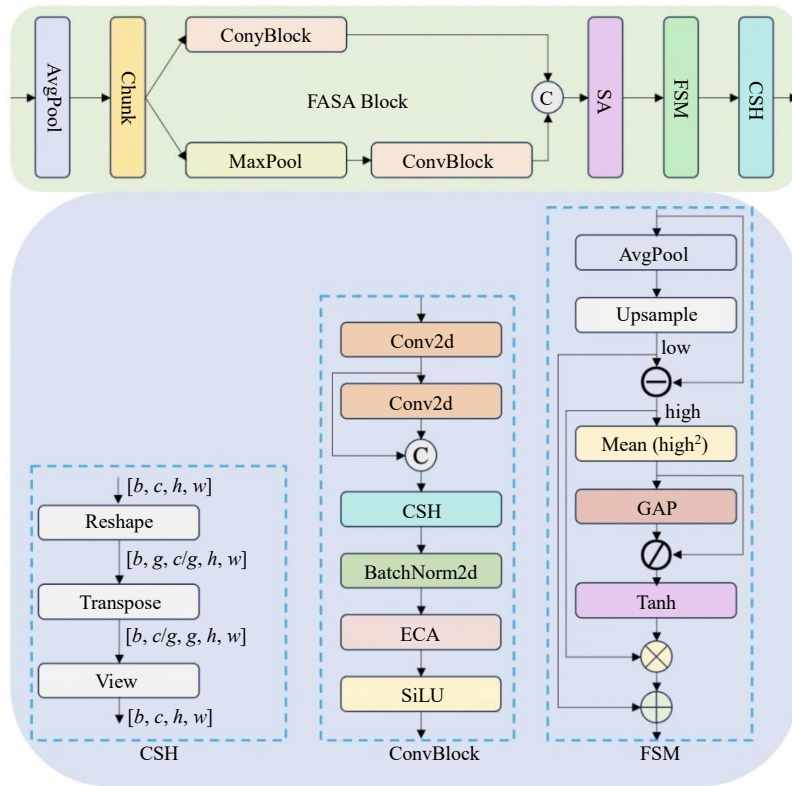


图 5 FASA 网络结构

Fig. 5 Network structures of FASA

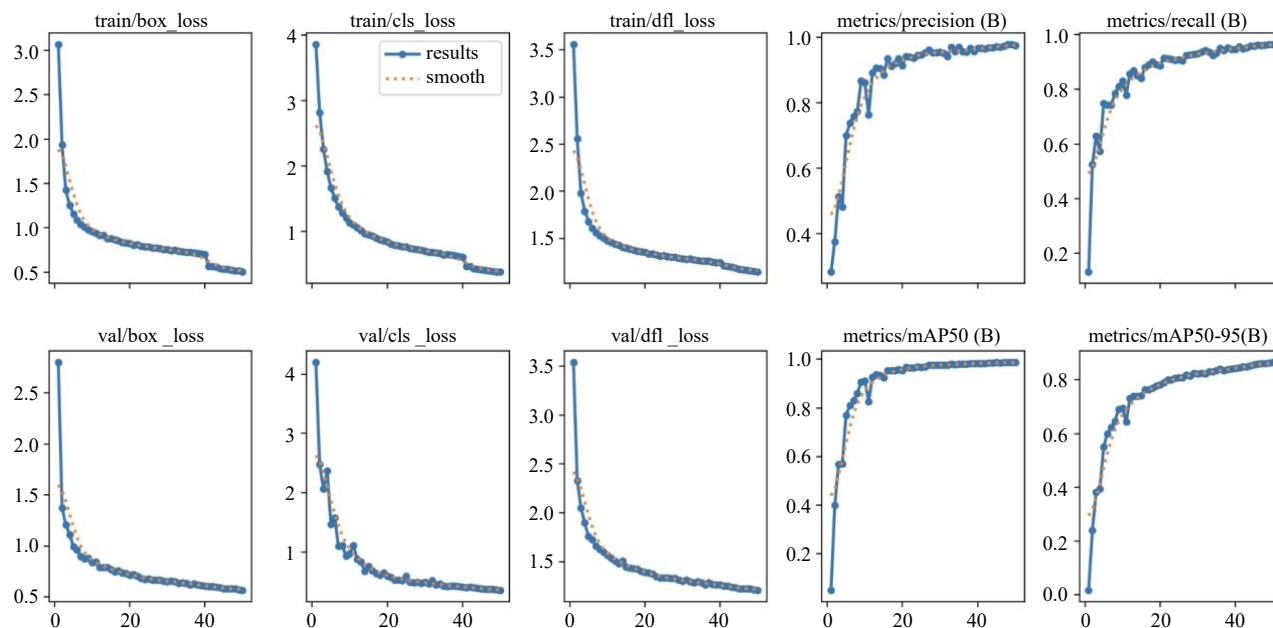
如图 5 所示，FASA 模块由 ConvBlock、SA^[20]以及 FSM 等核心组件协同构成，整体采用“并行建模、融合增强”的设计思路。首先，在特征变换阶段，FASA 采用改进的 ConvBlock 进行特征提取，ConvBlock 的输出端嵌入高效通道注意力（efficient channel attention, ECA）^[21]，实现对关键通道响应的自适应重标定，从而提升通道维度上的判别能力。其次，融合特征送入 SA，以增强目标区域的空间响应并抑制背景干扰。进一步地，提出 FSM 模块并嵌入在 FASA 输出端，FSM 显式分解特征中低频背景分量与高频细节，并依据高频能量分布对细节分量进行受约束的自适应缩放，从而在数值稳定的前提下强化纹理、边缘等高频信息表达。最终，经空间注意力与频域增强后的特征通过通道洗牌

（channel shuffle, CSH）通道洗牌进行再次融合输出。上述多阶段协同设计使 FASA 模块能够在空间、通道与频域 3 个层面同时提升特征表征能力，为后续检测头提供更加鲁棒且判别性更强的输入特征。

3 结果与分析

3.1 模型训练结果分析

如图 6 所示，FR-YOLO 在中药饮片数据集上的训练过程收敛良好且整体稳定。以 train 表示训练集，以 val 表示验证集，随着训练轮次增加，训练集边界框回归损失（train/box_loss）、分类损失（train/cls_loss）与分布焦点损失（train/dfl_loss）均快速下降并逐步趋于平稳，验证阶段对应的 val/box_loss、val/cls_loss 与 val/dfl_loss 亦保持一致的下降趋势，且训练曲线与验证曲线间差距较小，



X轴为训练的 epoch 数值, Y轴为各指标数值。

X-axis represents the training epoch value, and Y-axis denotes the values of each evaluation metric.

图 6 FR-YOLO 训练结果

Fig. 6 Training results of FR-YOLO

表明模型优化过程平滑、泛化性能较为稳定。与此同时, P 与 R 曲线在训练早期迅速提升并在中后期进入平台期, 说明模型能够有效学习中药饮片在复杂背景下的判别特征, 并保持较低的误检率与漏检率。进一步地, $mAP50$ 与 $mAP50-95$ 随训练持续上升并最终趋于稳定, 其中 $mAP50-95$ 的稳步提升反映了模型在更严格 IoU 阈值下仍具备较好的定位一致性与检测可靠性。总体而言, 该结果说明了 FR-YOLO 在形态多变、纹理细微及类别相似度较高的中药饮片数据中仍具备良好的训练稳定性与检测性能。

3.2 FR-YOLO 实验结果分析

FR-YOLO 在中药饮片数据集上的定量实验结果如表 2 所示。整体来看, 模型在 10 类饮片上取得了较高且较为均衡的检测性能, 所有类别的平均 P 与 R 分别达到 0.975 与 0.965, 表明 FR-YOLO 在复杂背景下能够有效抑制误检并保持较高的目标检出率。整体 $mAP50$ 达到 0.987, 说明模型在 IoU 阈值为 0.5 时具有稳定的定位与分类能力; $mAP50-95$ 为 0.866, 进一步反映在更严格的 IoU 区间内仍保持较好的综合检测性能。从不同类别的表现看, 荔枝核、甘草、板蓝根等类别取得了更高的 P 与 R , 例如, 荔枝核 $P=0.997$ 、 $R=0.988$, 甘草 $mAP50=$

表 2 FR-YOLO 实验定量结果

Table 2 Quantitative results of FR-YOLO experiment

类别	$P\uparrow$	$R\uparrow$	$mAP50\uparrow$	$mAP50-95\uparrow$
白术	0.973	0.947	0.989	0.866
板蓝根	0.990	1.000	0.989	0.888
草果	0.950	0.917	0.974	0.841
草豆蔻	0.965	0.937	0.986	0.902
冬虫夏草	0.978	0.955	0.981	0.744
茯苓	0.977	0.968	0.984	0.849
甘草	0.990	0.990	0.994	0.930
枸杞子	0.949	0.970	0.988	0.876
荔枝核	0.997	0.988	0.995	0.917
人参	0.984	0.982	0.987	0.850
所有类别 平均值	0.975	0.965	0.987	0.866

↑ 表示该指标数值越大越理想。

↑ indicates that the higher the value of this indicator, the better.

0.994、 $mAP50-95=0.930$, 板蓝根 $R=1.000$, 说明 FR-YOLO 对纹理特征相对清晰、目标形态较稳定的饮片类别具有较强的判别能力与定位一致性。相比之下, 冬虫夏草与草果在 $mAP50-95$ 上相对偏低 (冬虫夏草 0.744、草果 0.841), 这可能与该类饮片形态细长、局部纹理复杂及边界不清晰等因素有关, 导致高精度框回归更具挑战性。总体而言, FR-

YOLO在中药饮片多类别检测任务中表现出较强的泛化与鲁棒性：在保持高mAP50的同时，mAP50-95也维持在较高水平，说明该方法不仅能够实现可靠检测，还具备较好的定位性能，能够满足中药饮片自动化检测对精度与稳定性的需求。

图7展示了FR-YOLO在部分样本上的检测可视化结果。由图可知，模型对不同类别饮片（白术、板蓝根、草果、草豆蔻、人参等）均能精确识别，

预测类别或实例数量与标签保持一致，且在多目标密集分布场景下仍能较好地地区分相邻目标。此外，预测边界框与人工标注框整体高度重合，仅在个别样本的边缘区域存在轻微偏差，表明FR-YOLO具备较好的定位精度与目标覆盖能力。总体而言，该可视化结果从定性角度验证了FR-YOLO对复杂背景与细粒度纹理的中药饮片数据仍具有较强的泛化性与鲁棒性。



图7 FR-YOLO检测效果展示

Fig. 7 Demonstration of FR-YOLO detection effect

3.3 对比实验结果分析

对比实验定量结果如表3所示，FR-YOLO在中药饮片数据集上取得了更优的检测性能。与基线YOLOv8n相比，FR-YOLO的mAP50由0.985提升至0.987（提升0.2%），mAP50-95由0.864提升至0.866（提升0.2%），表明其在整体检测精度上实现了显著提升。FR-YOLO的R为0.965，高于YOLOv8n的0.958，也优于其他对比模型（如YOLOv11n为0.964、YOLOv10n为0.958、YOLOv9t为0.950），说明在复杂背景、遮挡与尺度变化等中药饮片实际场景中，FR-YOLO具有更强的目标检测能力。在模型复杂度方面，FR-YOLO同样展现出更显著的轻量化优势。相较于基线模型YOLOv8n，FR-YOLO的参数量由3.01 M降至1.82 M（下降

39.5%），FLOPs由8.2 G降至5.4 G（下降34.1%），表明该模型在显著降低计算复杂度与参数量的同时，仍能够保持甚至提升整体检测性能，说明FR-YOLO具有更高的检测效率。进一步与其他模型比较可知，FR-YOLO同样具有更低的模型复杂度优势，例如YOLOv13n的参数量与FLOPs分别为2.46 M和6.4 G，YOLOv11n分别为2.59 M和6.5 G，均高于FR-YOLO。上述结果表明，在中药饮片数据集上，FR-YOLO不仅具有最高的检测精度性能，其参数量和计算复杂度也最低，表明本研究方法在精度提升与模型计算开销之间实现了更优的平衡。

3.4 消融实验结果分析

FR-YOLO的消融实验结果如表4所示。相较于基线YOLOv8n (Base)，仅使用RDA-Bottleneck

表3 FR-YOLO 与其他模型的对比实验结果

Table 3 Comparative experimental results of FR-YOLO and other models

模型	$P\uparrow$	$R\uparrow$	mAP50 \uparrow	mAP50-95 \uparrow	参数量/M \downarrow	FLOPs/G \downarrow
YOLOv8n	0.979	0.958	0.985	0.864	3.01	8.2
YOLOv9t	0.955	0.950	0.981	0.850	2.66	11.1
YOLOv10n	0.967	0.958	0.982	0.857	2.71	8.4
YOLOv11n	0.969	0.964	0.983	0.854	2.59	6.5
YOLOv12n	0.959	0.948	0.983	0.843	2.52	6.0
YOLOv13n	0.961	0.954	0.984	0.853	2.46	6.4
FCOS	0.968	0.956	0.982	0.858	32.02	200.5
TOOD	0.972	0.960	0.984	0.862	31.98	201.1
RetinaNet	0.963	0.953	0.983	0.855	37.74	95.7
FR-YOLO	0.975	0.965	0.987	0.866	1.82	5.4

\uparrow 表示该指标数值越小越理想。

\downarrow indicates that the smaller the indicator value, the better.

表4 FR-YOLO 消融实验定量结果

Table 4 Quantitative results of FR-YOLO ablation experiments

模型	$P\uparrow$	$R\uparrow$	mAP50 \uparrow	mAP50-95 \uparrow	参数量/M \downarrow	FLOPs/G \downarrow
Base (YOLOv8n)	0.979	0.958	0.985	0.864	3.01	8.2
Base+RDA	0.972	0.961	0.983	0.862	2.30	6.5
Base+FASA	0.980	0.963	0.986	0.865	2.51	7.1
Base+RDA+FASA	0.975	0.965	0.987	0.866	1.82	5.4

时,参数量由 3.01 M 降至 2.30 M、FLOPs 由 8.2 G 降至 6.5 G,同时 R 由 0.958 提升至 0.961。该结果表明,RDA-Bottleneck 通过倒残差与深度可分离卷积可有效降低通道耦合冗余计算,验证了其轻量化设计的有效性。仅使用 FASA 时,mAP50 与 mAP50-95 分别提升至 0.986 与 0.865, P 提升至 0.963,同时参数量与 FLOPs 分别降低至 2.51 M 与 7.1 G。说明 FASA 能够在较低计算代价下能强化对显著区域的空间响应,并通过频域细节增强提升纹理与边缘等高频信息表征,从而更有效地抑制复杂背景干扰并提升细粒度类别的可分性,证明了其特征增强作用。当 RDA-Bottleneck 与 FASA 联合使用(Base+RDA+FASA,即 FR-YOLO)时取得最佳综合性能:参数量与 FLOPs 进一步降至 1.82 M 与 5.4 G,召回率提升至 0.965,mAP50 与 mAP50-95 分别达到 0.987 与 0.866。

上述结果表明,RDA-Bottleneck 更主要地体现在降低参数量与 FLOPs 方面,说明其在轻量化设计上具有较好的效果;FASA 在保持较低计算开销的同时,对 R 、mAP50 和 mAP50-95 指标表现出一定积极作用。两者联合后,虽然精度提升幅度较小,但模型参数量和 FLOPs 显著降低,说明本研究方

在较大幅度降低模型复杂度的同时,仍保持了较好的检测性能。

4 结论

本研究提出了一种用于中药饮片检测的轻量化模型 FR-YOLO,该方法在 YOLOv8n 框架基础上做了 2 个改进:(1)提出 RDA-Bottleneck 模块以替换 C2f 中的标准 Bottleneck,利用倒残差与深度可分离卷积降低冗余计算,并结合通道重标定提升有效特征响应;(2)提出 FASA 模块替换主体网络中的 Conv 单元,通过空间注意与频域细节增强协同强化纹理与边缘表征,从而提升模型对背景干扰的抑制能力与细粒度类别可分性。实验结果表明,在自建 10 类中药饮片数据集上,FR-YOLO 在保持检测性能基本稳定的同时取得了较好的轻量化效果。相较于 YOLOv8n,模型参数量与 FLOPs 分别降低 39.5%和 34.1%,mAP50 与 mAP50-95 均实现了 0.2% 的小幅提升。消融实验进一步表明,RDA-Bottleneck 与 FASA 在降低模型复杂度和保持检测性能方面具有一定互补作用,使模型在精度与效率之间取得了较好的平衡。

尽管本研究取得了一定成果,但仍存在提升空间。未来工作将从以下 2 点展开:(1)进一步扩充

饮片类别与采集场景, 构建更大规模、多样性的数据集, 并引入跨域测试以系统评估模型泛化能力;
(2) 面向实际应用部署, 将 FR-YOLO 部署至真实生产与调剂等业务场景中并开展工程化验证, 系统评估其在实际应用中的实时性、稳定性与可用性。

利益冲突 所有作者均声明不存在利益冲突

参考文献

- [1] 史亚博, 李铭轩, 傅饶, 等. 在线检测技术在中药饮片智能化生产与质量控制中的研究现状与展望 [J]. 中国现代中药, 2024, 26(5): 881-888.
- [2] 王超超, 张先超, 谷正昌, 等. 中药材及饮片检测中人工智能应用探讨 [J]. 中国工程科学, 2024, 26(2): 245-254.
- [3] 周明, 周金海, 张燕群, 等. 中药饮片性状质量智能检测关键技术研究 [J]. 世界科学技术—中医药现代化, 2023, 25(5): 1580-1589.
- [4] 张琴, 郭为安. 深度学习小目标检测算法综述 [J]. 计算机应用研究, 2025, 42(10): 2893-2904.
- [5] 刘晶, 张璐, 刘加峰. 基于 EfficientDet 算法的易混淆中药饮片识别 [J]. 北京生物医学工程, 2025, 44(4): 373-378.
- [6] 李云阳, 李根, 闫磊. 基于改进 YOLOv5 的中药饮片缺陷检测算法 [J]. 林业机械与木工设备, 2024, 52(5): 71-75.
- [7] 谷瑞, 宋翠玲, 李元昊. 基于 BIM-Res2Net-ICA 的中药饮片图像识别模型 [J]. 时珍国医国药, 2025, 36(6): 1192-1200.
- [8] 孙兴, 李杨, 毛天驰, 等. 改进 YOLOv8 在中药饮片自动识别标注中的研究与应用 [J]. 中草药, 2025, 56(1): 216-226.
- [9] Varghese R, M S. YOLOv8: A novel object detection algorithm with enhanced performance and robustness [A] // 2024 International Conference on Advances in Data Engineering and Intelligent Computing Systems (ADICS) [C]. Chennai: IEEE, 2024: 1-6.
- [10] Paszke A, Gross S, Massa F, et al. PyTorch: An imperative style, high-performance deep learning library [J/OL]. *arXiv*, (2019-12-13) [2025-12-02]. <https://arxiv.org/abs/1912.01703>
- [11] Wang C Y, Yeh I H, Mark Liao H Y. YOLOv9: Learning what you want to Learn using programmable gradient information [A] // Computer Vision-ECCV 2024 [C]. Cham: Springer, 2025: 1-21.
- [12] Chen H, Chen K, Ding G G, et al. YOLOv10: Real-time end-to-end object detection [A] // Advances in Neural Information Processing Systems 37 [C]. Vancouver BC: Neural Information Processing Systems Foundation Inc. (NeurIPS), 2024: 107984-108011.
- [13] Khanam R, Hussain M. YOLOv11: An overview of the key architectural enhancements [J/OL]. *arXiv*, (2024-10-24) [2025-12-02]. <https://arxiv.org/abs/2410.17725>.
- [14] Tian Y J, Ye Q X, Doermann D. YOLOv12: Attention-centric real-time object detectors [J/OL]. *arXiv*, (2025-02-18) [2025-12-02]. <https://arxiv.org/abs/2502.12524>
- [15] Lei M Q, Li S Q, Wu Y H, et al. YOLOv13: Real-time object detection with hypergraph-enhanced adaptive visual perception [J/OL]. *arXiv*, (2025-0905) [2025-12-02]. <https://arxiv.org/abs/2506.17733>
- [16] Tian Z, Shen C, Chen H, et al. Fcos: Fully convolutional one-stage object detection [A] // Proceedings of the IEEE/CVF international conference on computer vision [C]. Seoul: IEEE, 2019: 9627-9636.
- [17] Feng C J, Zhong Y J, Gao Y, et al. TOOD: Task-aligned one-stage object detection [A] // 2021 IEEE/CVF International Conference on Computer Vision (ICCV) [C]. Montreal QC: IEEE, 2022: 3490-3499.
- [18] Lin T Y, Goyal P, Girshick R, et al. Focal loss for dense object detection [A] // 2017 IEEE International Conference on Computer Vision (ICCV) [C]. Venice: IEEE, 2017: 2999-3007.
- [19] Hu J, Shen L, Sun G. Squeeze-and-excitation networks [A] // 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition [C]. Salt Lake City: IEEE, 2018: 7132-7141.
- [20] Woo S, Park J, Lee J Y, et al. CBAM: Convolutional block attention module [A] // Computer Vision-ECCV [C]. Cham: Springer, 2018: 3-19.
- [21] Wang Q L, Wu B G, Zhu P F, et al. ECA-net: Efficient channel attention for deep convolutional neural networks [A] // 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) [C]. Seattle WA: IEEE, 2020: 11531-11539.

[责任编辑 潘明佳]