doi:10.6041/j.issn.1000-1298.2025.10.044

基于 FDC－Mamba 的关龙胆根茎实例分割与表型参数提取

崔红光[1]　刘海涛[1]　马有泽[1]　黄文忠[2]　李宏博[3]　王铁军[1]

(1.沈阳农业大学工程学院，沈阳 110866；2.抚顺市农业农村发展服务中心，抚顺 113006；
3.沈阳农业大学园艺学院，沈阳 110866)

摘要：针对关龙胆根茎中茎痕与残留茎基表型特征高度相似，且根茎尺寸小、形态复杂，导致图像分割特征提取困难、识别精度不足等问题，本文提出了焦点调制-动态检测头-上下文引导-Mamba(Focal Modulation－DyHead seg－Context Guided－Mamba，FDC－Mamba)关龙胆根茎实例分割模型。首先，为解决关龙胆相邻根丝边界模糊、缠绕部位重叠问题，引入目标检测型 Mamba(Object detection Mamba，ODMamba)主干网络补充纹理细节，加强结构一致性；其次，通过融合 Focal Modulation 与 Context Guided 结构部分，增强多尺度感知能力和细节分割能力；最后，将 DyHead 结构结合辅助检测头(Auxiliary Head)训练策略，开发一种用于实例分割新训练结构 DyHead seg，提高信息传递效率、优化学习过程。与其他常用实例分割模型(YOLO 系列、Mask R－CNN、PointRend、HTC、SOLOv2、RT－DETR、HYPER)、不同特征金字塔架构模块(RepBN、AIFI、LSKA)、不同下采样结构模块(SRFD、ADown、CARAFE、EUCB、Gold－YOLO、HWD、PSConv、SODConv、WaveletPool)在关龙胆根茎数据集上进行对比，改进后模型完成了对关龙胆根茎实例分割，在根茎边缘和细小区域定位方面具有更高准确度，Box 类型和 Mask 类型精度 P、AP_{50}、AP_{95} 分别提升 6.52、5.09、5.44 个百分点和 4.49、2.68、1.16 个百分点。基于分割结果，提出了关龙胆根长、根部粗细度、含杂率和色度 4 种表型参数提取方法。试验结果表明，所提出模型分割精度(Mask 类型 P)达 87.12%，比基线模型高 4.49 个百分点。关龙胆表型参数提取结果与人工测量结果相对误差均在 5% 以内。本文对以关龙胆为代表的根茎类中药材表型特征提取具有较高的准确性，可为后续炮制工艺与装备研发奠定基础。

关键词：关龙胆；实例分割；表型参数提取；深度学习；FDC－Mamba

中图分类号：S126　　**文献标识码**：A　　**文章编号**：1000-1298(2025)10-0500-12　　**OSID**：

Instances Segmentation and Extraction of Phenotypic Parameters of *Gentiana scabra* Bunge Roots Based on FDC－Mamba

CUI Hongguang[1]　LIU Haitao[1]　MA Youze[1]　HUANG Wenzhong[2]　LI Hongbo[3]　WANG Tiejun[1]

(1. *College of Engineering*, *Shenyang Agricultural University*, *Shenyang* 110866, *China*
2. *Fushun Agriculture and Rural Affairs Development Service Center*, *Fushun* 113006, *China*
3. *College of Horticulture*, *Shenyang Agricultural University*, *Shenyang* 110866, *China*)

Abstract: Aiming to address the problems of challenges of difficult feature extraction and insufficient recognition accuracy due to the highly similar phenotypic characteristics of stem scars and residual stem bases in *Gentiana scabra* Bunge roots, as well as their small size and complex morphology, the Focal Modulation－DyHead seg－Context Guided－Mamba (FDC－Mamba) *Gentiana scabra* Bunge roots instances segmentation was proposed. Firstly, to solve the problems of blurred boundaries between adjacent root fibers and overlapping in entangled parts of *Gentiana scabra* Bunge, the object detection Mamba (ODMamba) backbone network was introduced to supplement texture details and enhance structural consistency. Secondly, by integrating the Focal Modulation and Context Guided structural, the multi-scale perception abilities and detail segmentation abilities were enhanced. Finally, the dynamic head structure was combined with the Auxiliary Head training strategies to develop a training structure for

收稿日期：2025－05－09　修回日期：2025－06－05

基金项目：辽宁高校基本科研业务费项目(LJKMZ20220998)

作者简介：崔红光(1977—)，女，副教授，博士，主要从事道地药材生产机械化技术与装备研究，E-mail：chg7763@syau.edu.cn

通信作者：王铁军(1991—)，男，副教授，博士，主要从事农业生物系统工程研究，E-mail：tiejunwang@syau.edu.cn

instance segmentation, DyHead-seg, which improved information transmission efficiency and optimized the learning process. The proposed model was compared with other commonly used instance segmentation models (YOLO series, Mask R－CNN, PointRend, HTC, SOLOv2, RT－DETR, HYPER), different feature pyramid architecture modules (RepBN, AIFI, LSKA), and different downsampling structure modules (SRFD, ADown, CARAFE, EUCB, Gold－YOLO, HWD, PSConv, SODConv, WaveletPool) on the *Gentiana scabra* Bunge roots datasets. The improved model successfully completed the instance segmentation of *Gentiana scabra* Bunge roots. FDC－Mamba model had high accuracy in the positioning of root and stem margins and small areas. The Box type and Mask type P, AP_{50}, and AP_{95} were increased by 6.52, 5.09, 5.44 percentage points and 4.49, 2.68, 1.16 percentage points, respectively. Based on the segmentation results, four phenotypic parameters extraction methods for *Gentiana scabra* Bunge sorting were figured put, including root length, root thickness degrees, impurity rate, and chromaticity. The experimental results showed that the proposed model achieved a Mask type P of 87.12%, which was 4.49 percentage points higher than that of the baseline model. The relative errors between the extraction results of *Gentiana scabra* Bunge phenotypic parameters and the manual measurement results were all within 5%. The research result had high accuracy in extracting phenotypic features of roots Chinese medicinal materials represented by *Gentiana scabra* Bunge, providing a foundation for subsequent processing technology and equipment research and development.

Key words: *Gentiana scabra* Bunge; instance segmentation; phenotype parameter extraction; deep learning; FDC－Mamba

0　引言

关龙胆属于龙胆科龙胆属多年生草本植物，主产于辽宁、吉林和黑龙江等地，以其干燥后根及根茎入药，具有清热燥湿、泻肝胆火的功能，上端有茎痕或残留茎基[1－3]。龙胆洗净后进行切段、干燥炮制工序前，需剔除机械损伤根茎、明确切段范围[4－5]，但目前该环节依赖于人工识别挑选，导致关龙胆生产成本高、生产效率低。快速识别关龙胆根茎并分割其切段位置、获取其表型参数可提升识别精准度，降低人工劳动强度、提高生产效率。

目前已有农业生产方面实例分割及表型参数提取方法研究，如病害识别、采摘定位、养殖监测等[6－11]。学者以 Swin Transformer、Mask R－CNN、AutoLNet 等网络模型优化实例分割细节特征提取效果[12－16]，采用 Kinect V3 传感器、Transformer 编码模块、神经辐射场等方法重建作物三维模型、获取表型参数[17－20]。AL RIZA 等[21]利用改进的 YOLO v8 实例分割技术评估可可豆发酵程度，通过集成选择性内核块（LSKBlock）取代 C2f 模块开发了 YOLO－CoLa 模型，提高检测精度。然而，尽管实例分割技术在农业生产分割任务中取得了进步，但在处理不规则块状的圆柱形扭曲、带有横皱纹的关龙胆根茎复杂图像时，形态复杂导致的图像分割特征提取困难与识别精度不足等问题亟待解决。苗全龙等[22]将快速神经网络、局部卷积、注意力机制等相结合，对 YOLO v8n－Seg 实例分割模型进行改进，实现双孢蘑菇目标实例分割，并提出双孢蘑菇子实体 4 种表型参数提取方法，为生长模型建立提供技术基础。研究表明，基于 YOLO 等典型框架改进算法模型，通过关键模块优化降低运算量、提升获取效率，可实现作物表型参数感知与提取[23－27]。

结合关龙胆机械化、智能化生产作业需求，本研究以清洗后关龙胆为研究对象，基于 Mamba 模型及 YOLO 框架提出一种高效率、高精度实例分割模型 FDC－Mamba，以及关龙胆表型参数获取计算方法，并通过对比试验验证模型精准度，以期为根茎类道地药材数字化生产提供技术支撑。

1　材料与方法

1.1　数据采集与标定

关龙胆图像均拍摄于辽宁省中药重点药用植物保存圃（123°56′N，41°83′E），采用搭配有英特尔 RealSense D435i 型深度实感摄像头和 1 个散射光源的图像采集装置，以固定高度俯视角度下拍摄关龙胆图像。图像采集装置如图 1 所示。

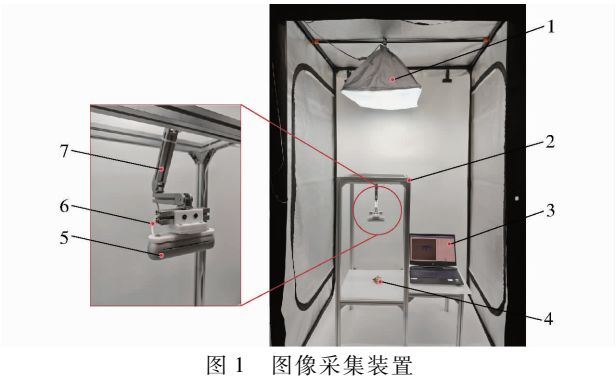


图 1　图像采集装置

Fig. 1　Image acquisition device

1. 补光灯　2. 采样平台　3. 信息采集系统　4. 关龙胆样本　5. 深度相机　6. 传输线　7. 相机支架

为得到关龙胆图像对应的空间分辨率 A，即每 A 个像素在实际中对应厘米值，需进行标定。具体方法为：将 3 个已知长度标定尺水平放置在拍摄区域并拍摄图像，用绘图软件测量标定尺起点和终点间的像素长度（欧几里得距离）后，计算像素数量和实际长度间的比例关系，即空间分辨率 A。测量 20 组标定尺空间分辨率，得到平均空间分辨率 A 为 28.3 像素/cm，表明实际距离 1 cm，图像上为 28.3 像素。

1.2 数据集与训练集参数设置

关龙胆根茎数据集包含 600 幅原始图像，分辨率 640 像素 ×640 像素，使用 Labelme 图像分割标注工具标注关龙胆茎痕和残留茎基，然后将原始图像按比例 8∶2 分割为训练集和验证集。同时，为避免训练过程中过拟合，提高模型泛化能力，采用随机组合旋转、翻转、镜像、改变亮度、高斯噪声等图像增强技术增强数据多样性，图 2 为部分增强后图像。数据增强后样本数量增加至 2 400 幅。



图 2 数据增强图像示例

Fig. 2 Examples of data enhanced images

使用 Windows 10 操作系统及 PyTorch 深度学习框架。设置环境信息：中央处理器为英特尔 i5 12600KF，显卡为 NVIDIA GeForce RTX 3060 8G，Torch 版本为 2.1.2 + cu118，编程语言为 Python 3.10.8。训练阶段超参数如下：输入图像尺寸为 640 像素 ×640 像素，批量大小为 8，不加载预训练权重，优化器选用随机梯度下降（Stochastic gradient descent，SGD），训练轮次为 300 个周期，使用马赛克在线增强。初始学习率为 0.001。其他训练参数均采用 Mamba - YOLO[28] 源代码中默认设置[21-22]。

1.3 实例分割模型改进

1.3.1 FDC - Mamba 模型

Mamba - YOLO[29] 是一种基于 YOLO v8 框架、引入状态空间机制的轻量级目标检测模型。该网络引入残差门控制模块（Residual gated block，RGBlock）新结构。使用 2 维选择性扫描（Selective scan 2D，SS2D）机制对输出特征进行细化，通过高维点积操作增强通道间相关性，提取特征表示，适用于关龙胆等细粒度纹理与长条形目标识别。其主干网络采用多层次堆叠的 ODSSBlock 模块（One dimensional state space block）替代传统 C2f 模块（Coarse to fine），在保证计算效率的同时增强特征提取能力。传统卷积模型在面对根系缠绕、多重交叉等复杂目标结构时存在局部感受野受限、细节特征提取不足等问题，导致分割边界模糊、断裂等现象，Mamba - YOLO 以状态空间建模为核心，在多层次语义融合基础上有较强结构保持能力与纹理一致性建模能力，具备状态空间建模优势，为本研究所涉及的精细实例分割与表型参数提取提供良好基础。因此，以 Mamba - YOLO - T 模型为基础，设计一种适用于关龙胆根茎实例分割的改进模型架构 FDC - Mamba，其总体架构如图 3 所示。

模型使用带有选择性状态空间模型（State space model，SSM）的 ODSSBlock 构建主干网络，采用简单下采样步骤（Simple Stem）将输入图像划分为多个图像块，使用视觉线索合并（Vision Clue Merge）模块完成下采样操作。将从主干中提取的多层特征（C_3、C_4、C_5），通过 Focal modulation 焦点调制策略，高效捕捉多尺度信息进行融合；采用 Context Guided 模块代替常规步长卷积以增强下采样阶段上下文表达。高层语义与底层空间特征由 ODSSBlock 提炼整合，最终输出特征图 P3、P4、P5 后送入检测头。在检测头部分，设计并使用融合多重注意力机制和语义分割任务的检测头 DyHead seg 生成检测结果。上述模块协同构成 FDC - Mamba 模型，为实现高精度根茎实例分割与后续表型参数提取提供结构支撑与特征保障。本架构通过将 Focal Modulation、DyHead seg、Context Guided 模块整合到基线模型中，有效捕捉全局依赖关系，无缝融合局部信息，为关龙胆分类检测提供一种基于机器视觉的稳定适用方案。

1.3.2 ODMamba 主干网络

ODMamba 由 Simple Stem、Vision Clue Merge 和 ODSSBlock 组成。网络首先通过 Simple Stem 模块进行下采样，得到空间分辨率为 $H/4 \times W/4$ 的 2 维特征图（H 表示输入图像高度维度，W 表示输入图像宽度维度）。然后，由 ODSSBlock 模块与 Vision Clue Merge 模块交替堆叠实现进一步下采样。在传统识别模型中，模型初始输入模块通常使用步长为

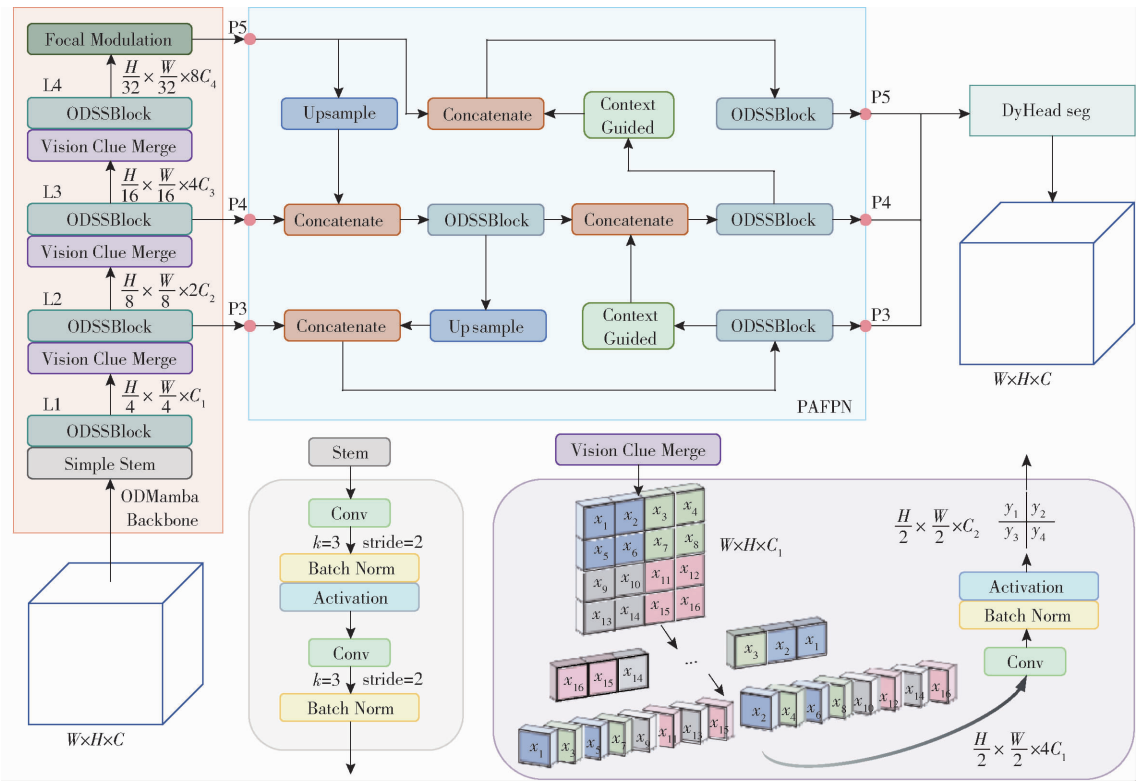


图 3　FDC－Mamba 总体架构图

Fig. 3　Overall architecture diagram of FDC－Mamba

4 的卷积实现，限制模型可优化性，影响模型整体性能。ODMamba 使用一种步长为 2、卷积核为 3 的 Simple Stem 层代替传统输入模块，平衡效率与性能。同时 ODMamba 使用一种新下采样卷积 Vision Clue Merge，更好适配 ODSSBlock 中 SS2D 对不同阶段特征图的选择性处理。与 CNN 和 Vision Transformer（ViT）不同，ViT 结构中步长为 2 的 3×3 卷积型下采样，Vision Clue Merge 移除了归一化操作并使用 4 倍压缩的逐点卷积（Pointwise Convolution）进行下采样，对二维特征图进行通道划分，保留前一层 SS2D 选择特征信息。

ODMamba 核心模块 ODSSBlock 如图 4 所示，主要由用于增强空间适应能力的 SS2D 模块和用于通道维度特征建模的 RGBlock 模块组成。二者均使用残差连接以增强表达能力和梯度流，以交替堆叠方式组成一个特征提取单元，分别在不同深度尺度上对输入图像进行局部与全局联合建模。其中，SS2D 模块主要分为 3 个阶段。将输入图像扩展为多个子图，每个子图沿不同方向提取特征信息，将不同方向信息作为序列整合，输入状态空间操作中，融合生成全局空间特征，实现图像中所有区域全面覆盖，并通过系统性方向变换增强对图像特征多维捕获能力。

RGBlock 模块承袭 VMamba 门控机制（Gated multilayer perceptron，Gated MLP），将输入特征 f'_A 和

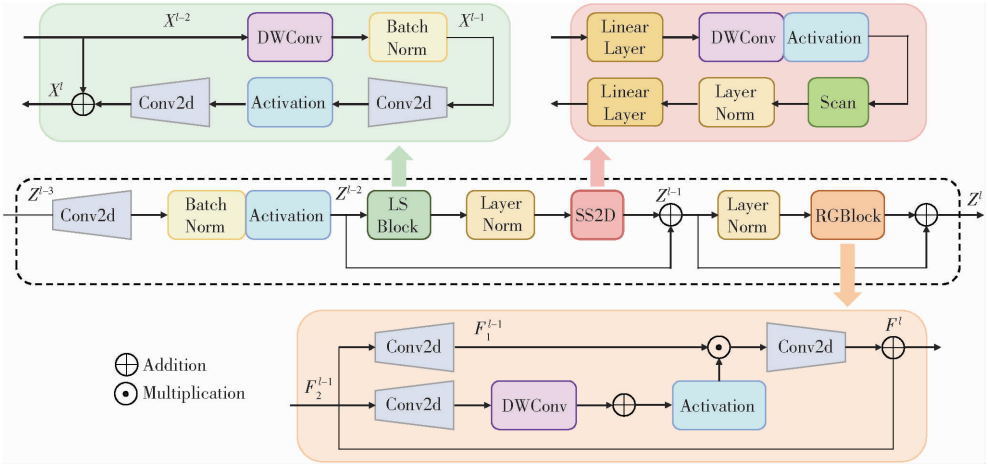


图 4　ODSSBlock 模块结构图

Fig. 4　Structure diagram of ODSSBlock

f'_R分别送入两条分支以保留局部特征 *Rl local* 与全局特征 *Rl global*，并在全局分支上引入深度可分离卷积（Depthwise separable convolution，DSC）作为位置编码模块；整体结构通过残差拼接进行信息融合，实现梯度高效反传。

在 FDC－Mamba 模型中，ODSSBlock 模块被嵌入至主干网络和多个下采样阶段，即通过 SS2D 模块补充纹理细节、RGBlock 模块加强结构一致性，二者联合可有效增强模型在复杂根茎图像中表达能力。试验中，该模块对相邻根丝边界模糊、缠绕部位重叠等问题具有良好适应性，在实例轮廓还原与根茎结构保持方面表现较优。

1.3.3 Focal Modulation 模块

前期研究发现背景干扰情况下，传统卷积神经网络难以同时处理细小目标和大目标。对比基于 Transformer 网络模型，通过自注意力实现全局特征建模架构虽表现较好，但计算成本高。因此，取消了传统架构中空间金字塔池化结构（Spatial pyramid pooling structure，SPPF）模块拼接特征图操作，引入 Focal Modulation 模块[30]，采用深度可分离卷积（Depthwise separable convolution，DWConv）和 Gated MLP 来建模上下文间交互关系，有效分离目标与背景信息。该模块通过逐层聚合局部上下文与不同尺度远程上下文，采用逐通道调制机制使模型聚焦关键区域，减小背景干扰。

如图 5 所示，Focal Modulation 模块主要通过 3 步操作增强网络多尺度感知能力。首先，输入通过全连接层（Linear）变成 query（q，查询投影函数）；然后，另一输入通过 Linear 后依次送入 3 层不同尺度卷积网络，并提取出不同感受野（Receptive-field）信息；经过 Gated MLP 过滤有效信息后，相加组成调制器（modulator）；最后与 q 相乘作为输出。这与自注意力机制（Self-attention，SA）通过 q 和 key（k）做每部分点乘相比，更快速高效捕捉全局信息。

模型引入 Focal Modulation 模块，在面对多尺度目标时能自适应地调整各特征通道权重。在关龙胆根茎实例分割任务中，Focal Modulation 模块有效提高目标细节分割能力，尤其在根茎边缘和细小区域定位方面表现出更高准确度。

1.3.4 Context Guided 模块

传统下采样方法易导致细节丢失，尤其在目标边界较模糊或形态复杂情况下，图像中细微特征易被消除。因此，引入 Context Guided 模块[31]，通过空洞卷积（Dilated Conv）与特征加权重标定相结合，在下采样过程中动态调整特征图，从而更好保持图像中细节和结构。

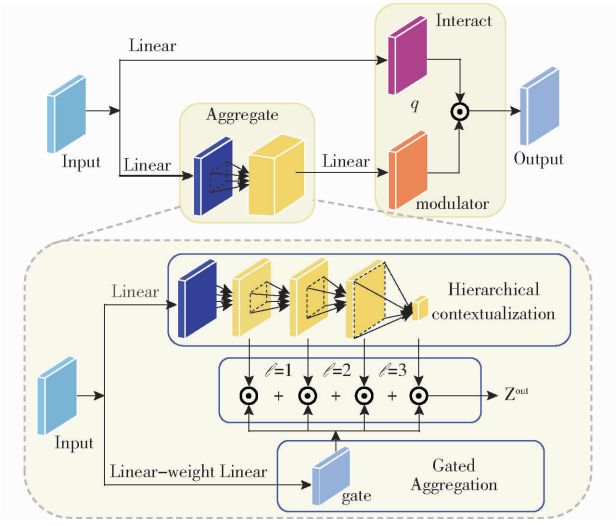


图 5 Focal Modulation 模块结构图

Fig. 5 Structure diagram of Focal Modulation block

如图 6 所示，Context Guided 模块将输入特征首先分别送入局部特征提取器 $f_{loc}(*)$ 和空洞率 r（%）的空洞卷积 $f_{sur}(*)$ 中，输出结果在通道维度通过联合特征提取层 $f_{joi}(*)$ 建模。之后，使用通道加权机制 $f_{glo}(*)$ 自适应地加强有效通道信息、抑制无效信息，最终输出 Y 为残差连接形式，将增强特征与原始输入相加。具体操作为

$$\boldsymbol{Y}=\boldsymbol{X}+\mathrm{Concat}(\mathrm{Conv}_{3\times3}(\boldsymbol{X}),\mathrm{DilatedConv}_{3\times3,r}(\boldsymbol{X}))\gamma \tag{1}$$

式中 $\boldsymbol{X}$——输入视觉特征图，$\boldsymbol{X}\in R^{H\times W\times C}$

γ——通道注意力权重

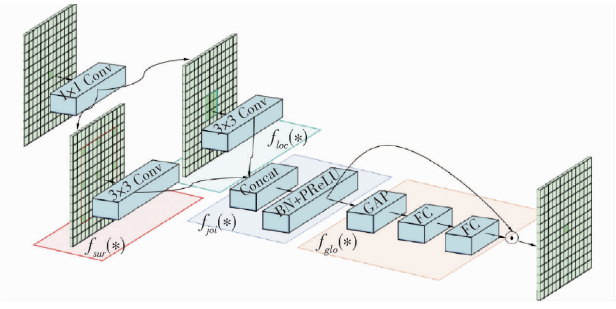


图 6 Context Guided 模块

Fig. 6 Context Guided block

此方式增强图像中有效特征，尤其在复杂背景和遮挡区域中的细节和目标形态。通过联合引导机制，模型能在不同图像背景下，有针对性地调整特征图重点，提升分割精度。

1.3.5 DyHead seg 模块

传统检测头虽能进行边界框回归，但无法有效捕捉目标细节和形态，尤其是对于细长目标或目标部分遮挡情况。因此，引入自设计 DyHead seg 模块，以 DyHead 模块[32]为主体，设计多视角注意力组合器，在 DyHead 模块基础上融合 Auxiliary Head 设计思想和固定输入通道操作，提高目标掩码预测精

度，适用于边界模糊、目标遮挡或目标形态不规则情况。

如图 7 所示，DyHead seg 模块结合动态卷积和多维度注意力机制，在不同层级、不同空间位置、不同任务通道类型 3 个维度上分别构建注意力机制，从而形成统一检测头结构，最大化提升性能。

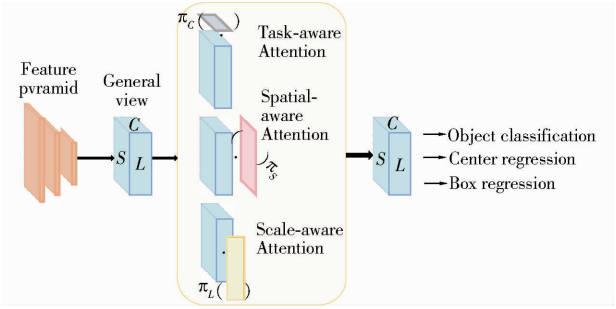


图 7　DyHead seg 模块

Fig. 7　DyHead seg block

该模块包含 3 种不同类型注意力机制，各机制侧重不同建模角度：尺度感知注意力机制（Scale-aware Attention）侧重跨特征层次建模对象尺度差异，空间感知注意力机制（Spatial-aware Attention）侧重对图像中空间位置进行自适应建模，任务感知注意力机制（Task-aware Attention）侧重在通道维度上区分分类、边界框回归等任务特性，进而提升检测头对目标细节表达能力。首先将主干网络输出特征表示为一个 3 维张量，其维度包括：来自不同语义深度特征图、图像中空间位置、特征通道数量，然后在这 3 种维度上分别引入注意力机制，均以残差结构集成入主网络，并与原始输入进行逐元素加和。

如图 8 所示，DyHead seg 同时引入基于 Auxiliary Head 设计方法。特征同时输入主检测头（Lead Head）和 Auxiliary Head 中，但仅 Lead Head 输出用于推理，Auxiliary Head 在训练时参与损失回传、不影响最终预测结果。图中 fine label 用于训练 Lead Head，而 Auxiliary Head 使用 fine label 中筛选过的 coarse 标签进行训练。为后续的实例掩码生成和表型分析提供了更加精确和稳定的支持。原始图像与分割后掩膜图像如图 9 所示。

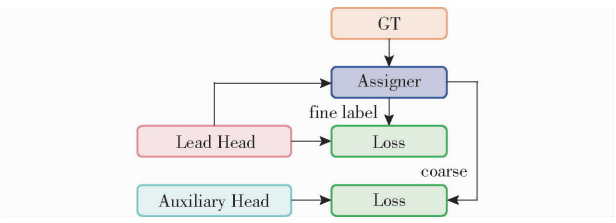


图 8　DyHead seg 引入基于 Auxiliary Head 辅助头示意图

Fig. 8　Schematic of DyHead seg incorporated an Auxiliary Head based on Aug

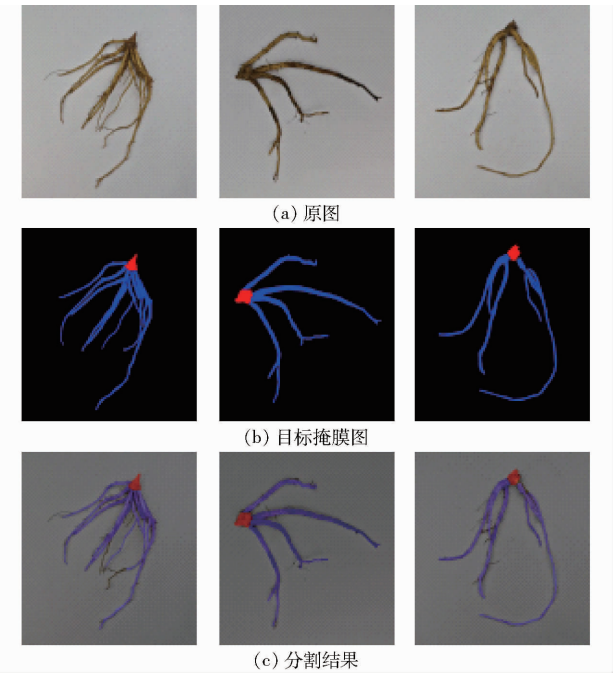


图 9　原图和分割掩膜图

Fig. 9　Original image and split mask image

1.3.6　评价指标

为客观衡量模型实例分割表现，采用精度（Precision，P）、mAP_{50} 和 mAP_{95} 等指标量化评估[33]。其中 P 表示不同目标类别上精度；mAP_{50} 表示重叠度（Intersection over union，IoU）阈值为 0.5 时平均精度；mAP_{95} 表示 IoU 阈值为 0.95 时平均精度。评估类别分为边界框（Bounding Box）和掩膜（Mask），用于评估模型 2 个不同层面能力。其中 Bounding Box 用于指导 Mask 生成区域粗定位参考，Mask 衡量模型对边缘、形状、姿态等目标精细理解。

1.4　关龙胆表型参数

采用根长、根茎粗细度均匀性、含杂率、色度等参数进行等级划分[34－35]。

1.4.1　根长

根长是反映关龙胆根系发育情况的重要表型参数之一，直观表征其生长潜力和营养状态。通过提取实例分割后图像中根茎轮廓及其两端像素坐标，计算端点间像素坐标差 D（像素）和实际根长 L（mm），计算式为

$$D = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \tag{2}$$

$$L = \frac{D}{A} \tag{3}$$

式中　(x_1, y_1)——根茎一端像素坐标

(x_2, y_2)——根茎另一端像素坐标

1.4.2　根茎粗细度均匀性

根茎粗细度影响其水分和养分吸收能力，是评估植物根系健康和生长重要参数。通过分割后图像上部、中部和下部位置根茎宽度，计算其均匀性。根

茎上部、中部和下部分别对应于根茎长度方向上宽度最小值、中间值和最大值位置，通过在该位置测量根茎像素宽度，得出实际宽度值。为减小测量误差，分别计算像素宽度平均值，参照式(3)将像素宽度平均值转换为实际宽度值。计算根部粗细度变异系数(Coefficient of variation，CV)，用于衡量根茎粗细度均匀性，CV 越小表示根部粗细越均匀，计算式为

$$CV = \frac{\sigma}{\mu} \tag{4}$$

式中 σ——根部各截面直径标准差，mm

μ——根部直径平均值，mm

1.4.3 含杂率

关龙胆炮制工艺中，根部药用价值最高，残留茎基药用价值较低，因此需明确根茎中茎基位置，并评估根茎中残留茎基比例。残留茎基比例以含杂率 R(%)表示，通过计算根茎面积 A_{root}(mm^2)与残留茎基面积 A_{stem}(mm^2)比值得出，即

$$R = \frac{A_{stem}}{A_{root}} \times 100\% = \frac{n_{stem}}{n_{root}} \times 100\% \tag{5}$$

式中 n_{stem}——残留茎基像素数量

n_{root}——根茎像素数量

1.4.4 色度

色度包括表面明度和色差，反映关龙胆健康状况及药效成分分布[36]。采用亨特白度公式[22]，将图像中 RGB 颜色值转换为 *XYZ* 颜色空间，从中提取出亮度分量，通过测量图像中每个像素亮度计算根茎区域内所有像素平均亮度，进而得出根茎整体明度。采用 CIE76 色差公式[37]，计算 CIE Lab 颜色空间中亮度和色彩差异来度量颜色变化。

2 结果与分析

2.1 消融试验

为验证优化模块在 FDC - Mamba 模型中有效性，在相同测试集上进行消融试验。试验中，将原始 Mamba - YOLO 模型作为基线模型进行对比，将 Focal Modulation、DyHead seg 和 Context Guided 3 个模块分别集成到基线模型中，以评估各自性能。消融试验对比结果如表 1 所示。

表 1 添加不同模块的消融试验结果

Tab. 1 Ablation experiment results of adding different modules %

DyHead seg	Focal Modulation	Context Guided	Box 类型			Mask 类型		
			P	mAP_{50}	mAP_{95}	P	mAP_{50}	mAP_{95}
×	×	×	84.56	81.96	49.47	82.63	78.21	36.83
√	×	×	87.50	82.58	53.86	83.81	77.88	36.34
√	√	×	91.04	88.29	54.86	85.43	79.70	37.36
√	√	√	91.08	87.05	54.91	87.12	80.89	37.99

注：√表示采用该模块，×表示未采用该模块。

结果表明，所集成的优化模块对模型检测性能具有一定影响，各模块对不同任务精度影响存在差异，具体体现为：当 Mamba - YOLO 模型集成 DyHead seg 模块后，Box 类型 P、mAP_{50}、mAP_{95} 和 Mask 类型 P 分别提升 2.94、0.62、4.39、1.18 个百分点，Mask 类型 mAP_{50}、mAP_{95} 基本持平。这表明 DyHead seg 模块能够有效增强目标检测中位置敏感性和尺度适应性，尤其在精确框回归(mAP_{95})方面效果显著，但对掩码精度提升有限。集成 DyHead seg 和 Focal Modulation 模块后性能提升较大，分别为 6.48、6.63、5.39、2.80、1.49、0.53 个百分点，该组合显著提高了精度。同时 Focal Modulation 模块在强化上下文感知和特征调制方面起到关键作用，进一步提升了模型对复杂场景适应能力。同时引入 3 个模块后，Box 水平小幅提升，但 Mask 水平分别提升 1.69、1.19、0.63 个百分点，表明 Context Guided 模块在引导空间注意力、优化掩码特征方面发挥了积极作用，进一步完善了模型语义分割能力。

2.2 模型对比试验

Mask R - CNN、PointRend、HTC、SOLOv2 等 4 种经典两阶段模型以及 YOLO 系列不同模型、引入注意力引导的超特征融合结构 HYPER 模型、基于 Transformer 的 DETR(RT - DETR)模型与 FDC - Mamba 模型在关龙胆根茎实例分割任务中的模拟对比试验结果如图 10 所示。

经典两阶段模型 Mask R - CNN、PointRend、HTC、SOLOv2 等检测与分割精度较低，说明 DETR 等新一代模型引入了注意力机制与端到端特征编码策略，在本研究场景中具备更强泛化能力，而两阶段结构依赖于 RoI 特征，对微小纹理、细粒度边界识别能力不足[38-39]。同时，相较于 HYPER 的 Box 类型 P、mAP_{50}、mAP_{50-95} 和 Mask 类型 P、mAP_{50}、mAP_{50-95} 分别提升 7.17、3.19、5.71 个百分点和、8.19、、2.52、0.51 个百分点；相较于 RT - DETR 分别提升 8.78、4.05、2.01 个百分点和 7.52、1.29、-1.61 个百分点，表明 FDC - Mamba 不仅在精度指标上优于

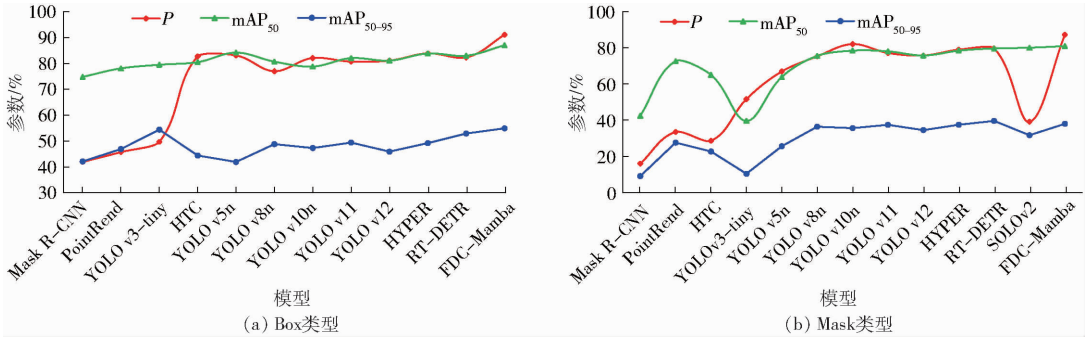


图 10　模型对比结果

Fig. 10　Model comparison results

传统 YOLO 系列与主流改进模型，在特征建模策略上也展现出更强的局部细节提取能力和多尺度语义理解能力，适用于形态复杂且目标密集的根茎类实例分割任务。

2.3　不同特征金字塔性能对比试验

为进一步验证 FDC－Mamba 模型中规避传统 SPPF 操作引入 Focal Modulation 模块的检测性能，对比了常见不同架构特征金字塔部分替换集成 DyHead seg 模块后 Mamba－YOLO 模型 SPPF 部分，结果如图 11 所示。选取 YOLO 系列结构重参数化模块（Re-parameterized batchnorm，RepBN）、Transformer 与 CNN 混合架构中自适应交互机制（Adaptive interaction feature integration，AIFI）和点目标检测中大尺寸选择性卷积注意力模块（Large selective kernel attention，LSKA）进行对比试验。

由图 11 可知，Focal Modulation 在大多数指标上表现最好，尤其在 Box 精度和 Mask 精度上优势明显，其中 Box 类型 P 比 AIFI、RepBN 分别提高 5.75、8.18 个百分点。在 Mask（P）方面，Focal Modulation 比 RepBN 提升幅度达 22.69 个百分点。AIFI 和 LSKA 在部分指标上也具有一定提升，如 AIFI Box 类 mAP_{50-95} 略高于 Focal Modulation 模块 0.17 个百分点，LSKA Mask 类型 mAP_{50-95} 接近 Focal Modulation，但整体稳定性稍差，尤其在 Box 类型指标中存在一定差距。表明 Focal Modulation 不仅在检测任务中保持较高精度，而且在分割任务中也能稳定输出，是关龙胆根茎分割任务中替代传统 SPPF 操作优势方案。

2.4　不同下采样结构性能对比试验

为进一步验证 FDC－Mamba 模型中 Context Guided 模块性能优势，选取 9 个其他类型下采样结构，分别替代集成 DyHead seg 和 Focal Modulation 模块后的 Mamba－YOLO 模型中的 Conv。选取空间重构引导下采样模块（Spatial reconfiguration feature downsampling，SRFD）、自适应下采样模块（Adaptive downsampling，ADown）、内容感知重组下采样模块

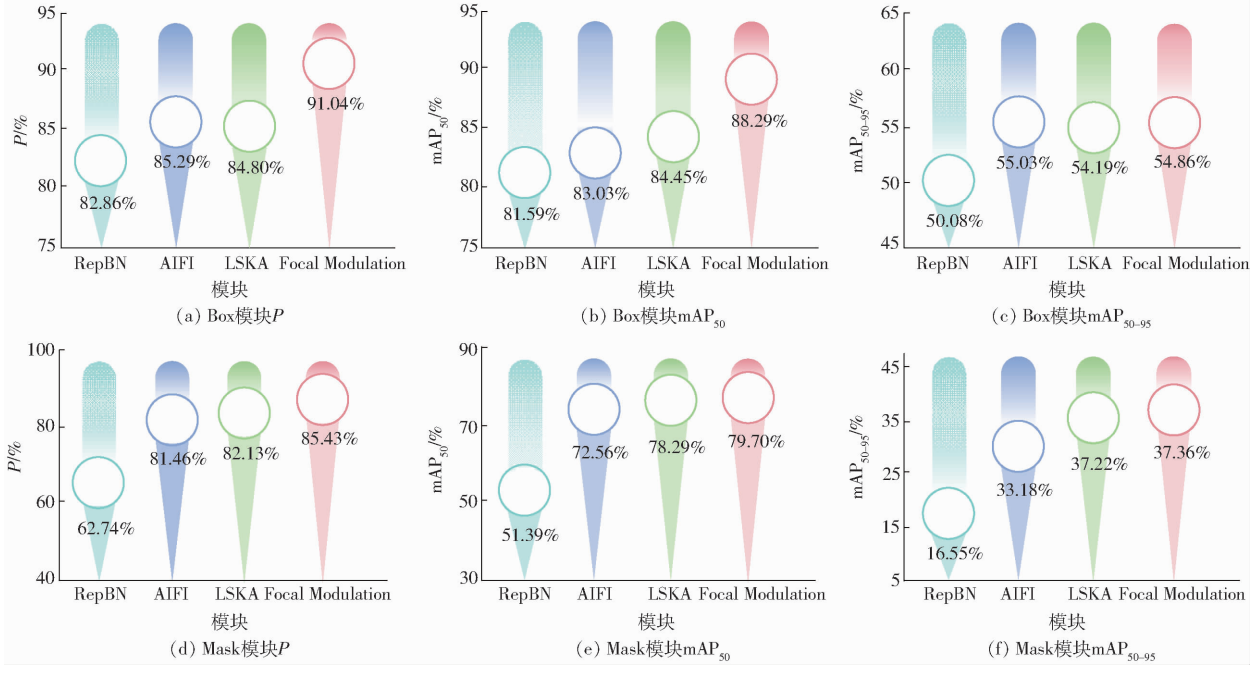


图 11　特征金字塔性能对比

Fig. 11　Feature pyramid performance comparison charts

(Content-aware reassembly of features, CARAFE)、增强型下采样卷积块(Enhanced under-sampling convolution block, EUCB)、Gold-YOLO 下采样模块(Gold YOLO downsampling module, Gold-YOLO)、混合小波下采样模块(Hybrid wavelet downsampling, HWD)、并行分支卷积模块(Parallel split convolution, PSConv)、空间金字塔深度可分卷积模块(Spatial pyramid depthwise convolution, SPDConv)和小波池化模块(Wavelet pooling module, WaveletPool)在相同训练集上进行对比试验,结果如图 12 所示。

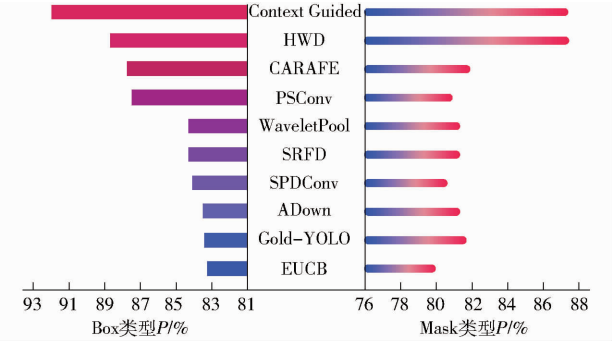


图 12 下采样结构性能试验结果对比

Fig. 12 Results of downsampling structure performance comparison experiment

由图 12 可知,Context Guided 更适配 FDC-Mamba 全局建模与动态调制机制。相较于第 2 位的 HWD,Context Guided 在 Box 类型 P 上提升 3.07 个百分点;在 Mask 类型 P 上,相较于 HWD 略低 0.28 个百分点,但相比第 3 位 CARAFE 提升 5.15 个百分点。此外,相较于常见的 Rep 类结构(如 SRFD、WaveletPool),Box 精度提升超过 6.30 个百分,Mask 精度提升约 5.70 个百分点,差距较为明显。这表明 Context Guided 模块在信息压缩同时能有效保留关键空间信息,并增强语义上下文引导作用,尤其在关龙胆根系边界复杂区域具有更强表达能力,相比其他下采样方式,取得更高质量特征表示,有效提升了整体模型检测与分割性能。

2.5 表型参数试验结果

选取 50 株关龙胆,对应选取其训练集中原始图像,并使用游标卡尺(精度 0.01 mm)、CR-400 型色彩色差计(柯尼卡美能达公司)等工具测量关龙胆根长、粗细度、色度和根茎面积等物理量。人工测量与系统估算间表型参数比较结果如图 13 所示。

根长测量中,人工测量根长平均值为 196.35 mm,与系统估测值相对平均误差为 -0.25%,标准差为 2.52 mm,所用算法略小于人工测量值。主要误差来源为关龙胆根茎呈曲线型,游标卡尺无法弯曲,只能尽量短距离测量,造成了不可避免误差,但相对误差在 2% 范围内,符合预期误差范围。

根茎粗细度测量中,人工测量根部直径平均值

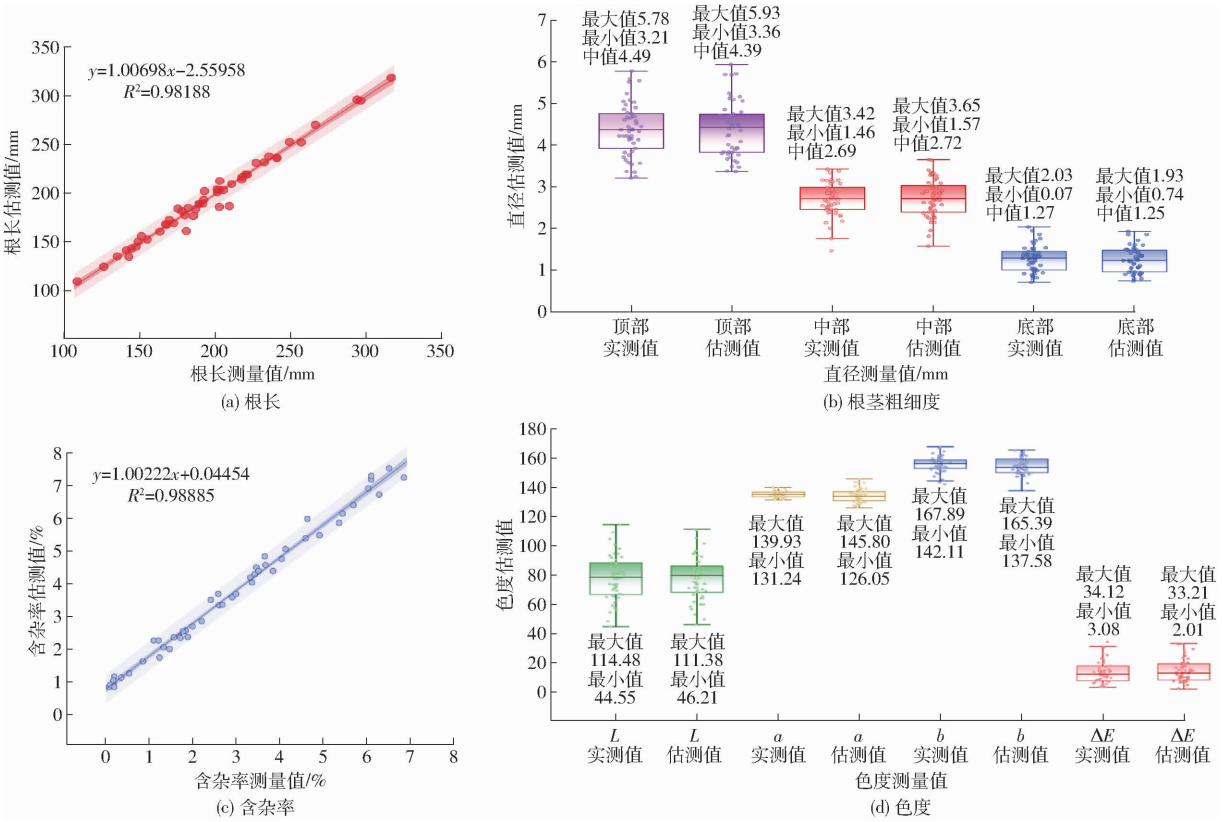


图 13 表型参数对比

Fig. 13 Comparison of phenotypic parameters

为 2.78 mm，最大误差为 0.31 mm（上部），最小误差为 0.32 mm（下部）。该误差产生与实例分割算法局限性有关，即图像分割时对非正上方根系图像与标准俯视图有差异，导致算法计算宽度时提取像素数量相对于关龙胆直径略有浮动。相对误差控制在 2% 范围内，符合预期误差范围，不影响实际炮制加工。

含杂率测量中，人工测量含杂率平均值为 2.91%，误差平均值为 6.34%。该误差来源主要是人工测量误差，使用卡尺测量根茎和残留茎基投影面积时，仅能测量大致轮廓面积，无法精确统计非规则区域。测量相对误差均在 5% 范围内，符合预期误差范围。

色度测量中，人工测量色度平均值为 14.67，系统测量色度平均值为 15.46，最小误差为 －4.92，最大误差为 6.17。模型测量值略高于人工测量值，但偏差很小，且相对误差在 5% 范围内，属于合理偏差，满足精度要求。

3　结论

（1）提出一种融合动态感知和辅助训练的 DyHead seg 模块，有效提升了模型对关龙胆根茎细节表达能力。试验结果表明，集成该模块后，Box 类型 P、mAP_{50}、mAP_{95} 和 Mask 类型 P 分别提升 2.94、0.62、4.39、1.18 个百分点，表现出较好分割效果。相比传统或压缩型检测头结构，该模块在提升检测与分割性能方面具有优异的特征表示并有效保留了关键信息。

（2）引入 Focal Modulation 和 Context Guided 模块增强了图像有效特征，通过多尺度聚焦和联合引导机制促进了建模上下文间交互关系，提高了对目标整体感知能力，尤其是在根茎边缘和细小区域定位方面表现出更高准确度。试验结果表明，FDC－Mamba 模型在 Box 类型和 Mask 类型 P、mAP_{50}、mAP_{95} 分别达 91.08%、87.05%、54.91% 和 87.12%、80.89% 和 37.99%，与基线模型对比，分别提升 6.52、5.09、5.44 个百分点和 4.49、2.68、1.16 个百分点。与同类型实例分割模型相比，在整体性能上均表现出显著优势，能更加准确地针对关龙胆根茎进行更精准分割。

（3）关龙胆表型参数提取方面，根长、粗细度均匀性、含杂率和色度等参数基于分割后图像进行提取，误差范围符合预期，可以满足关龙胆炮制设备部署和检测需求。

参　考　文　献

[1] 廖梅香，薛玉洁. 正交试验优选龙胆炮制工艺的研究[J]. 安徽农业科学，2020，48(8)：182－184.

[2] 刘越，孙晓，杨瑞瑞，等. 六叶龙胆化学成分研究[J]. 中国药学杂志，2020，55(21)：1780－1786.
LIU Yue，SUN Xiao，YANG Ruirui，et al. Chemical constuents from *Getinana hexaphylla* maxim. ex kusnez. [J]. The Chinese-Medicine Magazine of China，2020，55(21)：1780－1786.（in Chinese）

[3] 国家药典委员会. 中华人民共和国药典：一部[M]. 北京：中国医药科技出版社，2020.

[4] 龙胆草全程机械化生产技术规程：DB 2104/T 0037—2024[S]. 2024.

[5] 道地药材 第 69 部分：关龙胆：T/CACM 1020.69—2019[S]. 2019.

[6] 杨启良，陈成，雷炼，等. 基于 TD－BlendMask 的复杂环境三七叶片病害实例分割方法[J]. 农业机械学报，2025，56(4)：375－386.
YANG Qiliang，CHEN Cheng，LEI Lian，et al. TD－BlendMask-based approach for instance segmentation of *Panax notoginseng* leaf diseases in complex environments[J]. Transactions of the Chinese Society for Agricultural Machinery，2025，56(4)：375－386.（in Chinese）

[7] 张勤，庞月生，李彬. 基于实例分割的番茄串视觉定位与采摘姿态估算方法[J]. 农业机械学报，2023，54(10)：205－215.
ZHANG Qin，PANG Yuesheng，LI Bin. Visual positioning and picking pose estimation of tomato clusters based on instance segmentation[J]. Transactions of the Chinese Society for Agricultural Machinery，2023，54(10)：205－215.（in Chinese）

[8] 王福顺，王旺，孙小华，等. 基于改进 YOLO v8n－seg 的羊只实例分割方法[J]. 农业机械学报，2024，55(8)：322－332.
WANG Fushun，WANG Wang，SUN Xiaohua，et al. Sheep instance segmentation method based on improved YOLO v8n－seg [J]. Transactions of the Chinese Society for Agricultural Machinery，2024，55(8)：322－332.（in Chinese）

[9] KHAN Z，LIU Hui，SHEN Yue，et al. Deep learning improved YOLOv8 algorithm：real-time precise instance segmentation of crown region orchard canopies in natural environment[J]. Computers and Electronics in Agriculture，2024，224：109168.

[10] HAO Wangli，REN Chao，FAN Yulong，et al. YSD－BPTrack：a multi-object tracking framework for calves in occluded environments[J]. Smart Agricultural Technology，2025，11：100876.

[11] LI Haidong，HUANG Linsheng，RUAN Chao，et al. A dual-branch neural network for crop disease recognition by integrating frequency domain and spatial domain information[J]. Computers and Electronics in Agriculture，2024，219：108843.

[12] 倪纪鹏，朱立成，董力中，等. 基于 SwinS－YOLACT 的番茄采摘机器人实时实例分割算法研究[J]. 农业机械学报，2024，55(10)：18－30.
NI Jipeng，ZHU Licheng，DONG Lizhong，et al. Real-time instance segmentation algorithm for tomato picking robot based on

SwinS - YOLACT[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(10): 18 - 30. (in Chinese)
[13] MENG Kexin, XU Kexin, CATTANI P, et al. Camouflaged cotton bollworm instance segmentation based on PVT and Mask R - CNN[J]. Computers and Electronics in Agriculture, 2024, 226: 109450.
[14] 翁海勇, 李效彬, 肖康松, 等. 基于 Mask R - CNN 的柑橘主叶脉显微图像实例分割模型[J]. 农业机械学报, 2023, 54(7): 252 - 258, 271.
WENG Haiyong, LI Xiaobin, XIAO Kangsong, et al. Instance segmentation model for microscopic image of citrus main leaf vein based on Mask R - CNN[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(7): 252 - 258, 271. (in Chinese)
[15] 赵露露, 邓寒冰, 周云成, 等. 基于自生成标签的玉米苗期图像实例分割[J]. 农业工程学报, 2023, 39(11): 201 - 211.
ZHAO Lulu, DENG Hanbing, ZHOU Yuncheng, et al. Instance segmentation model of maize seedling images based on automatic generated labels[J]. Transactions of the CSAE, 2023, 39(11): 201 - 211. (in Chinese)
[16] OLAREWAJU M L. Real-time cucurbit fruit detection in greenhouse using improved YOLO series algorithm[J]. Precision Agriculture, 2024, 25: 347 - 359.
[17] 陈允琳, 兰玉彬, 韩鑫, 等. 基于 Kinect V3 传感器的叶菜类作物三维重建与表型参数获取[J]. 农业机械学报, 2025, 56(3): 101 - 110, 197.
CHEN Yunlin, LAN Yubin, HAN Xin, et al. Three-dimensional reconstruction and phenotypic parameters acquisition of leafy vegetables based on Kinect V3 sensor[J]. Transactions of the Chinese Society for Agricultural Machinery, 2025, 56(3): 101 - 110,197. (in Chinese)
[18] 谢立敏, 黄轶, 吴昊宇, 等. 基于 SwinPoinTr 的视角受限下杏鲍菇表型参数测量方法[J]. 农业机械学报, 2025, 56(3): 148 - 157.
XIE Limin, HUANG Yi, WU Haoyu, et al. Phenotypic parameter measurement method for *Pleurotus eryngii* based on SwinPoinTr under limited view conditions[J]. Transactions of the Chinese Society for Agricultural Machinery, 2025, 56(3): 148 - 157. (in Chinese)
[19] 胡程喜, 谭立新, 王文胤, 等. 基于改进 DeepLabV3 + 的轻量化茶叶嫩芽采摘点识别模型[J]. 智慧农业(中英文), 2024, 6(5): 119 - 127.
HU Chengxi, TAN Lixin, WANG Wenyin, et al. Lightweight tea shoot picking point recognition model based on improved DeepLabV3 + [J]. Smart Agriculture, 2024, 6(5): 119 - 127. (in Chinese)
[20] 朱磊, 江伟, 孙伯颜, 等. 基于神经辐射场的苗期作物三维建模和表型参数获取[J]. 农业机械学报, 2024, 55(4): 184 - 192, 230.
ZHU Lei, JIANG Wei, SUN Boyan, et al. Three-dimensional reconstruction and phenotype parameters acquisition of seeding vegetables based on neural radiance fields[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(4): 184 - 192, 230. (in Chinese)
[21] AL RIZA D F, TULSI A A, MOMIN A. Assessing cacao beans fermentation degree with improved YOLOv8 instance segmentation[J]. Computers and Electronics in Agriculture, 2024, 227(1): 109507.
[22] 苗全龙, 周扬, 李建涛, 等. 基于 YOLO v8 - ABSeg 的双孢蘑菇表型参数提取方法[J]. 农业机械学报, 2025, 56(3): 158 - 168.
MIAO Quanlong, ZHOU Yang, LI Jiantao, et al. Extraction method of phenotypic parameters of *Agaricus bisporus* based on YOLO v8 - ABSeg[J]. Transactions of the Chinese Society for Agricultural Machinery, 2025, 56(3): 158 - 168. (in Chinese)
[23] 王纪章, 姚承志, 周静, 等. 基于三维点云的黄瓜叶片分割与表型参数提取方法[J]. 农业机械学报, 2025, 56(3): 354 - 362.
WANG Jizhang, YAO Chengzhi, ZHOU Jing, et al. Cucumber leaf segmentation and phenotype extraction method based on three-dimensional point cloud[J]. Transactions of the Chinese Society for Agricultural Machinery, 2025, 56(3): 354 - 362. (in Chinese)
[24] 朱逢乐, 严霜, 孙霖, 等. 基于深度学习多源数据融合的生菜表型参数估算方法[J]. 农业工程学报, 2022, 38(9): 195 - 204.
ZHU Fengle, YAN Shuang, SUN Lin, et al. Estimation method of lettuce phenotypic parameters using deep learning multi-source data fusion[J]. Transactions of the CSAE, 2022, 38(9): 195 - 204. (in Chinese)
[25] 葛绍娟, 冀横溢, 詹宇, 等. 改进 YOLOv5s 后的轻量化猪只姿态识别方法[J]. 中国农业大学学报, 2025, 30(5): 179 - 189.
GE Shaojuan, JI Hengyi, ZHAN Yu, et al. Lightweight pig posture recognition method after improving YOLOv5s[J]. Journal of China Agricultural University, 2025, 30(5): 179 - 189. (in Chinese)
[26] TURGUT K, DUTAGACI H, ROUSSEAU D. RoseSegNet: an attention-based deep learning architecture for organ segmentation of plants[J]. Biosystems Engineering, 2022, 221: 138 - 153.
[27] SANGAIAH A K, YU Fannong, LIN Yibing, et al. UAV T - YOLO - Rice: an enhanced tiny YOLO-networks for rice leaves diseases detection in paddy agronomy[J]. IEEE Transactions on Network Science and Engineering, 2024, 11(6): 5201 -

5216.
[28] XIA Xue, ZHANG Ning, GUAN Zhibin, et al. PAB－Mamba－YOLO: VSSM assists in YOLO for aggressive behavior detection among weaned piglets[J]. Artificial Intelligence in Agriculture, 2025, 15(1): 52－66.
[29] WANG Zeyu, LI Chen, XU Huiying, et al. Mamba Yolo: a simple baseline for object detection with state space model[C]// Proceedings of AAAI Conference on Artificial Intelligence, 2025: 8205－8213.
[30] YANG Jianwei, LI Chunyuan, DAI Xiyang, et al. Focal modulation networks[C]// NeurIPS, 2022.
[31] TANG Sheng,WU Tianyi, ZHANG Rui, et al. Cgnet: a light-weight context guided network for semantic segmentation[C]// Proceedings of IEEE Transactions on Image Processing, 2020: 1169－1179.
[32] DAI Xiyang, CHEN Yinpeng, XIAO Bin, et al. Dynamic head: unifying object detection heads with attentions[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2021: 7373－7382.
[33] JIANG Lizhi, LI Changying, FU Longsheng. Apple tree architectural trait phenotyping with organ-level instance segmentation from point cloud[J]. Computers and Electronics in Agriculture, 2025, 229: 109708.
[34] 地理标志产品 清原龙胆: DB 2104/T0011—2022[S]. 2022.
[35] 中药材商品规格等级 龙胆: T/CACM 1021.144—2018[S]. 2018.
[36] 刘佳怡, 尹海波, 王丹, 等. 基于“辨状论质”理论的龙胆性状与质量相关性研究[J]. 天然产物研究与开发, 2024, 36(12): 2088－2101, 2144.
LIU Jiayi, YIN Haibo, WANG Dan, et al. Correlation between traits and quality of Gentianae Radix et Rhizoma based on the theory of “quality evaluation through morphological identification”[J]. Natural Product Research and Development, 2024, 36(12): 2088－2101, 2144. (in Chinese)
[37] ZHANG Huang, WU Fengfeng, XU Dan, et al. Effects of milling methods on the properties of glutinous rice flour and sweet dumplings[J]. Journal of Food Science and Technology, 2021, 58: 1848－1857.
[38] SAPKOTA R, AHMED D, KARKEE M. Comparing YOLOv8 and Mask R－CNN for instance segmentation in complex orchard environments[J]. Artificial Intelligence in Agriculture, 2024, 13: 84－99.
[39] OUF N S. Leguminous seeds detection based on convolutional neural networks: comparison of Faster R－CNN and YOLOv4 on a small custom dataset[J]. Artificial Intelligence in Agriculture, 2023, 8: 30－45.

(上接第 478 页)

[34] CHEN H, WANG T, CHEN T, et al. Hyperspectral image classification based on fusing S3－PCA, 2D－SSA and random patch network[J]. Remote Sensing, 2023, 15(13): 3402.
[35] TSAI C F, SUNG Y T. Ensemble feature selection in high dimension, low sample size datasets: parallel and serial combination approaches[J]. Knowledge-Based Systems, 2020, 203: 106097.
[36] 柴成武, 王方琳, 赵鹏, 等. 干旱胁迫对沙蒿叶片水分含量、光合特性及抗氧化酶的影响[J]. 西北林学院学报, 2025, 40(1): 42－50.
CHAI Chengwu, WANG Fanglin, ZHAO Peng, et al. Effects of drought stress on water content, photosynthetic characteristics and antioxidant enzymes of *Artemisia desertorum* leaves[J]. Journal of Northwest Forestry University, 2025, 40(1): 42－50. (in Chinese)
[37] CHENG M, PENUELAS J, MCCABE M F, et al. Combining multi-indicators with machine-learning algorithms for maize yield early prediction at the county-level in China[J]. Agricultural and Forest Meteorology, 2022, 323: 109057.
[38] FAN S M, QIN X Z, JIA Z, et al. ELM-based improved layered ensemble architecture for time series forecasting[J]. IEEE Access, 2019, 7: 97827－97837.
[39] DU H, SONG D, CHEN Z, et al. Prediction model oriented for landslide displacement with step-like curve by applying ensemble empirical mode decomposition and the PSO－ELM method[J]. Journal of Cleaner Production, 2020, 270: 122248.