

改进 YOLOv8 在中药饮片自动识别标注中的研究与应用

孙 兴¹, 李 杨¹, 毛天驰¹, 朱佳音¹, 刘 安², 闫志翻³, 马金刚^{1*}, 郭 丛^{2*}

1. 山东中医药大学医学信息工程学院, 山东 济南 250355

2. 中国中医科学院 中药研究所, 道地药材品质保障与资源持续利用全国重点实验室, 北京 100700

3. 华润三九医药股份有限公司, 广东 深圳 518000

摘要: **目的** 解决中药饮片种类繁多、形态相似, 人工识别耗时费力且易出错的问题。**方法** 构建了包含 201 类中药饮片的数据集, 并提出了一种轻量化改进的 YOLOv8 算法, 具体改进包括在 YOLOv8n 网络中引入 GhostC2f 模块以降低模型参数量, 采用 DySnakeC2f 模块以增强对纤细结构的灵敏度, 替换主干网络的池化层为 SimSPPF 模块以加快推理速度, 并加入坐标注意力 (coordinate attention, CA) 机制以增强对小尺寸目标的特征提取。**结果** 改进后的算法跨阈值平均精度 (50%~95%) 达到 84.16%, 较之前提高了 4.39%, 同时模型参数量减少了约 15%。改进的模型成功部署在电脑客户端和手机 APP 中, 构建了中药饮片自动化识别标注系统。**结论** 改进后的模型能够有效识别中药饮片, 同时系统支持自动数据扩充和升级, 从而为中药饮片的快速、准确识别提供了一种新方法。

关键词: 中药饮片; YOLOv8; 深度学习; 图像识别; 目标检测

中图分类号: TP18; R282.710.3 文献标志码: A 文章编号: 0253-2670(2025)01-0216-11

DOI: 10.7501/j.issn.0253-2670.2025.01.021

Research and application of improved YOLOv8 in automatic recognition and annotation of traditional Chinese medicine decoction pieces

SUN Xing¹, LI Yang¹, MAO Tianchi¹, ZHU Jiayin¹, LIU An², YAN Zhifan³, MA Jingang¹, GUO Cong²

1. College of Medical Information Engineering, Shandong University of Traditional Chinese Medicine, Jinan 250355, China

2. State Key Laboratory for Quality Ensurance and Sustainable Use of Dao-di Herbs, Institute of Chinese Materia Medica, China Academy of Chinese Medical Sciences, Beijing 100700, China

3. China Resources Sanjiu Pharmaceutical Co., Ltd., Shenzhen 518000, China

Abstract: Objective Addressing the issues of the diverse types and similar shapes of traditional Chinese medicine decoction pieces (TCMDPs), which could make manual identification time-consuming, labor-intensive, and prone to errors. **Methods** A dataset containing 201 classes of TCMDPs was constructed, and a lightweight improved YOLOv8 algorithm was proposed. The specific improvements include introducing the GhostC2f module in the YOLOv8n network to reduce model parameters, adopting the DySnakeC2f module to enhance sensitivity to fine structures, replacing the pooling layers of the backbone network with SimSPPF to accelerate inference speed, and incorporating the coordinate attention (CA) mechanism to improve feature extraction for small-sized targets. **Results** The improved algorithm achieved a cross-threshold mean average precision (50%—95%) of 84.16%, representing an increase of 4.39% compared to the previous version, while reducing the model's parameter count by approximately 15%. The enhanced model was successfully deployed on both computer clients and mobile apps, creating an automated recognition and annotation system for TCMDPs. **Conclusion** The improved model effectively identifies TCMDPs, while the system supports automatic data expansion and upgrades, providing a novel approach for rapid and accurate identification of TCMDPs.

Key words: traditional Chinese medicine decoction pieces; YOLOv8; deep learning; image recognition; object detection

中医药作为中国传统医学的重要组成部分, 历史悠久、渊源深远, 其起源可追溯至数千年前的古代中国^[1]。中药饮片是中医临床辨证施治必需的传统武器, 也是中成药的重要原料, 是由天然药材经过加工

收稿日期: 2024-09-15

基金项目: 中国中医科学院科技创新工程项目 (CI2023E002); 国家自然科学基金青年科学基金项目 (32202415); 山东省研究生优质教育教学资源项目 (SDYAL2022041)

作者简介: 孙 兴 (2000—), 男, 硕士研究生, 研究方向为医疗健康大数据。E-mail: xsun88@126.com

*通信作者: 马金刚 (1980—), 男, 硕士, 副教授, 从事医疗健康大数据研究。E-mail: ma_jingang@126.com

郭 丛 (1991—), 女, 硕士, 助理研究员, 从事中药质量评价研究。E-mail: cguo@icmm.ac.cn

制成的^[2]。中药饮片含丰富活性成分，具多种药效，对其准确、高效识别至关重要。在基层药检单位、药材市场、医院中药房乃至多数饮片企业的实际生产过程中，依赖传统人工经验的鉴别方法仍然是主要手段。而传统的人工分类方式需要丰富的专业知识和经验，且存在工作量大、效率低、易出错等问题，亟需新的技术手段来提高识别效率和准确性^[3]。

近年来，随着深度学习目标检测算法的快速发展，有助于传统医药领域更好地整合现代技术，推动中药饮片的自动化识别和应用。在中药饮片识别的研究领域，许多学者已经进行了大量的研究。例如，孙鑫等^[4]在 VGG16 网络上进行了 50 类中药饮片的识别，其结果准确度达到了 70%；谭代庆等^[5]融合残差网络和 Transformer 对 5 组易混淆的中药饮片进行构建，每组准确度均超过 95%；王健庆等^[6]采用 GoogLeNet 模型，其识别准确度为 92%；胡继礼等^[7]通过迁移学习的方法，在 Inception-V3 模型中结合了自适应度量，在 137 种中药饮片数据集上的识别精度达到了 88.3%。

虽然以上研究致力于中药饮片识别，但其往往专注于提升模型精度，却忽略了减少模型参数、提高推理速度的重要性，也未将改进后的模型部署到实际应用。为解决上述问题，并给中药饮片识别领域的未来工作提供参考，本研究主要从以下几个方面展开：首先，使用轻量化的 GhostC2f 模块和增强识别纤细结构的 DySnakeC2f 模块来替换 YOLOv8 网络中的部分 C2f 模块，从而减少模型参数量并提高对饮片纤细结构识别的灵敏度。同时，将主干网络中的池化层替换为 SimSPPF 模块，提高了模型的推理速度。然后，加入坐标注意力 (coordinate attention, CA) 机制，增强了模型对小尺寸饮片的检测准确性。最后，设计了一个中药饮片自动识别标注系统，该系统能够准确识别中药饮片，自动扩充数据并支持系统升级，从而实现中药饮片的全自动识别。本研究旨在将深度学习技术应用于中药饮片领域，为中医药现代化发展贡献新的力量^[8-9]。

1 材料与方法

1.1 实验环境

实验采用 Linux 操作系统，主要使用 Python 开发语言，基于 PyTorch 框架，并通过 CUDA 12.2 进行加速。CPU 选用 Intel® Xeon® Gold 6248R CPU@3.00 GHz, GPU 选用 NVIDIA 的 RTX A6000,

48 Gb 显存。在模型训练过程中，输入图像的尺寸设定为 640×640，采用 AdamW 优化器，训练持续 200 个周期，批量大小为 32，初始学习率为 0.01，并使用原 YOLOv8 的数据增强方法。

1.2 数据集

本研究建立了包含 201 种中药饮片共 3 879 张图像的数据集，数据图像来自中国中医科学院中药研究所门诊药房。为了确保图像的质量和一致性，所有图像均在标准化条件下使用佳能相机 (EOS 80D) 拍摄，背景统一为白色。同时，所有图像均调整为统一分辨率，以确保模型输入的一致性和训练效果。此外，通过去噪、旋转、平移和亮度增强等多种操作对原始图像进行预处理，生成了更加多样化的训练样本，以进一步提升数据集的整体质量和使用效果，丰富模型的训练数据^[10-12]。

数据集按照 7:2:1 的比值划分为训练集、验证集和测试集，涵盖根及茎切段类 (A) 45 种、根及茎切片类 (B) 56 种、果实或种子类 (C) 36 种、皮类 (D) 17 种、叶类 (E) 9 种以及其他类 (F) 38 种共 6 大类中药饮片，数据集具体样例见图 1。数据集中的每张图像包含单个或多个中药饮片样本，图 1 展示了部分样本在一张图像中的多种形态，包括不同的形状、角度和排列方式。这种多样化的图像使得模型能够更全面地学习和识别各种形态的中药饮片，从而提高了模型的泛化能力和实际应用中的表现。为进一步验证模型的泛化性，本研究还引入了从网络爬取的公开中药饮片数据，作为外部测试集进行可视化验证。这些外部数据为模型的性能提供了额外的验证，帮助评估其在不同拍摄条件和环境下的适应性，进一步证明其在处理多样化中药饮片时的稳定性与鲁棒性。

1.3 算法

1.3.1 YOLOv8n YOLOv8 算法是 Ultralytics 公司于 2023 年发布的一种目标检测算法，与之前版本相比，YOLOv8 引入了一些关键改进。图 2 以 YOLOv8n 网络为例，作为 YOLOv8 的轻量化版本，详细描述了网络结构的各个组成部分及其特点。首先，在骨干网络中，YOLOv8n 用 C2f 模块替代了 YOLOv5 的 C3 模块，增强了特征融合能力，并实现模型的轻量化。其次，在颈部网络中，YOLOv8n 去除了部分上采样后的卷积运算以简化结构。在头部网络，YOLOv8n 采用了独立的分类和边界框回归分支，这种解耦头结构提高了检测的准确性和效率。尽



A-根及茎切段类; B-根及茎切片类; C-果实或种子类; D-皮类; E-叶类; F-其他类。
A-root and stem segments; B-root and stem slices; C-fruits or seeds; D-dermatids; E-leaves; F-others.

图1 中药饮片样例

Fig. 1 Examples of traditional Chinese medicine decoction pieces

管 YOLOv8 在许多方面取得了显著进步,但在某些应用场景中仍存在挑战。例如,在移动设备有限的硬件性能下运行 YOLOv8n 时,其使用的 C2f 模块可能会限制设备的运行速度^[13]。此外, YOLOv8n 在动态检测中药饮片时,容易出现误检和漏检问题。为了适应中药饮片检测的实际需求,需对 YOLOv8n 模型进行优化。通过减少模型参数量、增强识别纤细结构的灵敏度、加快推理速度和提升识别精度,可以显著增强模型在实际应用中的表现。改进后的网络模型称为 YOLOv8-TCM,其结构如图 3 所示。

1.3.2 GhostC2f 在实际应用场景中,深度学习模型常需移植到嵌入式设备中使用,尤其后续研究中还需将模型部署到移动端。然而,当硬件算力不足时,深度学习的应用会受到极大限制。为了在 YOLOv8n 的主干网络中降低参数量并加快推理速度,引入了 GhostBottleneck 模块,这一模块可有效

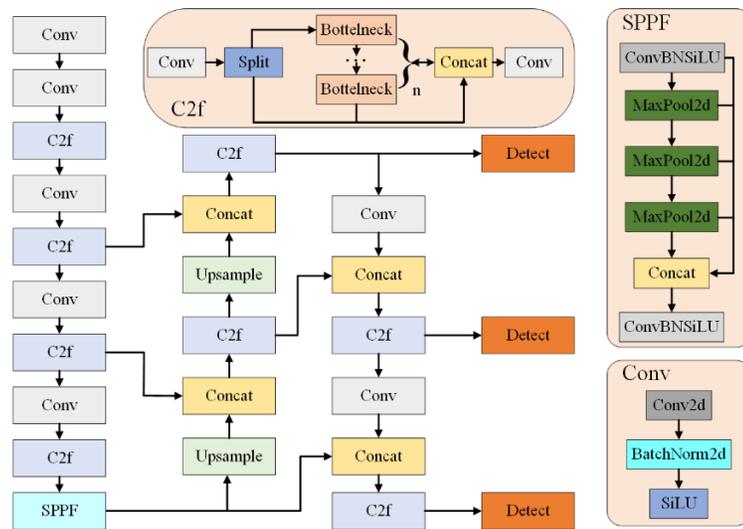


图2 YOLOv8n 网络结构

Fig. 2 YOLOv8n network architecture

减少 C2f 模块中的参数量,满足模型轻量化要求^[14]。

假定输入矩阵为 $A \in R^{h \times w \times c}$, 其中 c 是输入通道数, h 是输入数据的高度, w 是输入数据的宽度。在卷积层中生成 m 个特征图的计算公式如式(1)所示。

$$B = A * g + b \quad (1)$$

其中, $*$ 表示卷积运算, b 是偏差项, $B \in R^{h' \times w' \times m}$ 是输出的 m 个通道的特征图, $g \in R^{c \times k \times k \times m}$ 是该层的卷积滤波器,其卷积核大小为 $k \times k$, h' 和 w' 是输出数据的高度和宽度。普通卷积的计算量按式(2)计算。

$$m \times h' \times w' \times c \times k \times k \quad (2)$$

与普通卷积相比, Ghost 模块中的卷积分 2 步进行,其过程如图 4 所示。第 1 步先经过普通卷积生成少量特征图,然后通过 cheap operation 操作生成冗余特征图,最后将两者结合进行 concat 操作。假设每个基本特征对应着 s 个冗余特征,即 $s \ll c$, $d \cdot d$ 是线性运算的平均核大小, Ghost 模块中的卷积计算公式如下。

由上可知, Ghost 模块中的卷积计算量约为普通卷积的 $1/s$ 。与 ResNet 网络中的基本残差块类似, GhostBottleneck 模块由 2 个堆叠的 Ghost 模块构成。第 1 个 Ghost 模块作为扩展层,负责增加通道

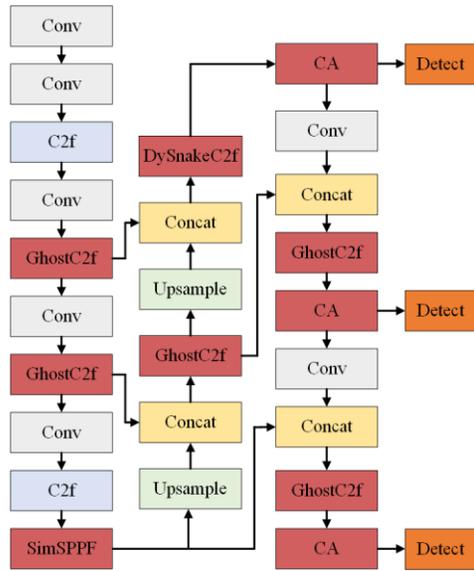


图3 YOLOv8-TCM网络结构
Fig. 3 YOLOv8-TCM network architecture

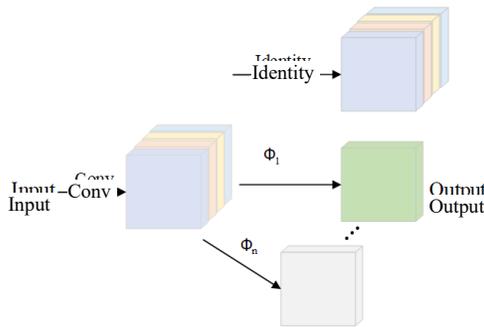


图4 Ghost模块
Fig. 4 Ghost module

$$\frac{m \cdot h' \cdot w' \cdot c \cdot k \cdot k}{\frac{m}{s} \cdot h' \cdot w' \cdot c \cdot k \cdot k + (s-1) \cdot \frac{m}{s} \cdot h' \cdot w' \cdot d \cdot d} \quad (3)$$

$$= \frac{c \cdot k \cdot k}{\frac{c \cdot k \cdot k}{s} + \frac{s-1}{s} \cdot d \cdot d} \approx \frac{s \cdot c}{s+c-1} \approx s$$

数；第2个 Ghost 模块则减少通道数，使之与输入通道数匹配。最后，通过 shortcut 连接这2个 Ghost 模块的输入和输出，实现了特征的有效传递。

GhostC2f如图5所示，其将瓶颈层的 Bottleneck 替换为 GhostBottleneck，从而显著降低了网络的参数计算量和计算成本。公式的推导和比较表明，

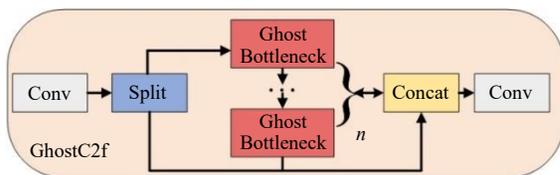


图5 GhostC2f模块
Fig. 5 GhostC2f module

GhostBottleneck 模块能够有效减少卷积层的计算开销，提高推理速度，同时保持较高的检测精度。

1.3.3 DySnakeC2f 对于一些中药饮片（如白薇、莪术、苦地丁和龙胆）常呈现细管或细丝状，模型在识别这些饮片时容易漏检纤细部分，导致识别结果的不完整。对此，本研究设计了一种将动态蛇形卷积（dynamic snake convolution, DSConv）融入 C2f 模块瓶颈层的 DySnakeC2f 模块，使其获得蛇形卷积功能^[15-16]。将该模块置于颈部网络的小目标检测头（P3）之前，能够有效提高模型对纤细目标的识别精度，确保对中药饮片的全面和准确检测。

DSConv 模块允许卷积核在特征学习过程中像蛇一样灵活适应目标结构，其通过引入连续性约束，每个卷积位置的选择参考前1个位置，从而实现了对细微管状结构的有效特征提取。该方法不仅增强了卷积核的形状适应性，还确保了特征感知的连贯性，解决了传统卷积核在小比例管状结构上可能失去感知的问题。在毛细血管的分割领域，该设计已取得显著成就。由于饮片的部分结构与毛细血管形状的较为相似，DySnakeC2f 模块在识别饮片纤细部分时能有效提高精度，使得模型在处理复杂的细丝结构时更加具有针对性和精确性。

1.3.4 SimSPPF 模块 在实际应用中，中药饮片的识别不仅需要高精度，还需要满足实时性要求，以便在繁忙的药房或加工厂中快速处理大量样本。为此，本研究提出了一种改进方案，通过将 YOLOv8 模型中的原始 SPPF 模块替换为更高效的 SimSPPF 结构^[17]来提升系统性能。SimSPPF 结构最初在 YOLOv6 中引入，其主要优点在于有效减少了计算复杂度和处理时间。具体而言，SimSPPF 通过连接3个5×5最大池化层来处理输入数据，生成固定大小的特征图。这些特征图不仅扩大了模型的感受野，还显著提高了特征表示能力，有助于精确识别中药饮片的细微差异。

与传统的 SPPF 模块相比，SimSPPF 采用了 ReLU 激活函数，而 SPPF 模块则使用 SiLU 激活函数。虽然 SiLU 函数在某些情况下表现良好，但其指数计算特性增加了计算复杂度，可能会影响实时性能。相反，ReLU 函数的计算效率更高，能够有效缓解梯度消失问题，加速模型收敛，进一步提升系统的响应速度。这些激活函数的公式如下。

$$\text{ReLU}: f(x) \begin{cases} x, x > 0 \\ 0, x \leq 0 \end{cases} = \max(0, x) \quad (4)$$

$$\text{SiLU}: f(x) = \frac{x^2}{1 + e^{-x}} \quad (5)$$

有实验数据显示, 单个 ConvBNReLU 模块的执行速度比 ConvBNSiLU 模块快 18%^[18]。这不仅证明了 ReLU 函数在计算效率上的优势, 也表明 SimSPPF 结构在提高模型整体性能方面的有效性。

1.3.5 CA 模块 仅依靠基本的检测算法, 仍然难以完全解决中药饮片识别中存在的各种问题, 特别是小目标的检测问题。小目标由于其特征较少, 容易在检测过程中被忽略。为了解决这一问题, 注意力机制被引入到深度学习模型中, 用于增强特征图中重要信息的表达和抑制不相关信息的干扰。CA 注意力机制^[19-20]是一种将位置信息嵌入到通道注意力中的方法, 能够在轻量级网络中实现更大区域的注意力聚焦。CA 注意力机制分为 coordinate 信息嵌入和 CA 生成 2 步, 结构如图 6 所示。

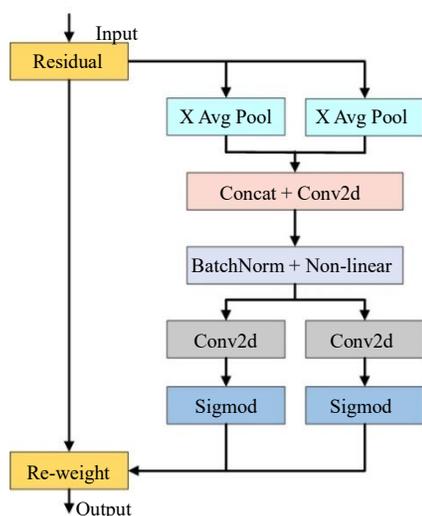


图 6 CA 模块

Fig. 6 CA module

CA 注意力机制简单灵活, 可以轻松地集成到现有模型的多个部分中, 同时不会增加过多的参数和计算负担^[21]。

1.4 模型评价指标

本实验所用到的评价指标有精确率 (precision)、召回率 (recall) 和平均精度均值 (mAP)。其中, TP 代表检测结果中的正确目标数量, FP 代表检测结果中的错误目标数量, FN 代表正确目标中被遗漏的数量。

精确率为预测正确的样本占有所有样本中的比例, 公式如下。

$$\text{精确率} = \text{TP} / (\text{TP} + \text{FP}) \quad (6)$$

召回率为所有正样本中成功检测出正样本的比例, 公式如下。

$$\text{召回率} = \text{TP} / (\text{TP} + \text{FN}) \quad (7)$$

n 个类别的平均精度均值公式如下。

$$\text{mAP} = \frac{1}{n} \sum_{i=1}^n \int_0^1 \text{精确率}(\text{召回率}) d(\text{召回率}) \quad (8)$$

跨阈值平均精度 (mAP@50-95%) 基于 mAP 计算, 将阈值从 50%~95% (每隔 5%) 的多个 mAP 取平均, 用于全面评估模型在不同匹配严格度下的检测性。

2 结果与分析

2.1 实验结果

本实验将改进后的 YOLOv8-TCM 模型与 SSD^[22]、Fast R-CNN^[23] 以及 YOLO 系列的多个基础模型进行了对比, 表 1 列出了模型在中药饮片数据集上的具体指标结果。从整体来看, YOLOv8-TCM 模型在准确度、mAP@50-95% 和参数量方面均优于其他模型, 取得了令人满意的效果。尽管 YOLOv8s 模型的召回率略高于改进模型, 但其参数量约为改进模型的 4 倍, 因此部署到设备的实用性有限。与 YOLOv8n 模型相比, YOLOv8-TCM 模型的 mAP@50-95% 提高了 4.39%, 同时参数量减少了约 15%。表 2 展示了各类饮片的准确率, 其中 75% 的中药饮片的精确率超过 90%。如图 7 所示, 在第 125 个 epoch 时, 上述模型的 mAP@50-95% 达到了 80% 左右, 随后逐渐趋于稳定。从精确率来看, YOLOv8-TCM 在第 80 个 epoch 之后的表现均优于其他模型。从收敛速度来看, YOLOv8s 模型的收敛速度最快, 这主要是由于其通道数参数设置远高于其他模型, 但最终的精度却低于 YOLOv8-TCM 模型, 且收敛曲线在初期波动较大, 稳定性不足。

2.2 结果分析

基于上述结果可知, YOLOv8-TCM 模型在多项性能上优于其他模型, 表明改进后的模型在更广泛的阈值范围内具有更高的精度。此外, YOLOv8-TCM 模型参数量的减少意味着在性能提升的同时, 计算资源的需求也有所降低, 有助于在实际应用中降低硬件成本。

从不同饮片形状来看, 皮类饮片因其通常呈长条状或正方形的规则形状, 表面光滑, 便于特征提取, 因此识别准确度最高。例如, 识别海桐皮的精确率为 95.1%, 白鲜皮为 98.7%, 陈皮、关黄柏和合欢皮则达到了 100%; 根及茎切段类饮片与皮类

表1 模型改进前后的指标对比

Table 1 Comparison of metrics before and after model improvement

模型	精确率/%	召回率/%	mAP@50-95/%	参数量/(×10 ⁶)
SSD	79.73	86.36	76.55	13.3
Fast R-CNN	87.42	89.03	80.56	17.1
YOLOv5n	84.42	88.81	78.18	3.0
YOLOv6n	80.86	85.92	76.81	4.8
YOLOv6s	86.89	90.55	78.95	16.7
YOLOv8n	88.41	89.95	79.77	3.4
YOLOv8s	92.26	96.78	82.72	11.2
YOLOv10n	90.24	92.02	81.85	3.1
YOLOv8-TCM	93.62	96.23	84.16	2.8

表2 改进算法对每种中药饮片的识别精确率

Table 2 Enhanced algorithm's accuracy in recognizing each type of traditional Chinese decoction piece

饮片	精确率/%	饮片	精确率/%	饮片	精确率/%	饮片	精确率/%	饮片	精确率/%
香薷	97.5	炙黄芪	97.4	清风藤	96.9	麦冬	92.7	赤芍	87.6
泽泻	91.6	夏枯草	68.4	山慈菇	96.3	白芍	94.3	黄芪	97.8
红花	96.7	地骨皮	89.3	焦麦芽	92.1	川芎	99.1	丹参	93.4
竹茹	95.3	熟地黄	96.0	金银花	88.1	黄连	97.6	天麻	98.6
紫苑	90.3	姜厚朴	83.3	海桐皮	95.1	麻黄	95.9	钩藤	96.3
龙葵	72.1	土茯苓	97.5	焦山楂	98.2	细辛	92.5	当归	77.8
高良姜	83.4	生山楂	99.5	大血藤	98.5	白英	94.6	防己	100.0
石见穿	80.6	党参片	97.9	鸡血藤	97.0	鳖甲	99.6	三七	95.1
吴茱萸	90.2	北沙参	98.1	墨旱莲	88.2	通草	88.8	砂仁	89.4
菟丝子	94.2	首乌藤	98.5	化橘红	100.0	木香	97.4	橘络	72.8
沙苑子	90.5	玉米须	94.2	醋龟甲	98.5	干姜	92.5	苏木	85.2
青葙子	93.8	盐知母	92.5	威灵仙	73.7	青蒿	96.8	锁阳	86.4
蒲公英	100.0	制远志	97.8	葫芦巴	83.5	牛膝	97.7	桂枝	96.9
络石藤	94.4	木蝴蝶	94.4	肉苁蓉	92.3	木瓜	83.2	枸杞	96.8
苦地丁	92.7	白鲜皮	98.7	板蓝根	98.5	郁金	94.6	香附	94.4
决明子	100.0	白头翁	89.8	巴戟天	89.5	莪术	96.0	人参	98.3
女贞子	70.3	川牛膝	100.0	石菖蒲	93.3	乌药	93.6	白及	100.0
地肤子	90.9	法半夏	98.5	炒山药	92.6	青椒	98.4	槟榔	96.9
炒谷芽	95.3	炒白芍	95.0	关黄柏	100.0	大黄	92.9	藿香	100.0
徐长卿	92.1	何首乌	93.4	山茱萸	81.3	射干	98.9	肉桂	93.4
金荞麦	97.9	金樱子	97.3	蔓荆子	77.5	桑椹	96.5	桑叶	85.8
猫爪草	97.3	炮姜炭	96.9	天花粉	91.6	白芷	100.0	地黄	94.8
棕桐炭	96.1	胆南星	96.7	银柴胡	95.5	杜仲	94.1	灵芝	99.0
薏苡仁	100.0	川楝子	91.3	珍珠母	75.5	茯苓	93.0	柴胡	97.2
浮小麦	95.6	石决明	83.5	麻黄根	91.4	芦根	96.0	水蛭	96.3
王不留行	95.8	合欢皮	100.0	生龙骨	95.5	地龙	87.5	豆蔻	94.1
朱炒苍术	79.4	生前胡	90.5	太子参	88.5	陈皮	100.0	甘草	97.2
麸炒枳壳	95.5	牡丹皮	100.0	冬瓜皮	93.2	香橼	96.8	云参	88.9
蜜桑白皮	96.1	淡竹叶	93.6	绵萆薢	95.8	天冬	88.6	秦艽	75.2
生薏苡仁	95.0	盐杜仲	92.3	生麦芽	87.1	续断	99.2	三菱	94.6
煅瓦楞子	93.2	白茅根	93.6	炒麦芽	92.5	仙茅	87.8	知母	98.6
广金钱草	96.8	野菊花	89.8	焦栀子	97.2	玉竹	99.3	百部	91.2
麸炒白术	93.6	仙鹤草	94.6	苍耳子	94.9	桔梗	86.1	防风	100.0
麸炒枳实	89.5	土鳖虫	100.0	盐黄柏	95.3	苦参	99.5	龙胆	98.5
麸炒椿皮	97.0	海螵蛸	98.4	红景天	92.8	白薇	92.7	滑石	99.6
麸炒山药	95.1	干石斛	97.6	炙甘草	100.0	僵蚕	97.8	海藻	95.1
干益母草	96.4	瓜蒌皮	85.3	清半夏	98.6	乌梅	95.6	茵陈	95.4
醋鸡内金	92.3	浙贝母	98.3	炙百部	96.2	大枣	96.8		
生鸡内金	92.9	伸筋草	85.9	鹿角霜	95.6	菊花	98.6		
制白附子	97.7	生栀子	87.8	百合	98.1	白术	93.6		
煅紫石英	94.9	海风藤	96.2	葛根	98.6	连翘	98.5		

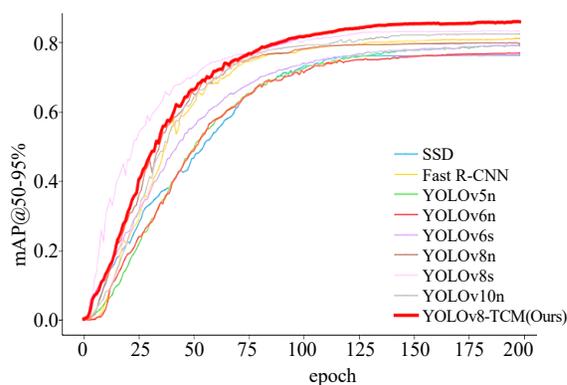


图 7 模型训练过程中 mAP@50-95% 的变化曲线

Fig. 7 Variation curve of mAP@50-95% during model training process

饮片特征相似，往往为条状或块状，其识别香薷、白茅根、龙胆和白薇的精确率均能超过 90%；根及茎切片类饮片因其纹理清晰，常带有明显的切割纹理，易于模型的识别，如干姜的识别精确率为 92.5%，大血藤为 98.5%，海风藤为 96.2%，白芍为 94.3%；果实和种子类饮片通常呈圆形、椭圆形或多面体，颜色鲜艳，因而具有较高的识别准确度，如连翘、金樱子、豆蔻和枸杞的精确率均高于 94%；相比之下，叶类饮片的形状较难统一，通常呈薄片状，边缘可能呈波浪状或锯齿状，颜色范围从绿色到褐色不等，这使得其识别准确度波动较大。例如，识别夏枯草准确率为 68.4%，识别伸筋草和桑叶的精确率约为 85%，而淡竹叶和海藻的精确率分别为 93.6% 和 95.1%。

从整体来看，尽管大多数中药饮片的识别准确率较高，但威灵仙、橘络、女贞子和夏枯草的预测精确率仅在 70% 左右，表现相对较差。此现象主要归因于这些饮片常呈现碎屑状、重叠严重，并伴随明显阴影，导致形态极不规则。例如，威灵仙和女贞子的碎屑结构及阴影效应显著降低了模型的识别能力，其重叠程度约达到了 50%，使得特征混淆与视觉干扰现象相对严重。此外，样本数量和类型的匮乏（如女贞子仅有 8 个样本，且摆放的类型较少）进一步影响了模型的学习效果。这一发现表明，模型在处理碎屑样本或重叠较严重的饮片时确实存在一定局限性。因此，后续研究将重点考虑如何增加样本数量、优化数据预处理及增强模型对特征不规则性适应能力，以提高识别精度。

2.3 可视化分析

为了更清晰直观地检验改进模型的性能，本研

究采用可视化方法展示了外部数据和特定模块的有效性。图 8 展示了常规环境下部分中药饮片的识别结果，包括枸杞、连翘、牡丹皮、白芷、大血藤和黄连等，该数据来源于公开网络数据的爬取。图中仅牡丹皮和黄连的识别存在轻微漏检，这主要是图像清晰度不足和光照角度问题所导致。而其余饮片的识别结果均较为准确，没有出现错检或漏检的情况，这进一步验证了模型在实际应用中具有较强的泛化性。

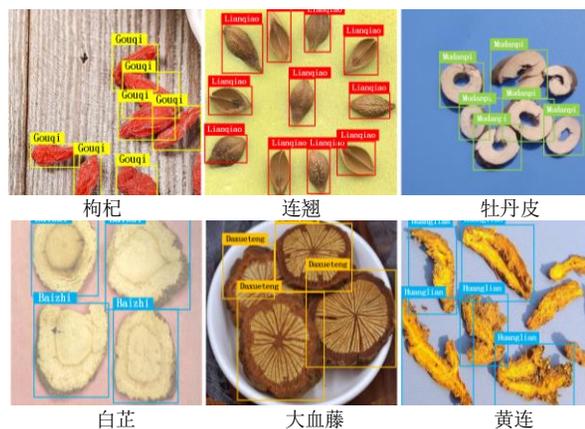


图 8 外部数据的可视化验证

Fig. 8 Visualization validation of external data

图 9 展示了特定模块加入前后的效果对比，其中左侧两列为 CA 模块，右侧两列为 DySnakeC2f 模块，上半部分为加入模块前的效果，下半部分为加入模块后的优化效果。由图 9 可知，改进前白头翁和党参片的小目标识别存在漏检现象，而在加入 CA 模块后，漏检情况得到了明显改善。此外，在未加入 DySnakeC2f 模块时，模型对莪术和生山楂的细管或细丝状结构识别不够灵敏，导致识别结果不完整，改进后效果得到了显著提升。

3 系统设计与实现

3.1 系统设计

3.1.1 设计目标 中药饮片识别系统的总体设计目标是在实际应用中实现中药饮片的准确识别、自动扩充数据与系统升级，具体可分为以下几点。

(1) 利用改进的 YOLOv8 模型生成预测坐标和标签文件，并结合 OpenCV 绘制目标检测框，以实现一键自动检测和标注中药饮片图像^[24]。

(2) 构建改进 YOLOv8 的识别模型，将其部署在手机 APP 和 Windows 客户端^[25-26]。在此系统中，用户能够实时对中药饮片的视频和图像进行检测。此外，用户在客户端还能对标注不准确的图像发起

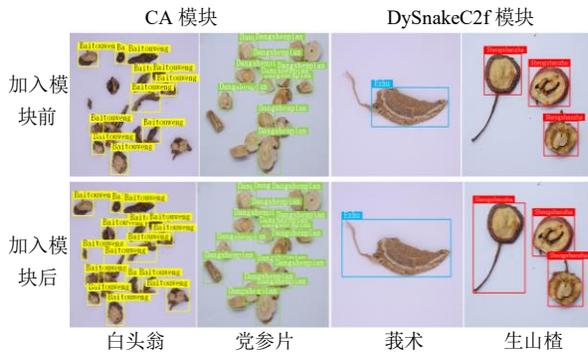


图 9 加入特定模块的前后对比

Fig. 9 Comparison before and after addition of specific modules

校正请求，专家在收到请求后会进行校正，并将结果返回给用户。

(3) 在系统中整合数据集自动划分与标注、更新模型和自动扩充数据等功能，旨在简化人工升级系统的操作流程。

3.1.2 总体架构 本系统分为用户层、应用层、逻辑层和数据层，具体架构如图 10 所示。用户层有 3 类角色，分别为用户、领域专家和系统管理员；应用层包括中药饮片识别系统的 APP 和客户端；逻辑层负责完成图像的检测和标注、数据集处理、模型训练等任务；数据层则存放模型文件、中药饮片图像和标签等数据。

3.1.3 流程设计 系统的流程设计如图 11 所示，具体步骤如下。

(1) 利用中药饮片识别系统的数据自动标注功

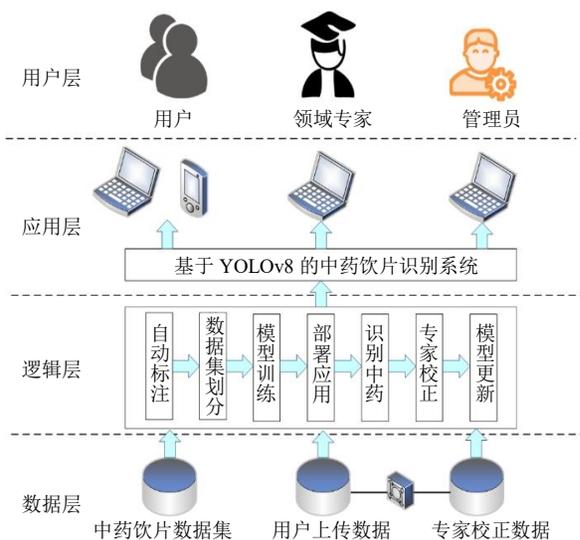


图 10 系统总体架构

Fig. 10 Overall architecture of system

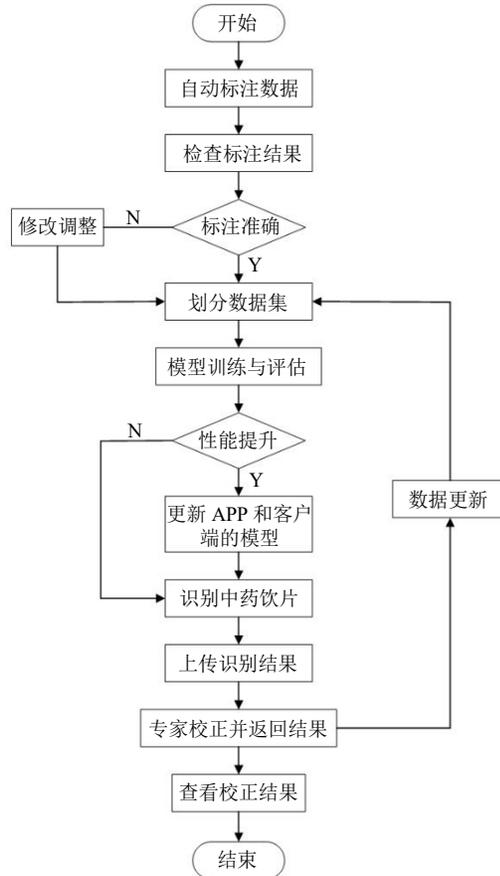


图 11 系统运行流程图

Fig. 11 Operation flowchart of system

能实现对中药饮片数据集的自动标注，为模型后续的训练提供依据。

(2) 检查中药饮片图像的标注是否准确，若不准确则要进行修改。

(3) 待数据集标注完成后，使用系统的数据集自动划分，按照设定比例将数据集划分为训练集、验证集和测试集。

(4) 在模型训练与评估中，可直接选择默认参数自动训练，也可以自己对参数做出灵活的调整。

(5) 若模型的性能得到提升，则可导出模型到手机 APP 和客户端，以供用户更精准地对中药饮片进行识别检测。

(6) 用户在识别中药饮片的过程中，若出现识别错误或新的中药饮片种类可将数据和情况通过服务器通信反馈给领域专家，等待专家的处理。

(7) 专家收到请求后审查中药饮片的识别结果并进行适当校正，将更新后的数据上传至服务器，实现数据的自动扩充，并确保数据得到实时更新。

(8) 用户查看校正结果，并根据实际需求决定是否通知管理员对系统进行升级。

(9) 系统的数据集会不断地扩充, 因此本系统需要不断地更新升级以更广泛准确地识别中药饮片。管理员可根据具体情况执行步骤(3)(4)(5)更新模型。

3.2 系统实现

本系统的主要功能是通过客户端实现的, 为了方便用户使用, 本系统还支持手机 APP 识别中药饮片的图像和视频。

3.2.1 手机 APP 为了方便用户的使用, 本研究将中药饮片的识别模型部署到 Android 端, 如图 12 所示, 其能够支持中药饮片的视频和图像检测。

同时, 客户端的模型更新功能可将新的模型导出至 APP, 以确保应用可以随时更新。

3.2.2 系统客户端 系统客户端可供用户、领域专家和管理员使用, 如图 13 所示, 其包含登录、注册、管理员主界面、自动标注、自动划分数据集、专家校正、模型训练和评估以及中药饮片识别 8 个界面。

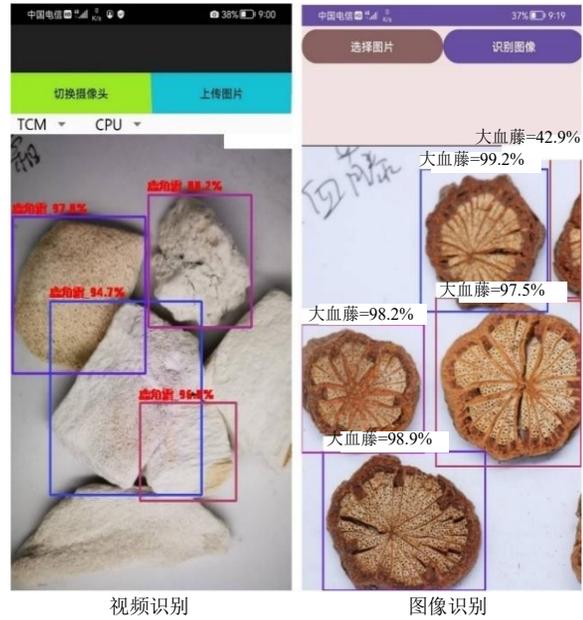


图 12 系统 Android 端操作示例

Fig. 12 Operation example of system Android client

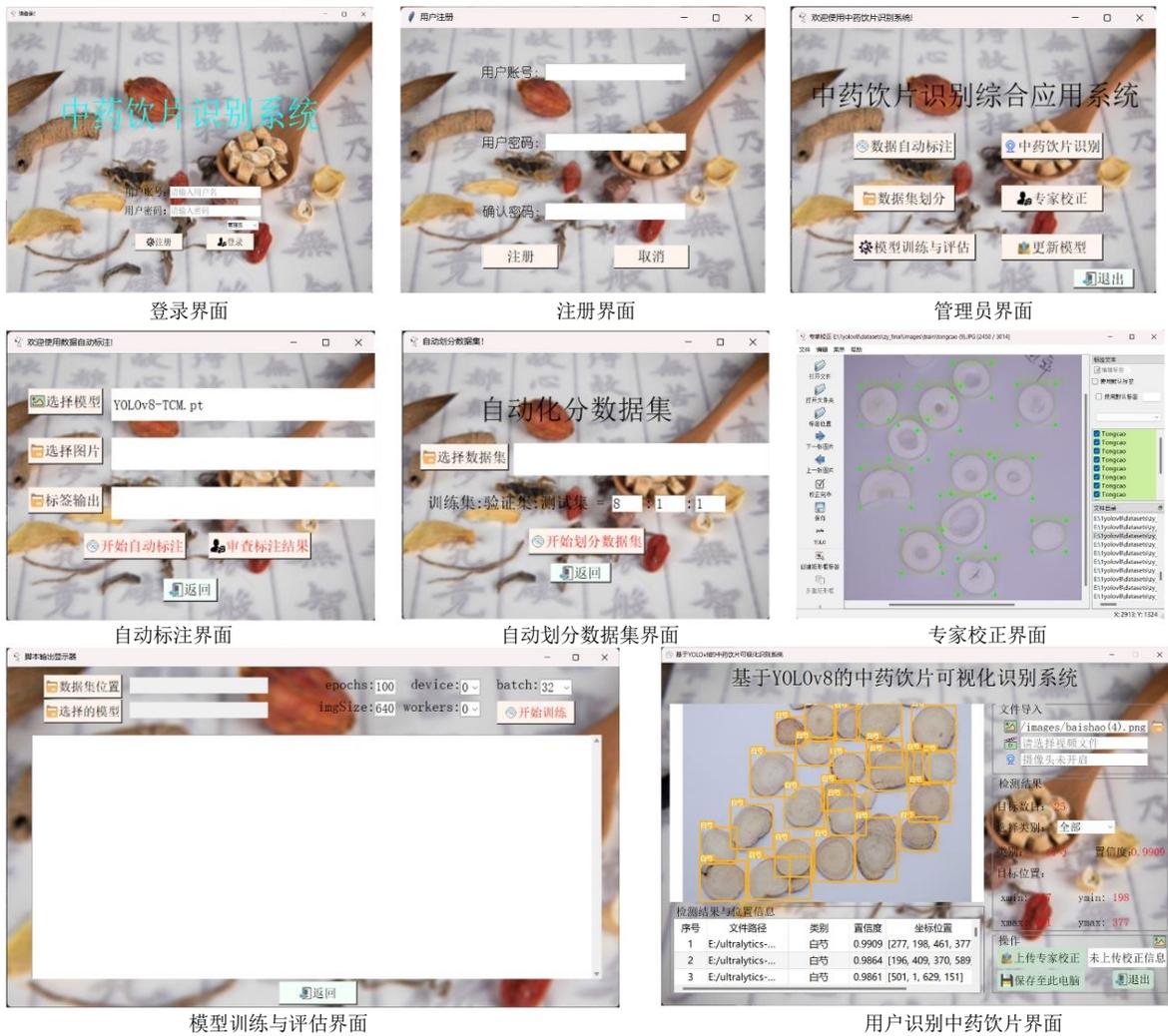


图 13 系统客户端操作界面

Fig. 13 Operation interface of system client

在登陆系统时,选择不同的角色将进入不一样的操作界面,其中管理员的权限最高,拥有系统所有功能的操作权限。管理员可先使用“自动标注”,将选择的中药饮片数据集进行初步标注,随后检查标注结果是否准确。然后,管理员通过“自动划分数据集”得到所需的数据集,再通过“模型训练与评估”进行模型的一键自动训练。若模型效果比之前好,则通过“更新模型”一键替换掉系统的模型,从而完成系统的更新。用户主要使用图 13 的“用户识别中药饮片界面”,其可以导入图片、视频或者打开本地摄像头来识别中药饮片。用户可以点击按钮“上传专家校正”将会通过服务器上传此次识别结果到专家系统,同时右侧的状态“未上传校正信息”将更新为“等待校正”,当专家处理完毕时,状态则再次更新至“矫正完毕”,用户可在系统查看校正后的数据。专家使用图 13 的“专家校正界面”,其可以对中药饮片图像的边界框、名称等进行灵活的修改校正。该模块主要嵌入了 Labellmg 工具^[27],这是一款开源的图像标注工具,其源代码和具体的使用文档可以在网址 <https://github.com/tzutalin/labellmg> 找到。此外,为了满足本系统的需求,本研究对 Labellmg 做出了以下改进:(1)解决了 Labellmg 只能一次性标注图像的问题,使其支持二次标注功能,用户可以随时打开数据集继续上次的标注工作。(2)增加了同步上传功能,标注结果可以实时保存到服务器,确保数据的及时更新。(3)在界面左侧新增了“矫正完毕”按钮,专家点击后将收到图像校正完成的通知。

3.3 系统测试

用户在使用客户端检测中药饮片时,可以选择将数据保存至本地或服务器中,从而实时更新数据。在系统测试时查看服务器可以发现,用户上传和专家校正时都会自动更新服务器上的数据。经测试,系统的自动标注、模型更新等功能一切正常。

4 结论

本研究建立了 1 个包含 201 种中药饮片的数据库,根据中药饮片的特点及便于移动端部署的需求,对 YOLOv8 算法进行了轻量化改进:采用 GhostC2f 模块和 DySnakeC2f 模块替换部分 C2f 模块,以实现模型轻量化并增强对管状或细丝状结构的识别灵敏度;使用 SimSPPF 模块取代主干网络的池化层,以提高模型推理速度;并引入 CA 注意力机制以提高模型对中药饮片的识别精度。另外,本

研究设计了中药饮片自动识别标注系统,该系统支持客户端和移动端应用,结合改进的 YOLOv8 模型,能够准确识别中药饮片图像和视频,同时支持系统的自动升级。本研究旨在将人工智能应用于中药饮片识别领域,减少人工识别工作量,进而推动中医药的发展^[28]。

尽管本研究取得了一定成果,但仍存在提升空间。未来工作可从以下几点开展:首先,数据集对模型识别精度至关重要,需不断提高中药饮片图像的质量和数量,建立一个高质量的大型中药饮片数据集;其次,模型的泛化性在面对复杂的实际情况(如遮挡、阴影和模糊等)时需仍能准确识别中药饮片^[29-32],故提高泛化性是一项十分迫切的研究;最后,进一步减少模型参数量,可以满足实际应用中的快速识别需求,并便于在小型设备上部署。

利益冲突 所有作者均声明不存在利益冲突

参考文献

- [1] 任程, 谢震宇, 刘淑仪, 等. 中医文化传播现状与调研 [J]. 中国中医药现代远程教育, 2021, 19(20): 49-52.
- [2] 张雅娟, 姜云耀. 中药材质量的研究方法和进展 [J]. 中医药学报, 2021, 49(7): 106-112.
- [3] Guo P Q, Brand E, Zhao Z Z. Chinese medicinal processing: A characteristic aspect of the ethnopharmacology of traditional Chinese medicine [J]. *Ethnopharmacology*, 2015: 303-316.
- [4] 孙鑫, 钱会南. 基于深度卷积网络的中药饮片图像识别 [J]. 世界科学技术—中医药现代化, 2017, 19(2): 218-222.
- [5] 谭代庆, 肖志鸿, 吴浩忠, 等. 融合改进残差网络和 Transformer 的易混淆中药饮片识别研究 [J]. 中国数字医学, 2023, 18(6): 42-50.
- [6] 王健庆, 戴恺, 李子柔. 基于深度学习的中药饮片图像识别研究 [J]. 时珍国医国药, 2020, 31(12): 2930-2933.
- [7] 胡继礼, 王永康, 阚红星. 基于深度迁移学习的中药饮片识别研究 [J]. 新乡学院学报, 2019, 36(3): 62-67.
- [8] Holzinger A, Langs G, Denk H, et al. Causability and explainability of artificial intelligence in medicine [J]. *Wiley Interdiscip Rev Data Min Knowl Discov*, 2019, 9(4): e1312.
- [9] Zhang P, Zhang D F, Zhou W A, et al. Network pharmacology: Towards the artificial intelligence-based precision traditional Chinese medicine [J]. *Brief Bioinform*, 2023, 25(1): bbad518.
- [10] Xu M L, Yoon S, Fuentes A, et al. A comprehensive survey of image augmentation techniques for deep learning [J]. *Pattern Recognit*, 2023, 137: 109347.
- [11] Zhang H Z, Huang M, Yang J J, et al. A data preprocessing method for automatic modulation classification based on CNN [J]. *IEEE Commun Lett*, 2021, 25(4): 1206-1210.

- [12] Chlap P, Min H, Vandenberg N, *et al.* A review of medical image data augmentation techniques for deep learning applications [J]. *J Med Imaging Radiat Oncol*, 2021, 65(5): 545-563.
- [13] Talib M, Al-Noori A H Y, Suad J. YOLOv8-CAB: Improved YOLOv8 for Real-time object detection [J]. *Karbala Int J Mod Sci*, 2024, 10(1): 5.
- [14] Han K, Wang Y H, Tian Q, *et al.* GhostNet: More features from cheap operations [A] // 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) [C]. Seattle: IEEE, 2020: 1577-1586.
- [15] Qi Y L, He Y T, Qi X M, *et al.* Dynamic snake convolution based on topological geometric constraints for tubular structure segmentation [A] // 2023 IEEE/CVF International Conference on Computer Vision (ICCV) [C]. Paris, France. IEEE, 2023: 6047-6056.
- [16] Niu Y N, Fan S H, Cheng X, *et al.* Road crack detection by combining dynamic snake convolution and attention mechanism [J]. *Appl Sci*, 2024, 14(18): 8100.
- [17] Norkobil Saydirasulovich S, Abdusalomov A, Jamil M K, *et al.* A YOLOv6-based improved fire detection approach for smart city environments [J]. *Sensors*, 2023, 23(6): 3161.
- [18] Wang X Q, Gao H B, Jia Z M, *et al.* BL-YOLOv8: An improved road defect detection model based on YOLOv8 [J]. *Sensors*, 2023, 23(20): 8361.
- [19] Hou Q B, Zhou D Q, Feng J S. Coordinate attention for efficient mobile network design [A] // 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) [C]. Nashville: IEEE, 2021: 13708-13717.
- [20] Xie C, Zhu H Y, Fei Y Q. Deep coordinate attention network for single image super-resolution [J]. *IET Image Process*, 2022, 16(1): 273-284.
- [21] Wen G, Li S B, Liu F C, *et al.* YOLOv5s-CA: A modified YOLOv5s network with coordinate attention for underwater target detection [J]. *Sensors*, 2023, 23(7): 3367.
- [22] Liu W, Anguelov D, Erhan D, *et al.* Ssd: Single shot multibox detector [A] // Computer Vision—ECCV 2016: 14th European Conference Proceedings, Part I 14 [C]. Amsterdam: Springer International Publishing, 2016: 21-37.
- [23] Ren S Q, He K M, Girshick R, *et al.* Faster R-CNN: Towards real-time object detection with region proposal networks [J]. *IEEE Trans Pattern Anal Mach Intell*, 2017, 39(6): 1137-1149.
- [24] Mishra S, Verma M V, Akhtar D N, *et al.* An intelligent motion detection using OpenCV [J]. *Int J Sci Res Sci Eng Technol*, 2022: 51-63.
- [25] Hyder R T S H, Siddiqui S S S A, Jonnalagedda M, *et al.* Android-based malaria detection using deep learning [A] // International Conference on Data Science and Applications [C]. Singapore: Springer Nature Singapore, 2023: 361-374.
- [26] Wajahat A, He J S, Zhu N F, *et al.* An adaptive semi-supervised deep learning-based framework for the detection of Android malware [J]. *J Intell Fuzzy Syst*, 2023, 45(3): 5141-5157.
- [27] Tzutalin D. LabelImg [J]. *GitHub Repository*, 2015, 6: 4.
- [28] Miao J F, Huang Y N, Wang Z S, *et al.* Image recognition of traditional Chinese medicine based on deep learning [J]. *Front Bioeng Biotechnol*, 2023, 11: 1199803.
- [29] Zhang C Y, Bengio S, Hardt M, *et al.* Understanding deep learning (still) requires rethinking generalization [J]. *Commun ACM*, 2021, 64(3): 107-115.
- [30] Chen W Y, Gong X Y, Wu J R, *et al.* Understanding and accelerating neural architecture search with training-free and theory-grounded metrics [J]. *IEEE Trans Pattern Anal Mach Intell*, 2024, 46(2): 749-763.
- [31] Azizi S, Culp L, Freyberg J, *et al.* Robust and data-efficient generalization of self-supervised machine learning for diagnostic imaging [J]. *Nat Biomed Eng*, 2023, 7(6): 756-779.
- [32] Su Z X, Yao K, Yang X, *et al.* Rethinking data augmentation for single-source domain generalization in medical image segmentation [J]. *Proc AAAI Conf Artif Intell*, 2023, 37(2): 2366-2374.

[责任编辑 潘明佳]