

基于卷积神经网络的中药材自动识别模型创建

林 兴¹

LIN Xing

摘 要

针对传统中药材人工识别效率低、主观性强、误差大,以及现有深度学习模型泛化能力不足、细粒度特征提取薄弱、对标注样本依赖度高的问题,文章提出一种注意力增强型多尺度卷积神经网络中药材自动识别模型。以150种常见中药材为研究对象,基于ResNet-50优化基础网络结构,嵌入SE通道注意力与空间注意力双机制,结合多尺度特征融合技术,实现中药材宏观形态与微观纹理特征的协同捕捉。融合迁移学习与半监督迭代训练策略,减少标注样本需求,提升模型泛化能力。实验结果表明,该模型在测试集上识别准确率达96.7%,在易混淆品种识别任务中表现优异,且参数量更优、部署灵活性强,具备良好的复杂场景鲁棒性,为中药材质量检测、市场监管及产业标准化发展提供高效技术支撑。

关键词

卷积神经网络; 中药材识别; 注意力机制; 多尺度特征融合; 迁移学习

doi: 10.3969/j.issn.1672-9528.2026.03.002

0 引言

中药材作为中医药产业的核心资源,其真伪与品质直接关系到临床疗效和用药安全。当前中药材识别仍以人工鉴别为主,受鉴别者经验、主观判断及外界环境影响,存在效率低、误判率高、一致性差等问题,部分易混淆品种鉴别准确率很低,严重制约了中医药产业的标准化发展^[1]。

随着人工智能与计算机视觉技术的飞速发展,深度学习图像识别技术为中药材自动鉴别提供了新路径,逐步实现了从传统机器学习到端到端智能识别的跨越。但现有研究在实际应用中仍面临三大核心挑战:一是中药材样本形态多样性强,同一品种因产地、加工方式、生长周期不同外观差异显著,导致模型泛化难度大;二是现有数据集多聚焦于常见品种,对稀有中药材及极端加工状态的样本覆盖不足,且标注规范性参差不齐,难以支撑模型深度训练;三是传统卷积神经网络(convolutional neural network, CNN)对中药材纹理、断面等细粒度特征的提取能力有限,单一注意力机制或特征融合方式难以全面捕捉宏观形态与微观纹理的协同特征,同时模型训练对标注样本依赖度高,在数据稀缺场景下性能衰减明显。

为此,本文提出一种注意力增强型多尺度CNN模型,通过构建多维度标准化数据集、优化双注意力机制与多尺度特征融合结构、融合迁移学习与半监督训练策略,针对性解

决上述问题,实现中药材高效精准识别,为中医药现代化、标准化发展提供技术支撑。

1 相关工作

中药材及饮片自动识别是中医药产业智能化发展的核心,深度学习技术推动相关研究在数据集构建、模型优化及应用落地方面取得显著进展,核心围绕数据支撑、精度效率优化、实用化开发三大方向展开,为本文研究提供了重要基础。

深度学习为中药材及饮片自动识别提供了全新路径,相关研究已在模型改进、数据集构建、应用开发及领域现状梳理等方面形成成果。

孙兴等^[1]针对中药饮片识别标注难题,构建含201类饮片的数据集,对YOLOv8n网络轻量化改进。通过引入GhostC2f模块降参、DySnakeC2f模块提升纤细结构灵敏度、SimSPPF模块加速推理,搭配CA注意力机制强化小目标提取,模型跨阈值平均精度达84.16%,参数量减少15%,并部署于电脑与手机端,搭建起可自动扩充升级的识别标注系统。

王甘红等^[2]聚焦实用需求,基于CNN开发多品类中药材识别模型及应用。历时4年采集含163种中药材的数据集,累计276 767张图像用于模型训练验证,对比5种CNN模型后,确定EfficientNetB0为最优模型,其验证集准确率达99.0%、AUC值0.994 2,测试集性能优异。基于该模型,采用PyQt5技术开发电脑端程序,实现中药材便携识别,适配基层与临床场景。

韩勇等^[3]围绕饮片精准识别,构建含50种饮片、15 622

1. 闽南理工学院信息工程学院 福建泉州 362700

[基金项目] 福建省中青年教育科研项目(JAT251208)

张图像的数据集，基于 Keras 框架优化 DenseNet-201 网络。通过稠密块与过渡层交替结构、全局最大池化层及丢弃法抑过拟合，模型在 43 种饮片上识别率超 90%，最高达 95.21%，验证了深度学习的有效性。胡晓东^[4]在硕士论文中，针对传统算法精度不足问题，优化深度学习网络与特征提取机制，提升饮片识别准确性，为算法创新提供参考。

王超超等^[5]从宏观视角梳理人工智能在中药材检测中的应用，总结其在分类、鉴别、成分测量等方面的应用场景，指出数据标准化不足、共享机制缺失、多模态融合薄弱等问题，提出智能化、精准化、快速化的未来方向。

综上，现有研究奠定了技术基础，但在细粒度提取、多场景适配等方面仍有优化空间，这也是本文研究核心。

2 中药材自动识别模型设计

2.1 总体框架

本文提出的中药材自动识别模型总体框架包括数据集构建与预处理、注意力增强型多尺度 CNN 模型设计、模型训练与优化三大核心模块，框架如图 1 所示。首先通过标准化流程构建多维度中药材图像数据集，解决样本单一、标注不规范问题；然后基于 ResNet-50 优化模型结构，嵌入注意力模块与多尺度特征融合机制，强化细粒度特征提取能力；最后结合迁移学习与半监督学习策略，减少标注样本依赖，提升模型泛化能力。

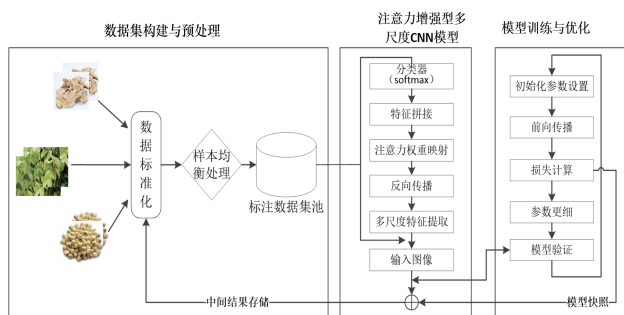


图 1 总体框架

2.2 多维度数据集构建与预处理

(1) 样本采集与标准化标注

选取 150 种常见中药材，覆盖根、茎、叶、花、果实等不同形态类型，从云南文山、宁夏中宁等 20 个主产区收集样本，涵盖鲜品、饮片、炮制品等多种加工方式，确保样本多样性。采用高清工业相机（分辨率 $\geq 2048 \text{ px} \times 1536 \text{ px}$ ）与标准化光照拍摄箱采集图像，严格控制拍摄距离、光照强度等参数，避免外界环境干扰。使用 LabelImg 工具进行标准化标注，标注内容包括中药材品种、产地、加工方式等信息，建立数据集元数据库，确保标注一致性与可追溯性。

(2) 图像预处理流程

设计去噪、光照校正、区域提取三级预处理流程，采用高斯滤波，卷积核大小 3×3 去除图像噪声，保留中药材纹理细节。其卷积核通过该滤波可保留中药材纹理细节，抑制高频噪声干扰。

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (1)$$

式中： $\sigma=1.0$ 为高斯核标准差； $x, y \in \{-1, 0, 1\}$ 对应 3×3 卷积核的像素坐标。

通过直方图均衡化算法解决光照不均问题，提升图像对比度，用公式表示为：

$$H(v) = \sum_{u=0}^v \frac{n_u}{N} \quad (2)$$

式中： $H(v)$ 是均衡化后灰度值 v 对应的映射函数； n_u 是原始图像中灰度值为 u 的像素数量； N 是图像的总像素数，求和范围是 $u=0$ 从灰度值到 $u=v$ 。

运用自适应阈值分割技术，基于中药材与背景的灰度差异实现目标区域自动提取，阈值排除无关背景干扰，突出核心特征区域。

$$T(x, y) = \text{mean}(I(x-r:x+r, y-r:y+r)) - (C) \quad (3)$$

式中： $r=5$ 为局部邻域半径； $C=2$ 为补偿系数； $\text{mean}(\cdot)$ 为邻域均值计算函数。

(3) 组合式数据增强策略

为解决样本数量不足与模型过拟合问题，采用基础增强、场景模拟增强的组合策略。基础增强阶段，通过随机翻转、旋转（ $-30^\circ \sim 30^\circ$ ）、缩放（ $0.7 \sim 1.3$ 倍）、色彩抖动（亮度 $\pm 15\%$ 、饱和度 $\pm 20\%$ ）及 MixUp 混合增强，将原始 5 万张数据集扩充至 15 万张。MixUp 公式为：

$$I_{\text{mix}} = \lambda I_i + (1 - \lambda) I_j \quad (4)$$

$$y_{\text{mix}} = \lambda y_i + (1 - \lambda) y_j \quad (5)$$

式中： I_{mix} 为混合后的样本图像； y_{mix} 为混合后的样本标签； λ 为混合系数，服从 $\text{Bata}(\alpha, \alpha)$ 分布， $\alpha=2$ ，用于控制两张图像的混合比例； I_i 、 I_j 为随机选取的两张原始样本图像。

场景模拟增强阶段，模拟实际应用中的复杂场景，添加轻微杂质、调整光照强度等，提升模型鲁棒性。最终将扩充后的数据集按 7:1.5:1.5 比例随机划分为训练集、验证集与测试集。

2.3 注意力增强型多尺度 CNN 模型构建

(1) 基础模型优化

以 ResNet-50 为基础网络，保留其残差连接结构，残差块可有效缓解深层网络训练中的梯度消失问题。

$$y = F(x, W) + 5x \quad (6)$$

式中： $F(x, W)$ 为残差映射； x 为输入特征，通过 shortcut、连接实现恒等映射。

针对中药材分类任务，将顶层全连接层输出维度调整为 150 类，适配 150 种中药材的识别需求。为实现模型轻量化与泛化能力提升，引入 Dropout 层（dropout rate=0.5）与 L_2 正则化（ $\lambda=0.001$ ）抑制过拟合， L_2 正则化项用公式表示为：

$$L_{\text{reg}} = \lambda \sum_{k=1}^K \|W_k\|_2^2 \quad (7)$$

式中： W_k 为网络第 k 层权重参数； K 为网络总层数。采用深度可分离卷积替换网络中 30% 的普通卷积，深度可分离卷积由深度卷积与逐点卷积组成。

(2) 注意力模块嵌入设计

为强化中药材细粒度特征提取能力，设计了通道注意力和空间注意力的双注意力机制，模型结构如图 2 所示。

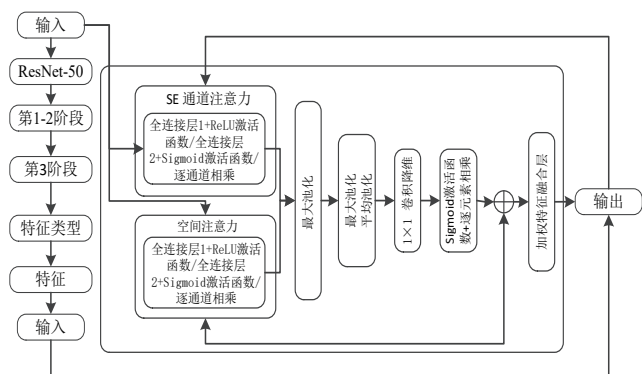


图 2 模型结构

通道注意力模块在网络第 3、4 阶段插入 SE（squeeze-and-excitation）通道注意力模块，通过全局平均池化将每个特征通道压缩为 $1 \times 1 \times C$ 的特征向量，再通过全连接层与 Sigmoid 激活函数生成通道注意力权重，动态调整各特征通道的重要性，强化与中药材品种相关的有效特征通道，抑制冗余信息；空间注意力模块：在网络顶层添加空间注意力模块，对通道注意力处理后的特征图，沿通道维度计算最大值与平均值，融合生成 $2 \times H \times W$ 的特征图，经卷积压缩为 $1 \times H \times W$ 的空间注意力图，通过与原始特征图逐元素相乘，突出中药材纹理、断面、斑点等关键局部特征，提升细粒度识别能力。

(3) 多尺度特征融合机制

设计横向连接的多尺度特征融合结构，充分利用不同网络层级的特征信息。

① 特征提取：从 ResNet-50 网络第 2、3、4 阶段分别提取低、中、高三个尺度的特征图 F_l 、 F_m 、 F_h ，其中第 2 阶段特征图保留中药材宏观形态信息，第 4 阶段特征图包含丰富的微观纹理细节。

② 维度统一：通过 1×1 卷积将三个尺度的特征图统一至相同通道维度（256 维），减少维度差异导致的信息损失。

卷积计算公式为：

$$\begin{aligned} F'_l &= \text{Conv}_{1 \times 1}(F_l, W_l), F'_m \\ &= \text{Conv}_{1 \times 1}(F_m, W_m), F'_h \\ &= \text{Conv}_{1 \times 1}(F_h, W_h) \end{aligned} \quad (8)$$

式中： W_l 、 W_m 、 W_h 分别为三个尺度的 1×1 卷积权重。

③ 自适应融合：采用自适应权重融合策略，通过全连接层学习各尺度特征的重要性权重，对统一维度后的特征图进行加权拼接，实现宏观形态与微观纹理特征的协同捕捉，提升模型对不同形态、大小中药材的适应能力。

2.4 模型训练策略优化

(1) 迁移学习初始化

采用迁移学习策略加速模型收敛，利用 ImageNet 数据集预训练的 ResNet-50 参数初始化基础网络，冻结底层前 7 层卷积层，仅微调第 8 层及以上卷积层、注意力模块与顶层分类器，减少模型训练对大规模标注样本的依赖，提升训练效率。

(2) 半监督迭代训练

针对中药材标注样本稀缺问题，设计预训练、伪标签迭代的半监督训练框架。

① 基础模型训练：使用标注样本训练基础模型，采用 Adam 优化器，初始学习率为 0.001，与交叉熵损失函数，设置 batch size=32，训练 30 个 epoch。

② 伪标签生成：用基础模型对未标注样本进行预测，选取置信度 ≥ 0.95 的样本作为高质量伪标签样本，与原始标注样本按 3:1 比例混合构建新训练集。

③ 迭代优化：每 5 个训练 epoch 重新评估未标注样本置信度，更新伪标签集，动态调整训练数据分布，迭代训练至模型性能收敛，进一步提升模型泛化能力。

(3) 超参数优化策略

采用余弦退火学习率调整策略，训练过程中学习率随 epoch 数周期性衰减，避免局部最优解。通过网格搜索法优化 Dropout 比例 0.3~0.7、 L_2 正则化系数 0.000 1~0.01 等超参数，选取验证集准确率最高的参数组合作为最优配置。

3 实验分析

3.1 实验环境

(1) 硬件环境：Intel Core i9-12900K 8 核 16 线程处理器（主频 3.2 GHz），32 GB DDR5 5 600 MHz 内存，NVIDIA RTX 3090 24 GB 显存显卡，1 TB NVMe 固态硬盘。

(2) 软件环境：操作系统为 Ubuntu 20.04 LTS，深度学习框架采用 PyTorch 1.12.1（搭配 torchvision 0.13.1），

图像处理依赖 OpenCV 4.5.5、Pillow 9.2.0，数据处理使用 NumPy 1.23.5、Pandas 1.5.3，模型评估采用 Scikit-learn 1.1.2、Matplotlib 3.5.2。

3.2 对比实验设计

为全面评估模型性能，选取准确率、精确率、召回率、 F_1 值、平均精度均值（mAP@0.5）、模型参数量为分类任务核心指标。为验证本文模型的优越性，设计两组对比实验：

（1）基础模型对比：将本文模型与传统 CNN 模型（AlexNet、VGG16、原始 ResNet-50）对比，验证基础网络优化的有效性。

（2）注意力机制消融实验：分别测试仅通道注意力、仅空间注意力、双注意力三种配置，验证双注意力机制的协同作用。

3.3 实验结果与分析

（1）基础性能评估

表 1 展示了本文模型与传统 CNN 模型的核心性能对比。整体来看，本文模型在准确率、精确率、召回率、 F_1 值和 mAP@0.5 等核心分类指标上均显著优于 AlexNet、VGG16 和原始 ResNet-50，同时参数量为所有对比模型中最小。这一结果表明，通过嵌入双注意力机制与多尺度特征融合结构，结合深度可分离卷积的轻量化改进，本文模型实现了识别精度与运行效率的协同提升，在保证高性能的同时具备更优的部署灵活性。

表 1 本文模型与传统 CNN 模型性能对比

模型	准确率 / %	精确率 / %	召回率 / %	F_1 值 / %	mAP@0.5 / %	参数量 / 10^6
AlexNet	82.1	81.7	81.5	81.6	83.2	61.1
VGG16	88.5	88.2	87.9	88.0	89.5	138.4
原始 ResNet-50	92.5	92.3	92.1	92.2	92.8	25.6
本文模型	96.7	96.5	96.3	96.4	97.2	17.6

（2）注意力机制消融实验

表 2 为注意力机制消融实验结果，验证双注意力机制的协同增益效果。

表 2 消融实验结果

模型配置	单位：%		
	准确率	易混淆品种识别准确率	特征提取效率特征图有效信息占比
无注意力机制	92.5	85.7	78.2
仅通道注意力（SE）	94.8	89.3	86.5
仅空间注意力	95.1	90.1	87.8
通道 + 空间双注意力	96.7	93.2	92.4

单独引入通道注意力或空间注意力均可提升模型性能，其中空间注意力对易混淆品种识别的增益更显著，原因是空间注意力精准聚焦中药材纹理、断面等局部特征，解决了近缘品种形态相似的识别难题。

双注意力机制结合后，准确率进一步提升，易混淆品种识别准确率也得到提升，证明通道注意力与空间注意力形成互补，实现了通道、空间维度的特征协同增强。

4 总结

本文针对中药材人工识别痛点及现有模型短板，构建了注意力增强型多尺度 CNN 中药材自动识别模型。研究以 150 种常见中药材为对象，从 20 个主产区收集多加工方式样本，经预处理与组合式数据增强扩充至 15 万张，搭建标准化多维度数据集。基于 ResNet-50 优化网络结构，嵌入双注意力机制与多尺度特征融合技术，强化宏观与微观特征协同捕捉能力。融合迁移学习与半监督训练策略，降低标注样本依赖，兼顾模型精度与轻量化。实验显示，模型测试集准确率达 96.7%，易混淆品种识别效果优异，核心性能优于传统 CNN 模型，且复杂场景鲁棒性强。本文为中药材质量检测提供了可靠技术支撑，后续将扩充数据集、优化模型性能，推动其在中医药产业落地，助力行业标准化发展。

参考文献：

[1] 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.

[2] 王甘红, 张子豪, 奚美娟, 等. 基于卷积神经网络建立中药材自动识别的人工智能模型及应用程序 [J]. 中国全科医学, 2025, 28(9): 1128-1136.

[3] 韩勇, 兰杰, 郭瑞瑶, 等. 基于深度学习的中药材饮片识别 [J]. 西北农业学报, 2023, 32(11): 1859-1867.

[4] 胡晓东. 基于改进深度学习算法的中药饮片图像识别研究 [D]. 吉林: 吉林农业大学, 2023.

[5] 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.

【作者简介】

林兴（1989—），男，福建福州人，本科，讲师，研究方向：信息管理与信息系统、计算机应用技术。

（收稿日期：2026-01-28 修回日期：2026-03-13）