

# 基于多分支神经网络的中药材菊花性状识别方法

## 摘要

随着近十年人工智能技术的高速发展，智能化、数字化数据分析和模式分类方法得到了长足发展。以深度学习为典型代表的人工智能技术已潜移默化地影响和改变各个行业，然而，基于人工智能的数字化自动化分析技术在中药材视觉检验领域发展缓慢。本文以药用菊花性状识别为例，开展以深度学习为基础的快速、自动化视觉检定技术，提出一种多分支卷积神经网络学习算法，实现检验员在实际场景下的多种特征融合的综合精确判断思维方式。所提出的多分支神经网络算法分别将底阶、中阶和高阶特征进行自适应提取，然后将三个分支上的数据特征进行自动汇聚，进而完成了从底层局部信息到高层抽象特征的全流程特征抽取和自动融合，实现了多层次视觉特征学习和智能综合决策。该论文从药用菊花的实际数据采集、存储、图像预处理、神经网络算法设计以及实验分析等角度来全面、系统地验证所提出算法的有效性。

## 引言

中医药作为我国独特的卫生资源、潜力巨大的经济资源、具有原创优势的科技资源、优秀的文化资源和重要的生态资源，在经济社会发展中发挥着重要作用。国务院印发的《中医药发展战略规划纲要（2016—2030年）》中将中医药发展列为我国深化医药卫生体制改革中的工作重点，其中，推动中医药发展的现代化建设是体制改革的重要发展目标。中医药在治病领域内有着独特的优势，并且发展空间广阔，中医药在临床方面的需求不断增加，也为中医药的发展提供了前所未有的机遇。

中药的性状鉴别是历经千百年传承的传统鉴别方式，它主要根据药材的外观特征直接利用感官得出真伪优劣的鉴别结果，具有快速、直观、经济等特点。虽然现代中药鉴定新技术蓬勃发展，但性状鉴定始终是研究开发中药鉴定新方法的前提和基础，具有不可替代的重要作用。

从古至今，性状鉴别是以经验为主导，依靠专家口述和参照图谱两种形式进行传承和发展。但作为生物体，中药本身具有个体差异，同时经过种植采收、炮制加工、存储等过程，形状、颜色、细节纹理等特征会发生改变，后辈在学习过程中按图索骥经常出现认知和理解偏差。同时某些鉴别特征难以用客观的量值形式进行文字描述，因此质量标准对性状规定多采用类比形式，如线纹海马的头部

描述为“头略似马头，具冠状突起，具管状长吻，口小”，依赖人员的主观判断长短、大小和整体状态，无法实现客观评价，可能带来错误判断的严重后果。

具体来讲，性状鉴别任务主要是通过样品多个角度的图像，确定特定药材的种类。在传统的人工性状鉴别任务中，由于药材具有不同的种类，例如药用菊花主要包含毫菊、滁菊、贡菊、杭菊和怀菊五类，而它们之间的形态差异较小，因而人工判别需要依赖大量经验，耗费时间长，新人在按图索骥的学习中难以避免遇到认知和理解偏差。

随着机器学习和深度学习技术的不断发展，基于自动化人工智能技术（特别是计算机视觉技术）的自动化智能识别系统不仅可以消除传统诊断检验员的主观意识和经验积累的影响因素，还能实现中医药药材的智能分类。

针对鉴别关键特征难量化、鉴别主观性强及经验传承困难等问题，本项目提出基于多视图特征融合的中药材性状图像识别的技术，采用大数据驱动的深度学习识别方法，克服单一特征的零和博弈评价的缺点，模拟人类模糊综合判断的方式，实现中药材的高性能、自动化智能鉴别。

## 下面是采集及算法部分

### 基于多分支神经网络的中药材性状鉴别算法

中药材性状的传统鉴别方式，以经验为主导，主要根据药材的外观特征，利用人的感官得出真伪优劣的鉴别结果，具有快速、直观、经济等特点。中药材本身具有个体差异，同时经过种植采收、炮制加工、存储等过程，形状、颜色、细节纹理等特征会发生改变；后辈在学习过程中按图索骥经常出现认知和理解偏差；依赖人员的主观判断长短、大小和整体状态，无法实现客观评价，可能带来错误判断的严重后果。

本论文提出一种基于多分支神经网络模型模仿人类实际药品鉴定过程中需要多种药材特征综合决策的思想，利用多种特征融合的中药材性状图像识别技术实现自动化中药材鉴定，特别是基于以大数据预训练模型为基础的多个分支深度神经网络模型为主要技术，克服单一特征的无法综合多种不同角度的特征实现自适应特征融合判断的难点，实现中药材的快速、自动智能鉴别。

本阶段项目，以药用菊花为示范，搭建了药用菊花性状识别的自动化智能识别系统，主要包括样本采集、算法设计和系统识别模型结果展示。

#### 1 样本采集

### 1.1 采集条件

为控制统一的光照、背景等环境条件，我们制作了简易的拍摄箱，原型图如图 2-2 所示。

箱子主要包含四部分：12 个 LED 灯管，两个 1080P 分辨率高清 USB 摄像头，置物台，电路连接线及开关。12 个 LED 灯管分布在箱子立方体的每条边上，使发出的白色光源均匀充分；拍摄设备采用 1080P 分辨率高清 USB 摄像头，连接电脑，实现同步拍摄和保存功能；箱子包含一个箱门，用于放置药材样本；箱子可封闭，内部全白色，置物台底面也是白色，构成全白的封闭空间；箱子左侧是电路板板和开关，实现对摄像头、灯管和箱门线路的控制。

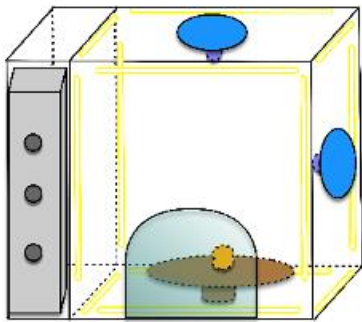


图 2-2 简易拍摄箱子图

### 1.2 采集实例

药用菊花分为五大类：亳菊，滁菊，贡菊，杭菊，怀菊。具体如图 2-3 所示。

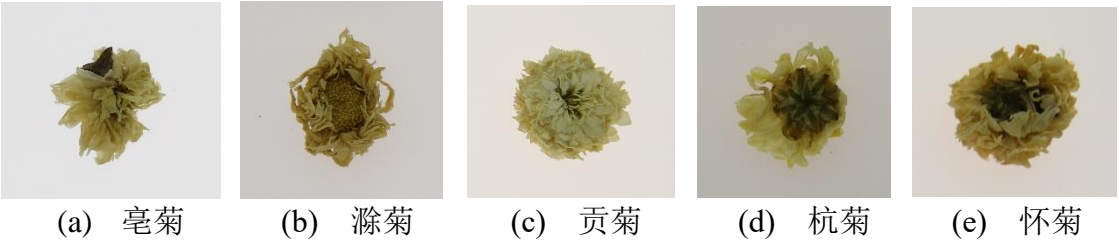


图 2-3 五类药用菊花样例

图像采集，分为正面，侧面和底面三个角度拍摄，提供药材多视角的特征信息，有利于提高模型识别的准确率和泛化能力。五种不同种类的中药菊花，保存在样本袋里，以贡菊为例，具体如图 2-4 所示。



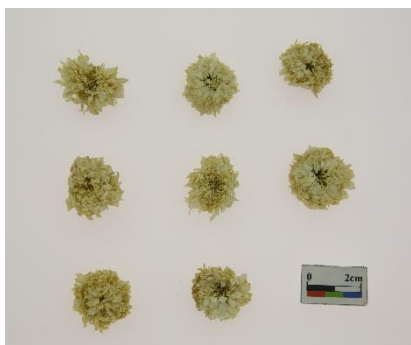
图 2-4 贡菊样本袋图

拍摄步骤具体如下：

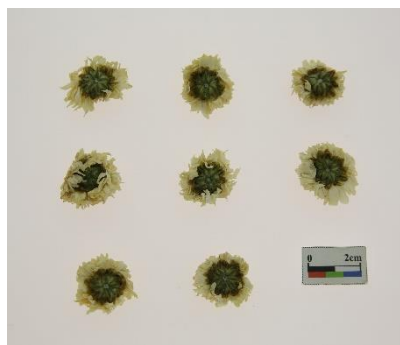
将 USB 摄像头连接电脑，打开 LED 灯管电源，从装有中药菊花的样本袋里，随机轻轻的捏取任意朵菊花样本，分散无序的放在置物台上（尽量分散不要有重叠部分），关闭箱子，在电脑上点击拍摄按钮，获取拍摄图像，下载到本地保存。

药用菊花采集图像，以贡菊的正面、侧面和底面为样例，如图 2-5 所示。

现阶段，采集图像数量，每类菊花有 460 张图片，共计  $460 \times 5 = 2300$  张图片，其中正面，侧面和底面所占比例约为：6:2:1。



(a) 贡菊正面拍摄样例



(b) 贡菊底面拍摄样例



(c) 贡菊侧面拍摄样例

图 2-5 贡菊多视角拍摄图像样例

## 2. 算法设计

### 2.1 图像预处理：检测与分割

下面以图 2-5(a)中贡菊正面拍摄图像为例，介绍图像预处理检测和分割的算法。

### 2.1.1 基于区域选择性搜索（Selective Search）的菊花检测分割算法

#### (1) 分层分组算法：

主要分为以下几个步骤：

- a) 计算所有相邻区域之间的相似性。
- b) 合并两个最相似的区域。
- c) 计算合并区域和相邻区域的相似度。
- d) 重复 b、c 过程，直至图像变为一个区域。

最后输出物体位置的可能结果。

#### (2) 多样化策略

一张图像中包含的信息非常丰富，图像中的物体有不同的形状、尺寸、颜色、纹理等关系，根据颜色，纹理和尺度，采用不同的相似度进行衡量，适应不同图像的不同特征关系，达到更好的检测分割效果。

#### (3) 算法检测分割效果

算法效果，如图 2-6 所示。

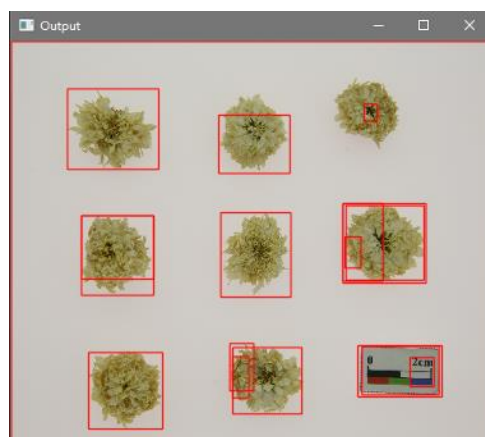


图 2-6 基于区域选择性搜索算法的检测分割效果图

#### (4) 算法分析

由于中药菊花纹理复杂，颜色与背景相似，导致基于区域选择性搜索的检测分割算法效果不好，存在漏检、多检等情况。

### 2.1.2 基于 COCO 数据库目标检测预训练模型 Detectron2 的菊花检测分割算法

#### (1) COCO 数据库：

COCO 数据集由 5 个关键字段构成，检测对象包含 80 个类别。Detectron2 是一个比较成熟的主流检测框架，我们在 COCO 数据集上进行训练，然后对本任务中药菊花进行检测分割。

#### (2) 算法检测分割效果：

算法效果，如图 2-7 所示：



图 2-7 基于 COCO 数据库的预训练模型 Detectron2 的检测分割效果图

### (3) 算法分析：

由于 COCO 数据库的对象分类跟本任务的中药菊花有着很大区别，缺乏对中药菊花的检测训练，导致算法效果不好，存在漏检、误识别等情况。

## 2.1.3 基于背景去噪的菊花检测分割算法

### (1) 彩色图像阈值分割（去背景）

彩色图像由 RGB(红绿蓝) 三通道组合而成，为了能够精准的检测出菊花药材的位置并切割出来，我们首先进行去背景操作，减少除菊花药材以外的任何背景因素的干扰。



图 2-8 去背景处理效果图

设置像素阈值为 125，对所有位置进行检测，当前位置像素值如果大于此阈值，则认为是背景因素，然后将此位置像素值设置为 0，也就是黑色，有利于跟菊花药材的像素值进行区分。去背景处理效果图如图 2-8 所示。

针对存在误识别色彩对比卡的情况，算法做了进一步的优化。

减少干扰色彩对比卡的干扰，防止误识别情况，详细分析色彩对比卡的像素值以及与环境的对比，针对色彩对比卡设置特定的像素阈值，不断优化，去除色彩对比卡的影响。

### (2) 检测物体轮廓

对(1)中得到的去背景图像进行二值化处理，便于更好的分析物体的性状和轮廓。然后开始进行菊花药材轮廓的检测，并在原图中画出检测的轮廓结果，检测结果如图 2-9 所示。



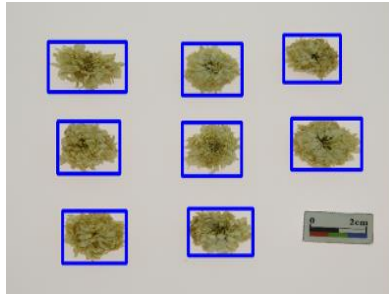


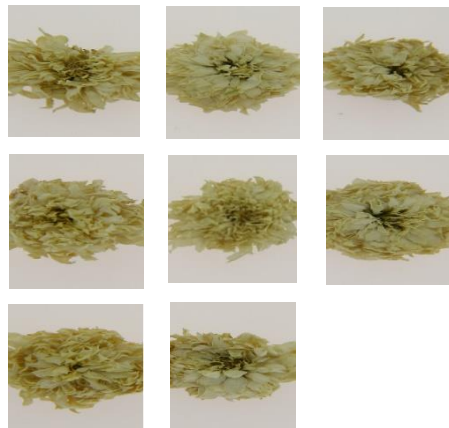
图 2-9 基于背景去噪算法的菊花药材轮廓检测效果图

针对对于尺寸较小菊花样本漏检的情况，算法做了进一步的优化。

减少漏检，有时对于很小的菊花样本，会当作背景识别不出来，在检测物体轮廓时，将识别框阈值进行调整，不断优化，最后调整到适应于不同大小的菊花样本，能够更好的更全面的实现对样本的检测，减少漏检的情况。

### (3) 切割目标

对(2)中得到的轮廓检测图，根据检测到的轮廓的坐标，进行每个菊花药材的



切割，切割效果如图 2-10 所示。

图 2-10 基于背景去噪算法的菊花药材切割效果图

### (4) 速度提升：

在去背景操作中，为了更准确更快速的识别背景，针对像素阈值设置了一个辅助矩阵，称为“mask”矩阵，利用计算机硬件对于矩阵进行快速运算，代替之前对于像素值繁琐的两层循环机制，将分割速度提高了 3-4s，由之前的 4s 降低到 1s 以内，实现了速度质的提升，同时保证了分割检测的准确性。

### (5) 算法分析：

由于中药菊花纹理复杂，尺度不一，颜色与背景差异不明显，对基于背景去噪的算法设置合理的阈值，针对不同的复杂情况进行优化，可以有效的减少背景等因素的干扰，更准确的判断出中药菊花的位置，可以获得更好的检测分割效果。

## 2.2 多分支融合细粒度分类模型

### 2.2.1 多分支融合细粒度分类模型整体框架

根据项目需求，我们需要实现五类药用菊花的自动化精细分类，这一图像分类任务在计算机视觉领域中属于细粒度图像分类（Fine-grained Image Classification），旨在对属于同一基础类别的图像（如猫、狗、花、鸟等）进行更加细致的子类划分。但是这五类药用菊花之间的类间差异非常细微，直接使用通用的图像分类模型来对其进行分类会取得不理想的效果。

因此，我们提出了多分支融合细粒度分类模型 MBFN（Multi-Branch Fusion Network），这一模型采用了多分支深度残差网络结构提取并融合不同层次的图像高层特征，最终得到更综合、更具有判别性的图像高层语义特征，使得整个分类模型在药用菊花细粒度分类任务中取得了很好的效果。

MBFN 模型的具体架构如图 2-11 所示，主要由两个部分组成：多分支特征提取模块、特征融合模块和分类模块。

多分支特征提取模块主要包含了三个不同深度的深度残差网络，其中每个深度残差网络可视为一个独立的分支，可以提取到特定深度的图像高层特征。每个分支的尾部都有一层全连接层，用以对提取到的图像高层特征进行线性变换，得到压缩后的图像高层特征。最终，我们可以从输入的药用菊花图像中提取得到三种不同层次的图像高层特征。

特征融合模块对前面提取得到的三种不同层次的图像高层特征进行相加，得到更综合、更具有判别性的图像高层语义特征。

分类模块主要由一层全连接层和一层 Softmax 层组成，其中全连接层对前面融合得到的特征进行进一步地线性变换，以进行不同维度的信息交互和融合，得到具有药用菊花类别信息的图像高层语义特征。最后的 Softmax 层对这一特征进行归一化，最终得到当前输入的药用菊花图像被分类为各种类别的概率，其中概率最大的类别即为模型输出的预测类别。

（插入图）

图 2-11 MBFN 模型架构图

在训练 MBFN 模型的过程中，我们采用交叉熵损失函数

### 2.2.2 分类模型实际应用情况

以图 2-5(a)中贡菊正面拍摄图像为例，最后的药用菊花分割检测以及识别结果，如图 2-13 所示。



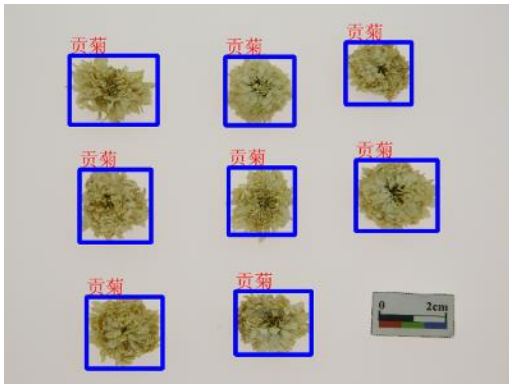


图 2-12 菊花药材切割检测与识别结果图

2.3 系统识别结果

2.3.1 数据集划分

基于以上图像采集设备可采集到一批药用菊花的图像数据，作为我们的基准数据集 ChryFG\_Dataset。ChryFG\_Dataset 包含了五类药用菊花的图像，其中亳菊图像 41 张、怀菊图像 438 张、贡菊图像 350 张、滁菊图像 667 张、杭菊图像 307 张。我们对 ChryFG\_Dataset 按照 3:1 的比例将每一类药用菊花进行划分，得到训练集和测试集，用以模型的后续训练和优化。

2.3.2 模型识别结果

本文提出的 MBFN 模型已迭代优化至第五版，每一历史版本的分类准确率，如表 2-1 所示。可以看到，模型经过不断的优化，已经将分类的准确率从 88.3% 提高到了 97.6%。

不同版本的 MBFN 模型	分类准确率
MBFN_v1	88.3%
MBFN_v2	92.4%
MBFN_v3	93.6%
MBFN_v4	94.5%
<b>MBFN_v5</b>	<b>97.6%</b>

表 2-1 不同版本的 MBFN 模型的分类准确率

此外，如表 2-2 所示，与经典的图像分类模型(VGG16、AlexNet、GoogLeNet)

相比，我们提出的 MBFN\_v5 模型取得了更高的分类准确率，可以更准确地对每一类药用菊花进行识别。其中 VGG16、AlexNet、GoogLeNet 和 MBFN\_v5 模型都加载了预训练权重进行初始化，再在 ChryFG\_Dataset 数据集上进行训练微调，以加快模型的收敛速度，同时取得更高的准确率。

图像分类模型	分类准确率
VGG16	87.4%
AlexNet	92.7%
GoogLeNet	93.4%
<b>MBFN_v5</b>	<b>97.6%</b>

表 2-2 不同图像分类模型的分类准确率

其中，MBFN\_v5 模型在测试集上的混淆矩阵如图 2-14 所示，纵坐标表示数据原本的类别，横坐标表示模型预测出的类别，比如第一行第一列，代表的意义是，原本为亳菊的菊花药材，被模型识别为亳菊的数目有 10 个。矩阵对角线的元素，就代表了模型对每类药用菊花正确分类的数目。

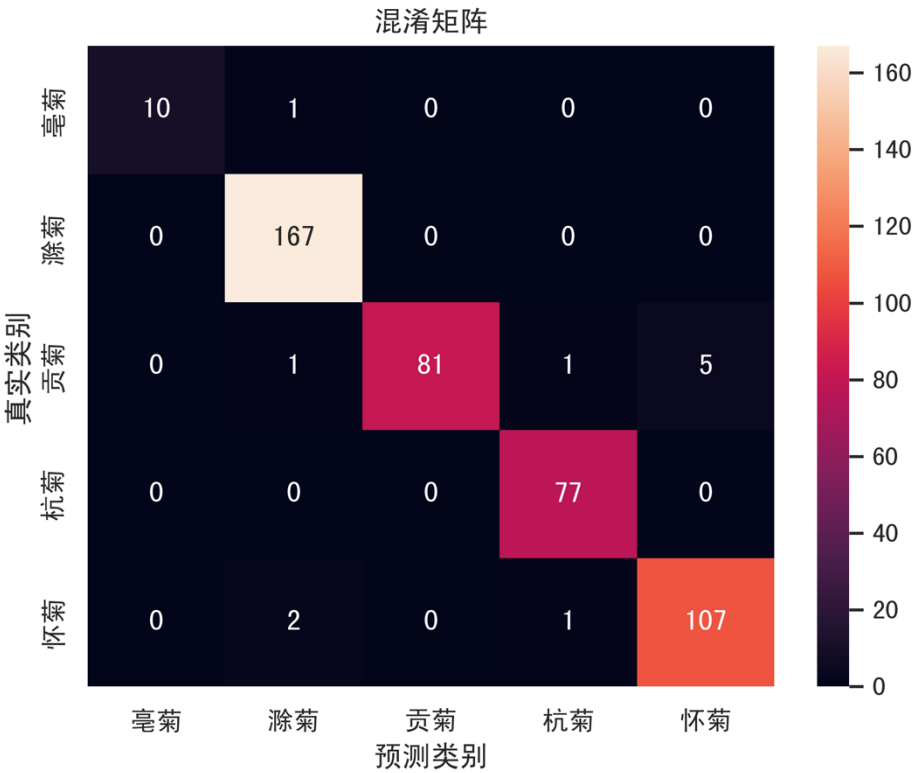


图 2-13 MBFN\_v5 模型在测试集上的混淆矩阵

