

III. Инфо-коммуникационные технологии

For citation: Liu Ao. Design of image recognition algorithms for Chinese Herbal medicine plants based on deep learning // Grand Altai Research & Education — Issue 1 (26)'2026
DOI: 10.25712/ASTU.2410-485X.2026.01) — EDN: <https://elibrary.ru/CIXDJJ>

УДК: 004.421.6

DESIGN OF IMAGE RECOGNITION ALGORITHMS FOR CHINESE HERBAL MEDICINE PLANTS BASED ON DEEP LEARNING

Liu Ao¹

¹ Hubei Digital Textile Equipment Key Laboratory, Wuhan Textile University, Wuhan, 430073, China
E-mail: 1476563758@qq.com

Abstract. Chinese herbal medicine identification is of great significance to the pharmaceutical industry and the sustainable development of herbal production. Traditional identification relies on feature engineering with expert participation, featuring high labor costs and decreasing recognition accuracy as the number of species increases. This paper designs a recognition algorithm based on deep learning, constructs a sample set expanded to 45,236 images through data augmentation, and selects the EfficientNet B2 model based on the TensorFlow and Keras frameworks. With an accuracy of 90.66%, it provides a high-performance solution for Chinese herbal medicine identification and opens up new paths for the application of deep learning in this field.

Keywords: Chinese herbal medicine; image recognition; convolutional neural network; deep learning

基于深度学习的中草药植物图像识别算法设计

刘奥¹

¹ 武汉纺织大学, 湖北省数字化纺织装备重点实验室, 中国武汉, 430073
E-mail: 1476563758@qq.com

摘要: 中草药识别对医药工业及草药生产的可持续发展具有重要意义。传统识别方法依赖专家参与的特征工程, 存在人力成本高, 识别率随种类增加而下降的问题。本文设计了一种基于深度学习的识别算法, 构建样本集并通过数据增强扩充至 45236 张图像, 基于 TensorFlow 和 Keras 框架选用 EfficientNet B2 模型。该模型准确率达 90.66%, 为中草药识别提供了高性能解决方案, 也为深度学习在该领域的应用开辟了新路径。

关键词: 中草药; 图像识别; 卷积神经网络; 深度学习

0 引言

本文设计了一种基于深度学习 [1] 的识别算法, 构建样本集并通过数据增强扩充至 45236 张图像, 基于 TensorFlow 和 Keras 框架 [2] 选用 EfficientNetB2 模型。该模型准确率达 90.66%, 为中草药识别 [3] 提供了高性能解决方案, 也为深度学习在该领域的应用开辟了新路径。

1 卷积神经网络 (CNN)

1.1 卷积神经网络结构

卷积神经网络是含有卷积操作且具有深度结构的前馈神经网络, 本质上是多层的神经网络结构, 包含卷积层, 池化层和全连接层。

CNN 的核心是卷积层 (Convolutional Layer), 它通过卷积操作提取图像中的特征。卷积操作是指将一个滤波器 (Filter) 与输入图像进行卷积运算, 得到一个

新的特征图 (Feature Map). 滤波器可以看作是一种特征探测器, 它可以检测图像中的某种特定特征, 例如边缘, 纹理等. 卷积层是一种运算方式, 离散的卷积定义为

$$(f \cdot g)(n) = \sum_{\tau=-\infty}^{\infty} f(\tau)g(n-\tau) \quad (1)$$

卷积层作为卷积神经网络处理的第一个环节, 主要用于捕捉输入图像有用的信息, 之后利用内核窗口对图像进行选择, 检测特定的图像模式以减少参数的总数. 卷积层有多个卷积核, 利用多种方式实现对图像的增强和特征提取作为池化层的输入层, 卷积层捕获输入图像后, 计算连接输入局部区域的输出神经元, 对每个区域点乘各自的权重. 其计算表达式如公式 2 所示.

$$Z_i^k = f(W_i^k \otimes Z_{i-1}^k + b_i^k) \quad (2)$$

其中,

Z_i^k — 表示 i 层 k 卷积核形成的特征图,

W_i^k — 表示 i 层 k 卷积核形成的权重矩阵,

b_i^k — 表示 i 层 k 卷积核形成的偏置.

卷积计算过程示意图如图 1 所示。

卷积层的输出送入下一层池化层 (Pooling Layer). 池化层的作用是对特征图进行降采样, 减少特征图的大小, 从而减少计算量和参数数量, 同时也可以提高模型的鲁棒性.

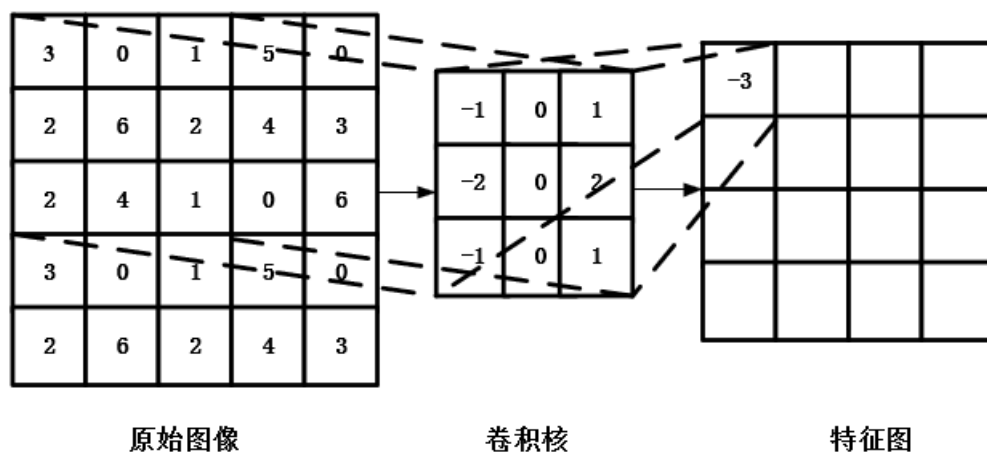


图 1. 卷积计算过程示意图

Figure 1. Schematic diagram of the convolution calculation process

除了卷积层和池化层, CNN 还包含全连接层 (Fully Connected Layer) 和激活函数 (Activation Function). 全连接层将卷积层和池化层的输出展开成一维向量, 然后通过全连接层进行分类或回归等任务. 激活函数则用于引入非线性因素, 增强模型的表达能力.

1.2 激活函数 Swish

激活函数用于将神经元输入转换为输出, 核心作用是引入非线性, 避免多层线性网络退化为单层线性结构, 让深度学习能拟合复杂数据分布. Swish 在 ReLU 基础上改进的激活函数, 可有效缓解梯度消失, 其公式为:

$$f(x) = x \cdot \text{sigmoid}(\beta * x) \quad (3)$$

$$\text{sigmoid}(x) = \frac{1}{1+e^{-x}} \quad (4)$$

Swish 函数是一个平滑的函数 [4], 它的导数在整个实数域上都是连续的, 这使得它在反向传播时计算梯度更加高效. 此外, Swish 函数是一个非单调的函数, 这意味着它可以更好地适应不同的数据分布. 在一些数据集上, Swish 函数比 ReLU 函数具有更好的性能. Swish 函数的导数可以用它自己和 Sigmoid 函数的值来表示, 这使得它在反向传播时计算梯度更加高效. 在一些深度神经网络中, Swish 函数被证明比 ReLU 函数具有更好的性能.

2 基于深度学习的中草药图像识别算法实现

2.1 数据采集

本研究使用爬虫脚本, 于百度图库收集了大量中草药图片, 经过人工优选, 去重后, 数据增强前数据集概况如表 1 所示.

表 1. 中草药数据分布

Table 1. Distribution of Chinese herbal medicine data

中草药名	样本量
当归	587
冬青	618
景天	721
六月雪	582
马蹄莲	720
蜀葵	546
文竹	790
辛夷	494
泽漆	496
紫菀	676

2.2 数据预处理

在使用深度学习进行训练的过程中, 为了增加训练数据和增强模型的鲁棒性, 本课题通过镜像翻转, 亮度变换以及旋转图片进行了数据增强. 在数据增强之后, 将数据划分为三部分: 训练集, 验证集, 测试集. 本文在具体划分时按照时间先后顺序, 前 60% 的数据作为训练集, 中间 20% 的数据作为验证集, 后面 20% 作为测试集.

2.3 水平镜像翻转

设原图像中像素点的位置为 X, Y , 经过水平镜像翻转变换后新图像对应的位置为 X', Y' , $width$ 为图像的宽. 那么:

$$X' = width - X - 1 \quad (5)$$

$$Y' = Y \quad (6)$$

镜像翻转之后的图像如图 2 所示.

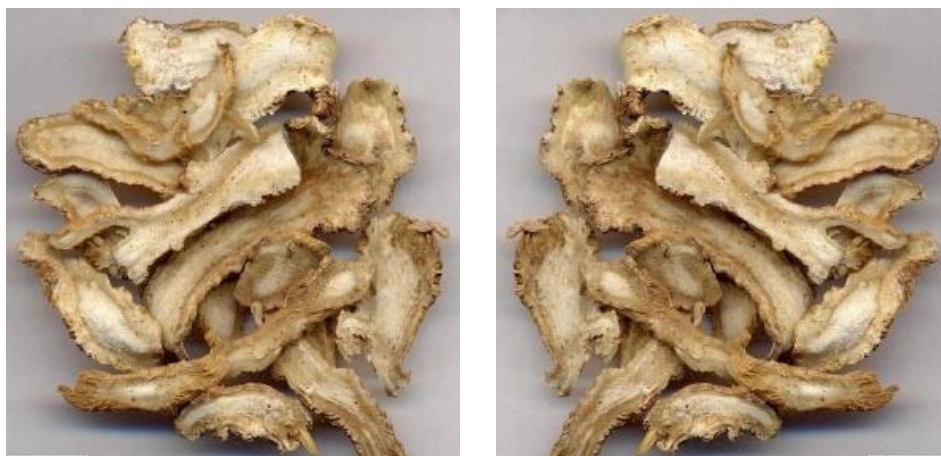


图 2. 原图和翻转后图片对比

Figure 2. Comparison of the original image and the flipped image

2.4 亮度变换

图像的亮度是影响图像特征的因素之一, 在本课题的研究中, 对图像的亮度进行了变化操作. 实现原理是对彩色图像均衡化, 首先将彩色图像从 RGB 色域转换到 HIS 色域, 并分离 HIS 分量, 之后对 I 分量进行均衡化, 并替换原 I 分量, 最后将 HIS 色域换成 RGB 色域. 改变亮度后的图像如图 4 所示.



图 3. 原图和改变亮度后图片对比

Figure 3. Comparison of the original image and the brightness-adjusted image

2.5 图像旋转

对图像按照一定的角度, 在图像的顺时针方向上进行旋转 90° , 180° . 旋转后图片如图 5 所示.



图 4. 原图和旋转后图片对比
Figure 4. Comparison of the original image and the rotated image

3.3 实验环境

3.1 软硬件环境配置

本文模型的训练和测试过程在 Microsoft Windows 11 操作系统下进行, CPU 型号为 AMD Ryzen 7 4800H with Radeon Graphics, GPU 型号为 GeForce GTX 1650, GPU 的可用显存为 6144MiB, 实验采用的深度学习框架为 TensorFlow 和 Keras, 实验的软硬件环境配置的具体信息如表 2 所示.

表 2. 软硬件环境配置

Table 2. Software and hardware environment configuration

配置项	项目值
操作系统	Microsoft Windows 11 专业版
CPU	AMD Ryzen 7 4800H with Radeon Graphics
GPU	GeForce GTX 1650
可用显存	6144MiB
深度学习编程框架	TensorFlow, Keras

3.2 实验结果和分析

本实验在训练模型时, 将训练轮次设置为 10 轮, 每一轮使用训练集对模型进行训练, 并在训练后用验证集验证训练出模型的准确率, 然后通过 Adam 优化器选取在每一轮中最优的模型, 在训练完模型之后, 用测试集对训练出的模型进行测试.

训练与验证结果如图 5, 6 所示. 测试集测试结果如图 7 所示.

```
Epoch 10/10
936/935 [=====] - ETA: 0s - loss: 0.0879 - acc: 0.9684 - precision: 0.9728 - recall: 0.9648
Epoch 10: val_acc did not improve from 0.89116
```

图 5. 模型训练结果
Figure 5. Model training results

```
val_loss: 0.4582 - val_acc: 0.8858 - val_precision: 0.8931 - val_recall: 0.8809
```

图 6. 模型验证结果
Figure 6. Model verification results

loss: 0.3387 - acc: 0.9066 - precision: 0.9161 - recall: 0.8996

图 7. 验证集准确率

Figure 7. Accuracy on the validation set

通过对实验数据的分析, 可以得出训练后的模型在训练集上的准确率达到 88.58%, 说明模型的准确率很高, 实验结果基本满足预期。

测试集的准确率和损失率如图 8 和图 9 所示, 其中横坐标代表模型的迭代次数, 纵坐标分别为准确率和损失率. 通过可视化结果可以看出准确率和损失率的数值都逐步收敛, 由此可以推测模型准确率趋于稳定。

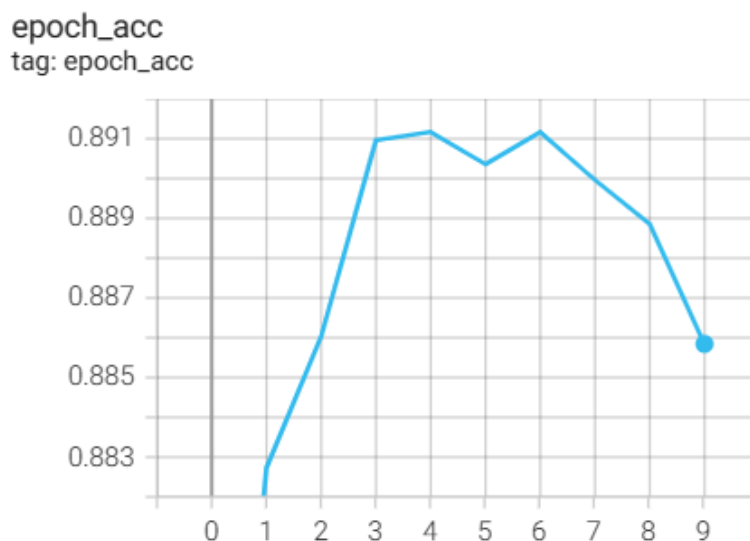


图 8. 测试集准确率 (基于 TensorBoard)

Figure 8. Accuracy on the test set (based on TensorBoard)

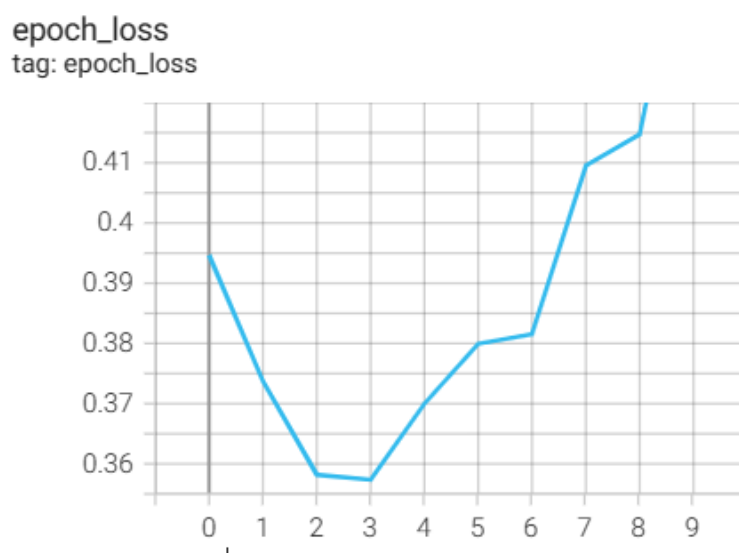


图 9. 测试集损失率 (基于 TensorBoard)

Figure 9. Loss rate on the test set (based on TensorBoard)

3.3 类别对比

混淆矩阵可可视化模型对各类别的分类性能, 每行表示真实类别, 每列表示预测类别, 能直观反映分类混淆情况. 中草药细粒度 [5] 分类中, 物种特征相似, 生长期形态差异易造成误判. 图 10 为迁移学习下 EfficientNet B2 在 10 种中草药测试集上的识别结果, 实验表明该模型识别准确率整体较高.

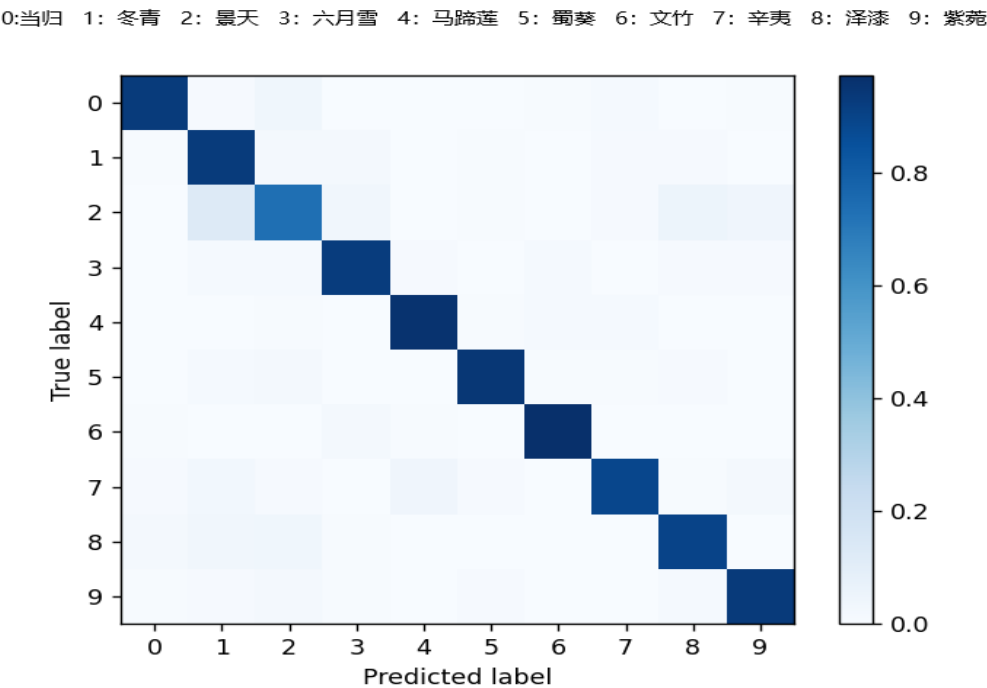


图 10. 混淆矩阵
Figure 10. Confusion matrix

3.4 模型对比

本节对比了 VGG19, ResNet152, MobileNet V2 及 EfficientNet B2/B7 等卷积神经网络模型, 各模型测试结果见表 3.3. VGG19 采用多层 3×3 卷积, 参数量更低, 准确率为 75.66%. ResNet152 通过残差连接缓解梯度消失, 准确率 88.44%. MobileNet V2 采用深度可分离卷积实现轻量化高效推理, 准确率 89.24%. EfficientNet 通过自适应缩放与 MBConv+SE 模块提升性能, 准确率最高, 其中 EfficientNet B2 为 90.66%. 将五种模型进行对比后的结果如表 3 所示.

表 3. CNN 模型对比结果
Table 3. Comparison Results of CNN Models

模型名称	Accuracy	Precision	Recall
VGG19	75.66%	84.33%	67.87%
ResNet152	88.44%	89.32%	88.21%
MobileNet V2	89.24%	89.84%	88.97%
EfficientNet B7	89.92%	91.03%	88.97%
EfficientNet B2	90.66%	91.61%	89.96%

3.5 验证数据增强的作用

本文采用镜像翻转, 亮度变换, 旋转等数据增强方法扩充中草药图像数据集, 以提升模型泛化能力. 对比实验显示, 未增强时模型训练波动大, 准确率仅 87.20%; 增强后模型收敛稳定, 准确率提升至 90.66%. 该方法有效缓解了样本量不足引发的过拟合, 增强了模型鲁棒性与抗干扰能力.

4 结束语

本文设计并实现了基于深度学习的中草药图像识别算法, 核心工作总结如下: 首先, 梳理课题研究现状, 概述了算法实现所用的框架与模型; 其次, 通过镜像翻转, 旋转, 亮度变换完成数据集扩充, 并按 3:1:1 划分训练, 验证与测试集, 采用迁移学习训练 EfficientNet B2 模型, 最终在测试集上实现 90.66% 的识别准确率; 最后, 借助混淆矩阵可视化模型对 10 种中草药的识别效果, 并通过对比数据增强前后及经典模型的实验结果, 验证了本算法中数据增强的有效性和所选模型的高准确率优势.

参考文献

- [1] 张露. 基于深度学习的植物叶片图像识别方法研究 [D]. 北京林业大学, 2019. DOI: 10.26949/d.cnki.gblyu.2019.000548.
- [2] 王明, 张倩. 我国基于深度学习的图像识别技术在农作物病虫害识别中的研究进展 [J]. 中国蔬菜, 2023(03):22-28. DOI: 10.19928/j.cnki.1000-6346.2023.2007.
- [3] 王丽君, 淮永建, 彭月橙. 基于叶片图像多特征融合的观叶植物种类识别 [J]. 北京林业大学学报, 2015, 37(01):55-61. DOI: 10.13332/j.cnki.jbfu.2015.01.006.
- [4] 曹凤莲. 基于深度学习的植物叶片识别方法研究 [D]. 重庆大学, 2017.
- [5] 周婷. 基于深度学习的民族药植物图像识别研究 [D]. 江西中医药大学, 2022. DOI: 10.27180/d.cnki.gjxzc.2022.000018.

References

- [1] Zhang L. Research on plant leaf image recognition method based on deep learning [D]. Beijing Forestry University, 2019. DOI: 10.26949/d.cnki.gblyu.2019.000548.
- [2] Wang M., Zhang Q. Research progress of deep learning-based image recognition technology in crop disease and pest recognition in China [J]. China Vegetables, 2023(03):22-28. DOI: 10.19928/j.cnki.1000-6346.2023.2007.
- [3] Wang L.J., Huai Y.J., Peng Y.C. Ornamental foliage plant species recognition based on multi-feature fusion of leaf images [J]. Journal of Beijing Forestry University, 2015, 37(01):55-61. DOI: 10.13332/j.cnki.jbfu.2015.01.006.
- [4] Cao F.L. Research on plant leaf recognition method based on deep learning [D]. Chongqing University, 2017.
- [5] Zhou T. Research on ethnic medicinal plant image recognition based on deep learning [D]. Jiangxi University of Traditional Chinese Medicine, 2022. DOI: 10.27180/d.cnki.gjxzc.2022.000018.