

基于 MobileNet V2 迁移学习的中药材图像识别

刘雪纯, 刘大铭*, 常佳鑫, 王博

(宁夏大学物理与电子电气工程学院, 宁夏银川, 750000)

(宁夏大学沙漠信息智能感知重点实验室, 宁夏银川 750000)

摘要: 中药材对人体疾病的预防及控制具有重要的作用, 然而普通百姓对药材知识了解过少, 可能滥用草药从而带来不可控的后果。因此, 对药材进行精准识别是一项紧迫的任务。本文将改进的深度卷积神经网络应用到药材识别中, 提出了基于 MobileNet V2 模型进行迁移学习的中药材识别系统。该模型将标准卷积改进为深度可分离卷积形式, 在 ImageNet 数据集上进行了预训练。通过对 MobileNet V2 的预训练模型参数进行多次微调, 在全连接层采用 softmax 激活函数实现药材分类, 使用交叉熵损失函数定量表达与理想模型之间的差异, 最后使用 adam 优化算法实现最优梯度下降, 实现了 12 类中药材图片识别, 最终在验证集上精度可达 93%, 表明本文的中药材图像识别系统具有较好的实际应用场景。同时 MobileNet V2 的轻量级模型在训练时间上也比普通卷积神经网络节省了 2 倍以上的时间, 在药材识别上的鲁棒性和泛化性得到了较大提升。

关键词: 深度卷积神经网络; 药材识别; MobileNet V2; 迁移学习

*基金项目: 宁夏大学研究生创新项目(GIP2021007); 宁夏自然科学基金资助项目(2021AAC03113)

第一作者: 刘雪纯, 女, 1997 年生, 甘肃庆阳人, 硕士研究生; 研究方向为智能仪器仪表研究。E-mail: 12021130682@stu.nxu.edu.cn

通讯作者: 刘大铭, 男, 1969 年生, 宁夏银川人, 硕士, 教授; 研究方向为智能仪器仪表研究。E-mail: ldm@nxu.edu.cn

0 引言

随着中华民族上下五千年的发展历程, 中医已经成为一个较为完善的体系, 但是近些年随着西方医学的发展, 对中医在科学界的冲击较大, 国内众人已经怀疑中医的科学性和适用性。2015 年, 屠呦呦在“青蒿素-中医送给世界的礼物”^[1]一文中提到了青蒿素对治疗人类疟疾的重要作用, 研创了新型抗疟药, 并获得了诺贝尔奖; 2020 年, 中国工程院院士钟南山团队^[2]提出连花清瘟显著抑制新冠病毒, 其影响病毒形态并在体外发挥抗炎活性, 这些贡献使人们重新关注中医。随着中医的广泛普及, 多数人开始使用中药材养生, 但是药材的种类繁多, 一些类间差异较小, 容易误用药材, 由此带来不良后果。因此, 对药材进行精准识别是一项长期而又艰巨的任务。早期的药材识别都是基于物理或者化学特征, 往往需要人工先对大类进行一定的区分, 再依赖特定的器材对中药材进行分类。这些方法能在对应的任务上取得令人满意的成果, 但是不足之处在于学习和识别过程依赖特定器材, 操作复杂, 对日常识别药材极不方便, 也不便于对新数据进行预测。2014 年, Changjiang Liu^[3]等采用了反向传播神经网络 (BPN) 对中药材进行识别, 通过提取药材目标区域 (region of interest, ROI) 的颜色, 纹理和形状特征, 将获取的特征组成的特征向量输入到反向传播神经网络 (BPN) 中, 并将类标签作为神经网络的输出。但是该方法对相互遮挡的和杂乱背景的药材图像识别精度不高, 不适合实际应用。2016 年, Xin Sun^[4]等人提出了卷积神经网络 (CNN) 进行药材的识别和检索, 训练了一个基于 VGGNet 的网络模型, 采用三元组损失对识别网络进行微调, 确保找出不同类别中相似外观的图像, 该网络提高了在杂乱背景下药材识别精度, 使得药材识别技术更适合实际应用, 但是平均识别精度为 71%, 平均检索精度为 53%, 药材识别的精度仍然有待提高。要获得较为精准的识别结果, 则需要大量的训练样本, 但药材识别还属于发展初级阶段, 并没有公开的较为完善的海量数据集。因此, 针对目前药材识别精度低下和数据集稀少的问题, 本文基于深度学习的理论知识, 提出了一种基于 MobileNet V2 模型进行迁移学习的中药材识别系统, 迁移学习在有限的数据集下可以更快提取图像的特征。通过对 MobileNet V2 的预训练模型参数进行多次微调, 在全连接层采用 softmax 激活函数实现药材分类, 使用交叉熵损失函数定量表达与理想模型之间的差异, 最后使用 adam 优化算法进行最优梯度下降, 实现了 12 类中药材图片识别, 并得到了 93% 的识别准确率。

1 建立 MobileNet V2 迁移学习分类模型

1.1 MobileNet V2 模型介绍

MobileNet V2 作为一款轻量级兼备精度和速度的模型，近几年被人们广泛使用，MobileNet 网络结构在传统的卷积神经网络的基础上，加入了深度可分离卷积（Depthwise Separable Convolution）的思想，从而极大减少了网络训练的参数。其中，传统的卷积神经网络是输入特征矩阵的通道数与卷积核的通道数相匹配，卷积核的个数和输出特征矩阵的个数相匹配，如图 1 所示：

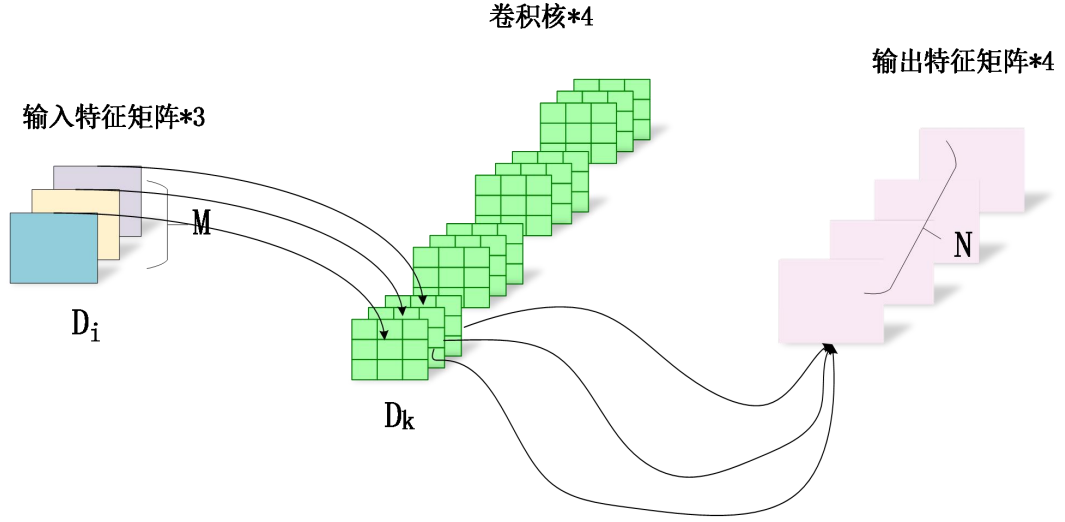


图 1 卷积神经网络

Fig. 1 Convolutional neural networks

因为每个卷积核都与输入特征矩阵的每一个维度进行卷积，所以运算量较大。公式为：

$$D_i * D_i * M * N * D_k * D_k \#(1)$$

而在深度可分离卷积中，将传统卷积方式分解为深度卷积和 1×1 逐点卷积，如图 2。图（a）是深度卷积（Depthwise Conv），卷积核的通道数不随输入特征矩阵改变，每一个卷积核的通道数都是 1，其深度卷积的计算公式为：

$$D_i * D_i * M * D_k * D_k \#(2)$$

图（b）是逐点卷积（Pointwise Conv），逐点卷积和普通的卷积一样，只不过卷积核的大小是 1×1 ，计算公式^[5]为：

$$D_i * D_i * M * N \#(3)$$

深度可分离卷积公式为：

$$D_i * D_i * M * D_k * D_k + D_i * D_i * M * N \#(4)$$

深度可分离卷积与传统卷积运行的比例为：

$$\frac{D_i * D_i * M * D_k * D_k + D_i * D_i * M * N}{D_i * D_i * M * D_k * D_k} = \frac{1}{N} + \frac{1}{D_k^2} \#(5)$$

因为 MobileNet 网络模型中的大部分都是 3×3 的卷积核，因此可分离卷积方法比普通的卷积运算量减少了 8-9 倍，极大地缩短了网络训练的时间、减少了参数更新的成本。

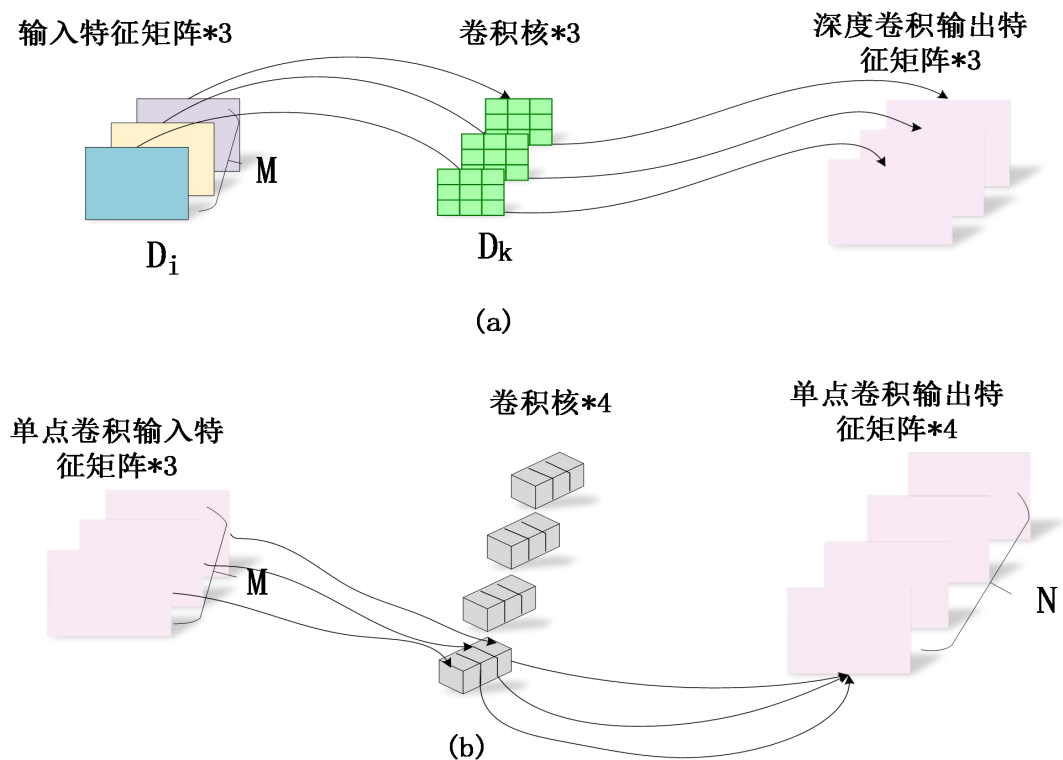


图 2 深度可分离卷积神经网络

Fig. 2 Deeply separable convolutional neural networks

1.2 迁移学习策略

在深度学习中，使用卷积神经网络（CNN）训练数据时，需要海量数据集，这就对硬件要求很高，此时将源领域中学习到的有用知识迁移到目标领域就变得很有必要。迁移学习是指将一个领域上训练好的模型通过适当的方法将其调整到一个新的领域中^[6]。

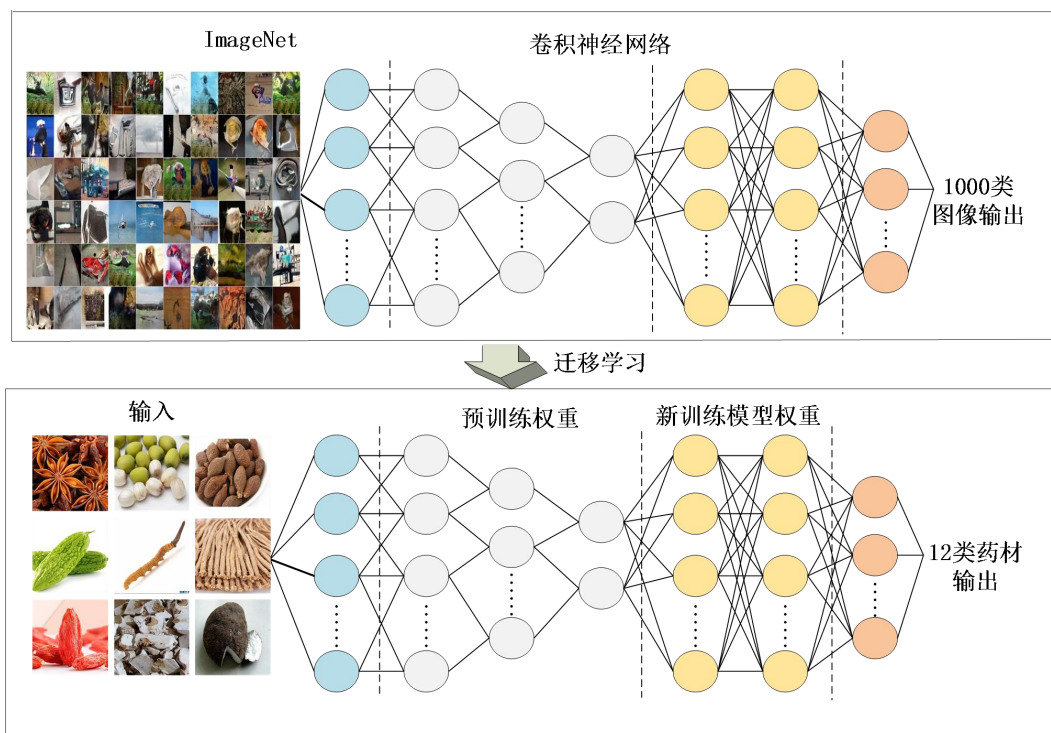


图 3 迁移学习原理图

Fig. 3 Transfer learning schematic

本文使用的是在 ImageNet 数据集上训练好的预训练模型，ImageNet 数据集中含有大量的标注数据，其中包括 1000 种类别物体，因此，全连接层的输出是 1000 个神经元。卷积神经网络（CNN）的卷积层和池化层都是图像特征提取的过程，迁移学习就是冻结卷积神经网络的卷积层和池化层的参数，在新建立的数据集上，直接使用训练好的神经网络对图像进行特征提取^[7]，然后将提取的特征向量作为输入训练一个全新的单层全连接神经网络，并在此网络的基础上进行微调，最终在全连接层输出 12 类的中药材类别，从而降低训练时间，其泛化性和鲁棒性得到了很好的验证，图 3 是迁移学习的原理。迁移学习的优点是能够很好地利用现有大规模数据集训练得到的成功特征结合自身特殊数据集性质进行有针对性地进行再训练，既有效节约了训练时间，又保持相对较高的准确率，是值得深入研究的方向。

1.3 网络优化设计

为了使迁移学习的模型在新的数据集上有更好的泛化性和鲁棒性，本文将全连接层的参数进行了微调，通过多次对比实验，最终决定使用以下参数进行优化。

1.3.1 激活函数

在卷积神经网络运算过程中，卷积求和都是线性运算，为了引入非线性因素，使得模型训练地更加精准，在全连接层加入了激活函数来提高模型的性能。激活函数对于整个神经网络来说十分重要，在反向传播过程中，更新主要取决激活函数的导数^[8]。常见的神经网络激活函数是 sigmoid^[9]，其一般用于两分类任务，表达式为：

$$f(x) = \frac{1}{1 + e^{-x}} \#(6)$$

函数见图 4

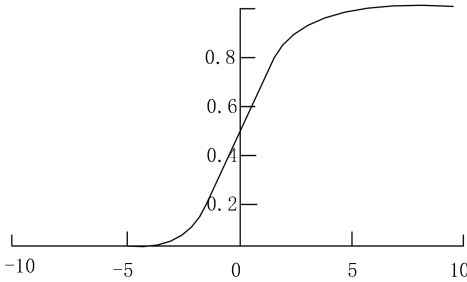


图 4 sigmoid 函数

Fig. 4 sigmoid function

由图 4 可知，激活函数在大于 5 以及小于-5 的时候，导数趋向 0，这就是常见的梯度消失，为了解决梯度消失的现象，可以采用 ReLU 函数^[10]，表达式：

$$ReLU(x) = \max(0, x) \#(7)$$

函数见图 5

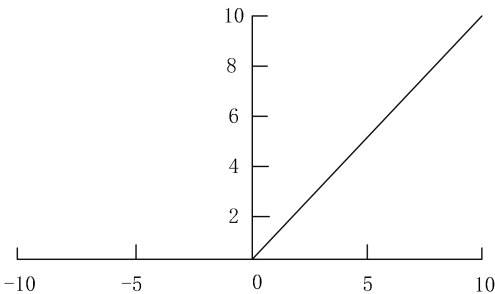


图 5 relu 函数

Fig. 5 relu function

由图 5 可知，激活函数只在大于 0 的时候有定义，且大于 0 以后的导数均是 1，因此 relu 函数解决了梯度消失的现象，但是在卷积神经网络实现分类问题时，结果是一个概率值，并将概率较高的分类结果输出，而 ReLU 函数的最大值可以无穷大，为解决以上问题，我们采用 sigmoid 函数的推广形式 softmax 函数与 relu 函数相结合

的思想进行训练，只在迁移学习之后的全连接层加入 softmax 函数^[11]。softmax 函数在 sigmoid 基础上实现图像多分类，并为每个输出分类的结果都赋予一个概率值，表示属于每个类别的可能性，且多个分类的预测值和为 1。引入指数函数使机器学习中差距大的数值距离拉的更大。表达式为：

$$\text{softmax}(x) = \frac{e^{x_i}}{\sum_i e^{x_i}} \#(8)$$

本文是对 12 类药材进行分类，为了避免梯度消失，最终选择在全连接层加入适合多分类任务的 softmax 函数。

1.3.2 损失函数

在神经网络模型中，损失函数就是网络中的概率模型与理想中的概率模型差异的定量表达。softmax 函数与交叉熵损失函数共同使用模型精确度更高，本文选用交叉熵损失函数^[12]，表达式为：

$$H(X) = - \sum_{i=1}^n P(x_i) \log(Q(x_i)) \quad (X = x_1, x_2, x_3, \dots, x_n) \#(9)$$

其中， $P(x_i)$ 表示真实值， $Q(x_i)$ 表示求出的 softmax 值。

1.3.3 优化算法

随机梯度下降算法是一种有效的优化方法，通常，目标函数是随机的，批量化地存入一组数据进行梯度下降运算，找出最优解。公式为：

$$w_t = w_{t-1} - \alpha * g_t \#(10)$$

公式（10）中， α 是学习率， g_t 是 t 时刻参数 w_t 的损失梯度。由公式（10）知，随机梯度下降算法权重更新过程中学习率不能随着改变，有可能在最优点附近来回震荡，陷入局部最优解，不能提高优化效率。

Adam 旨在结合两种流行的优化算法的优点：自适应梯度算法(AdaGrad)和均方根传播(RMSProp)，AdaGrad 维护参数的学习速率，适用于稀疏梯度；RMSProp 根据最新权重梯度的平均值计算学习率，这意味着 RMSProp 算法在非平稳问题上表现良好，适用于波动环境。Adam 结合两者优点，其参数更新的大小对梯度的重新缩放是不变的，学习率近似地由超参数决定，它不需要固定的目标，与稀疏梯度一起工作^[13]。

公式如下：

$$m_t = \gamma * m_{t-1} + (1 - \gamma) * g_t \#(11)$$

$$n_t = \beta * n_{t-1} + (1 - \beta) * g_t^2 \#(12)$$

$$\widehat{m}_t = \frac{m_t}{1 - \alpha^t} \#(13)$$

$$\widehat{n}_t = \frac{n_t}{1 - \beta^t} \#(14)$$

$$w_t = w_{t-1} - \frac{\alpha}{\sqrt{\widehat{n}_t} + \varepsilon} * \widehat{m}_t \#(15)$$

其中，公式(11)是一阶动量，更新有偏一阶矩估计，公式(12)是二阶动量^[14]，更新有偏二阶矩估计，公式(13)是修正一阶矩偏差，公式(14)是修正二阶矩偏差，公式(15)是梯度更新， α 是学习率， $\frac{1}{\sqrt{\widehat{n}_t} + \varepsilon}$ 目的是减少 α 的值，

使 α 在训练过程中的值越来越小， ε 一般取 10^{-7} ，防止分母为零，实现自适应学习率的特点。
本文综合考虑，决定采用 Adam 自适应学习率的优化算法，可以实现动态的改变学习率。

2 数据集收集及实验环境

2.1 数据集收集

本文的数据集由两部分构成，通过网络爬虫和中药店采集的方式获取图片，构建了 12 类中药材数据集，其中包括易混淆药材。由 3520 张彩色图片构成，每类图片在 100–500 张不等，其中所有图片都经过筛选，为 jpg 格式^[15]，在训练前统一为 224*224 的大小，以加快训练速度，并按照 8: 2 的方式分成了训练集和验证集，表 1 是各类药材数据集的样本分布。为了增强训练结果的泛化能力，本文在实验前使用了数据增强的方法来增加图片的数量以及多样性。

表 1 药材数据集样本分布
Tab. 1 Sample distribution of herbal dataset

名称	数量	每张图片大小/KB	数据集大小/MB	训练集	验证集
人参	487	35~2530	55.3	390	97
党参	207	31~2180	30.3	166	41
八角	299	33~2430	37.1	240	59
冬虫夏草	163	36~2989	21.0	131	32
枸杞	327	30~3389	32.7	262	65
石斛	153	33~3680	25.3	123	30
胖大海	509	32~3249	62.1	408	101
苦瓜	330	30~3389	38.5	265	65
茯苓	333	35~2530	29.0	267	66
茴香	152	33~2430	18.0	122	30
莲子	311	36~2989	36.9	249	62
黄连	249	30~3389	34.6	200	49

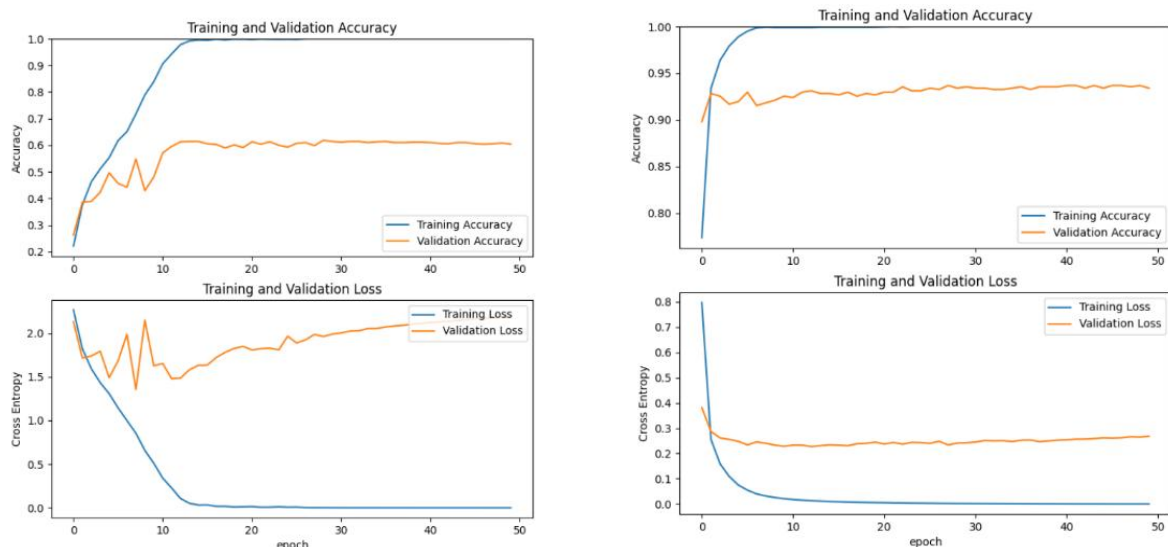
2.2 实验环境

表 2 试验平台
Tab. 2 Test platform

硬件环境	11th Gen Intel(R) Core(TM) i5-1135G7 @ 2.40GHz	2.42 GHz
操作系统	Windows 10	
试验平台	Keras	
基础配置	Anaconda Python3.9 版本	

3 实验结果及分析

本文将 12 类药材图片输入卷积神经网络进行训练，在训练过程中冻结原模型的卷积层和池化层，直接使用训练好的权重参数进行特征提取，在全连接层通过微调参数使其适用于药材识别，经过 50 轮的迭代更新，得到损失函数最小，权重最好的神经网络。为验证本文提出的基于迁移学习的模型优越性，我们分别做了传统卷积神经网络和基于迁移学习的深度可分离的卷积神经网络的对比实验，图 6 是传统卷积和迁移学习卷积的性能对比。表 3 给出了两种算法的药材识别准确率和损失函数。



(a) CNN

(b) 基于迁移学习

图 6 CNN 和基于迁移学习 mobilenet 的准确率和损失率性能对比

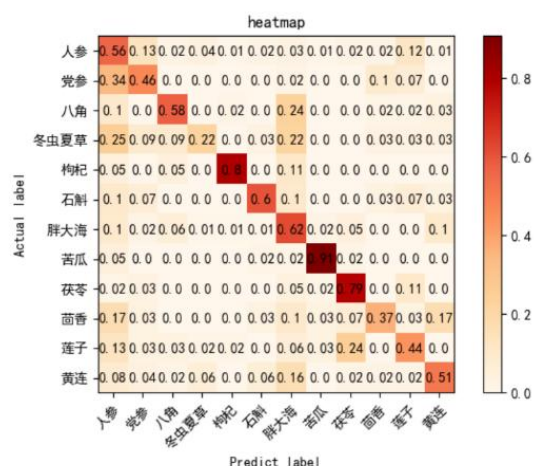
Fig. 6 Accuracy and loss rate performance comparison of CNN and migration learning based mobilenet

表 2 药材识别准确率和损失率

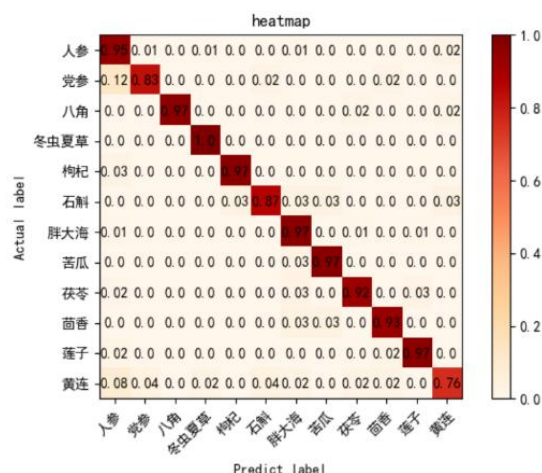
Tab. 2 Herb identification accuracy and loss rates

网络	准确率	损失率
CNN	60.41%	2.82%
迁移学习	93.40%	0.3%

由图 6 可知，在前 10 轮的训练中，传统的卷积神经网络在训练集上的准确率开始稳步上升，在训练集上的损失函数也开始逐渐收敛，但是验证集上的准确率以及损失率却出现波动现象，在训练 10 轮以后准确率和损失函数才开始收敛，但是准确率和损失函数仍然不能达到医学药材的鉴定要求，这是因为训练的次数还远远不够，未达到收敛要求。本文提出的基于迁移学习的卷积神经网络的准确率和损失率在训练开始的前 3 轮就已经开始收敛，极大地提高了训练时效。验证集上的准确率从传统的卷积神经网络的 60% 上升到 93%，损失率从原来的 2% 降低至 0.25%，其原因是使用了迁移学习，本文在模型训练时冻结了 MobileNet v2 模型中卷积层和池化层的参数，使其不参与新模型的训练，直接利用预训练模型提取到的特征参数进行学习，提高了模型训练的效率，减少了训练的时间。其次是 MobileNet v2 中加入了深度可分离卷积的思想，参数量有所降低，网络模型更加轻量化，使其能够达到移动设备实时性的要求。



(a) CNN



(b) 基于迁移学习

图 7 CNN 和基于迁移学习 mobilenet 的正确和错误识别的概率图

Fig. 6 Probability plots of correct and incorrect identifications for CNN and migration learning based mobilenet

图 7 中显示了 12 类中药材被准确识别和错误识别的概率，(a) 是传统卷积神经网络的识别概率图，其中有很多种类的药材识别概率不超过 50%，远不能达到药材识别大面积使用的精度。(b) 是基于迁移学习的识别概率图，由图 (b) 看出，精度已经明显提升，特征明显的药材如冬虫夏草全部正确分类，胖大海，八角等正确分类的概率在 97%，但对于特征相似的药材，如党参和人参这两类易混淆的药材被正确分类的概率是 83% 和 95%^[16]，人参中有 1% 的概率被认为是党参，黄连被正确分类的概率仅为 76%，有 8% 的概率被误认为人参，4% 的概率被误认为党参。

图 8 是将本文中提到的基于 MobileNet V2 模型预测结果进行可视化测试，取一组药材照片传入模型中进行分类识别，可以看出，各种药材模型都能准确的识别，因此，本文提出的模型可以帮助中医研究者以及居民完成药材识别，具有学术意义和实际应用价值^[17]。





图 8 药材识别效果图

4 结论与展望

本文对比了传统卷积神经网络和基于迁移学习的算法性能，搭建了一种基于 MobileNet V2 预训练模型的中药材图像识别系统。通过迁移学习对传统卷积神经网络进行优化，采用微调参数的方式在模型输出端使用了 softmat 激活函数以及 adam 优化算法进行药材识别，实验结果表明我们所提出的方法在 12 类药材中可以达到 93%

的精度,说明了方案的有效性,可以应用于实际药材识别的场景中。但是对于人参,黄连等易混淆药材,识别精度还有待提高,未来仍要做大量研究来提高精度。

参 考 文 献

- [1] Tu Y .Artemisinin-A Gift from Traditional Chinese Medicine to the World (Nobel Lecture).[J]. Angewandte Chemie, 2016, 47(41):10210-10226.
- [2] R Li, Y Hou, J Huang, et al. Lianhuaqingwen exerts anti-viral and anti-inflammatory activity against novel coronavirus (SARS-CoV-2)[J]. Pharmacological Research, 2020, 156:104761.
- [3] Xuling, Wu, Wei, et al. Chinese Herbal Medicine Classification Based on BP Neural Network[J]. Journal of software, 2014, 9(4):938-943.
- [4] Sun X , Qian H , Deng Z. Chinese Herbal Medicine Image Recognition and Retrieval by Convolutional Neural Network[J]. Plos One, 2016, 11(6):e0156327.
- [5] 高淑萍, 赵清源, 齐小刚,等. 改进 MobileNet 的图像分类方法研究[J]. 智能系统学报, 2021, 16(1):10.
- [6] 李林, 李小舜, 吴少智. 基于迁移学习和显著性检测的盲道识别[J]. 计算机工程与应用, 2018.
- [7] 宋益盛,林志杰.基于迁移学习和数据增强技术的物种识别[J].现代计算机,2019(14):57-63.
- [8] 吕梦棋, 张芮祥, 贾浩,等. 基于改进 ResNet 玉米种子分类方法研究[J]. 中国农机化学报, 2021, 42(4):7.
- [9] 邵星灵,王宏伦.基于改进 sigmoid 函数的非线性跟踪微分器[J].控制理论与应用,2014,31(08):1116-1122.
- [10] Cao Y, Gu Q. Generalization Error Bounds of Gradient Descent for Learning Over-Parameterized Deep ReLU Networks[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 34(4):3349-3356.
- [11] 万磊,佟鑫,盛明伟,秦洪德,唐松奇.Softmax 分类器深度学习图像分类方法应用综述[J].导航与控制,2019,18(06):1-9+47.
- [12] 任进军,王宁.人工神经网络中损失函数的研究[J].甘肃高师学报,2018,23(02):61-63.
- [13] Kingma D , Ba J . Adam: A Method for Stochastic Optimization[J]. Computer Science, 2014.
- [14] 庄奕珊. 基于深度神经网络的中药材识别[D].华南理工大学,2018.
- [15] 刘南杉,裴云强,蒋皓,韩永国,吴亚东,王赋攀,易思恒.基于 VD-MobileNet 网络的 WebAR 生活垃圾分类信息可视化方法[J/OL].图学学报:1-11[2022-05-05].<http://kns.cnki.net/kcms/detail/10.1034.T.20220314.1548.002.html>
- [16] Wang W, Tian W, Liao W, et al. Identifying Chinese Herbal Medicine by Image with Three Deep CNNs[C]// CCEAI 2021: 5th International Conference on Control Engineering and Artificial Intelligence. 2021.
- [17] 陈智超,焦海宁,杨杰,曾华福.基于改进 MobileNet v2 的垃圾图像分类算法[J].浙江大学学报(工学版),2021,55(08):1490-1499.

MobileNet V2 migration learning based image recognition of Chinese herbs

Liu Xuechun, Liu Daming*, Chang jiaxin, Wang Bo

(School of Physics and Electrical and Electronic Engineering, Ningxia University, Yinchuan, 750000, China)

(Key Laboratory of Intelligent Sensing of Desert Information, Ningxia University, Yinchuan, 750000, China)

Abstract: Chinese herbs play an important role in the prevention and control of human diseases, yet the general public has too little knowledge of herbs and may misuse them with uncontrollable consequences. Therefore, accurate identification of herbs is an urgent task. In this paper, an improved deep convolutional neural network is applied to herb identification, and a migration learning system for herb identification based on the MobileNet V2 model is proposed. The model improves the standard convolution into a deep separable convolutional form and is pre-trained on the ImageNet dataset. By fine-tuning the parameters of the pre-trained model of MobileNet V2 several times, using softmax activation function in the fully connected layer to achieve herb classification, using cross-entropy loss function to quantitatively express the difference between the ideal model and the ideal model, and finally using adam optimisation algorithm to achieve optimal gradient descent, the recognition of 12 classes of herbal images was achieved, and the final accuracy on the validation set could reach 93%, while The lightweight model of MobileNet V2 also saves more than two times in training time compared with ordinary convolutional neural networks, and the robustness and generalization in herbal medicine recognition are greatly improved. Thus, it shows that the herbal image recognition system in this paper has good practical application scenarios.

Keywords: deep convolutional neural networks; identification of medicinal herbs; MobileNet V2; Transfer learning