文章编号:2096-1472(2023)09-0038-04
DOI:10.19644/j.cnki.issn2096-1472.2023.009.007

基于改进型 AlexNet 的中药材图像识别

李万虎,吴丽丽

(甘肃农业大学信息科学技术学院,甘肃 兰州 730070)
✉ 894083430@qq.com; wull@gsau.edu.cn

摘 要:文章以百部、白芍、枸杞、黄精、姜黄、蒲黄、蛇床子、益母草、郁金和枳壳片共 10 种中药材为研究对象,使用 Python 爬虫算法构建数据集,使用随机缩放、随机剪切、水平翻转对数据集进行增强,对 AlexNet 网络模型进行改进,具体改进内容为在原 AlexNet 网络模型的第三层(卷积层)之后插入一层,为新的第三层,同时在原 AlexNet 网络模型中引入岭回归和迁移学习,建立基于改进型 AlexNet 网络模型的 10 种中药材的图像识别方法,该模型的平均识别准确率达 95.4%;数据集足够大时,可以有效地提高图像识别准确率,该模型也可应用于绝大部分需要识别的中药材类别的场景。

关键词:卷积神经网络;AlexNet;中药材;图像识别
中图分类号:TP391 **文献标志码**:A

Image Recognition of Chinese Medicinal Materials Based on Improved AlexNet

LI Wanhu, WU Lili

(*College of Information Science and Technology, Gansu Agricultural University, Lanzhou* 730070, *China*)
✉ 894083430@qq.com; wull@gsau.edu.cn

Abstract: In this paper, a total of 10 Chinese herbal medicines, including Radix stemonae, Paeonia lactiflora palls, Wolfberry, Rhizoma polygonati, Turmeric, Cattail pollen, Fructus cnidii, Motherwort, Curcuma turmeric, and Aurantium aurantium, are taken as the research objects, Python crawler algorithm is used to build data sets, and random scaling, random clipping, and horizontal flipping are used to enhance the data sets. AlexNet network model is improved by inserting one layer after the third layer (convolutional layer) of the original AlexNet network model, and it becomes a new third layer. At the same time, ridge regression and transfer learning are introduced into the original AlexNet network model to establish an image recognition method for 10 kinds of Chinese medicinal materials based on the improved AlexNet network model. The average recognition accuracy of this model is 95.4%. When the data set is large enough, it can effectively improve the accuracy of image recognition. The proposed model can also be applied to most of the scenarios where the categories of Chinese herbal medicine need to be identified.

Key words: convolutional neural network; AlexNet; Chinese medicinal materials; image recognition

0 引言(Introduction)

我国是中药材的发源地,药材资源丰富,并且 95%以上的中药材为植物药材。中药历经数千年发展,展现了它强大的生命力和我国医药特色优势。新型冠状病毒感染暴发以来,中医药一直在抗击感染中发挥了重要作用,一些中医药疗法在病例的救治和康复中起到良好效果。因此,我们要坚定不移地推动中医药的发展。

目前,传统图像处理和卷积神经网络均可用于中药材识别。孙鑫等[1]通过构建白薇等 50 种常见中药饮片图像数据库,运用 Softmax 损失训练网络模型,平均识别精度达到 70%。屈凌波等[2]利用人工神经网络对中药进行模式识别,预测率最高能达到 93.7%。刘勍等[3]介绍了图像处理技术对中药材的检验过程,然后对中药材图像鉴定与识别技术面临的问题及发展前景做了探讨与展望。徐飞等[4]通过正交试验优化卷积神经网络,其图像识别准确率为 90.9%。刘加峰等[5]建立的中药饮片图像检测识别系统,平均识别率高于 80%,但只能识别三种中药饮片。

分析以上研究可知,在对中药材识别分类研究上,上述方

收稿日期:2023-02-10
基金项目:甘肃省科技计划项目(20CX9NA095)

www.rjgczz.com

法在特定的中药鉴定场景中取得了较好的效果，但对细节特征的关注度不高，鉴别准确度仍然不足。同时，网络模型设计复杂，难以迁移，其鉴定结果不可量化，与实际应用存在差距。本文提出的基于改进型 AlexNet 的中药材图像识别模型可以解决以上问题。

1　材料与方法(Materials and methods)

1.1　数据集

本研究以百部(Radix stemonae)、白芍(Paeonia lactiflora palls)、枸杞(Wolfberry)、黄精(Rhizoma polygonati)、姜黄(Turmeric)、蒲黄(Cattail pollen)、蛇床子(Fructus cnidii)、益母草(Motherwort)、郁金(Curcuma turmeric)、枳壳片(Aurantium aurantium)共 10 种中药材为对象，数据集通过 Python 爬虫算法爬取百度图片里的相关关键字，每一类选取 500 张，共 5 000 张构成数据集，按照占图像总数量的 80%、10%、10%的比例划分出训练集 4 000 张、验证集 500 张、测试集 500 张，将该数据集命名为 CMM，如图 1 所示[6]。

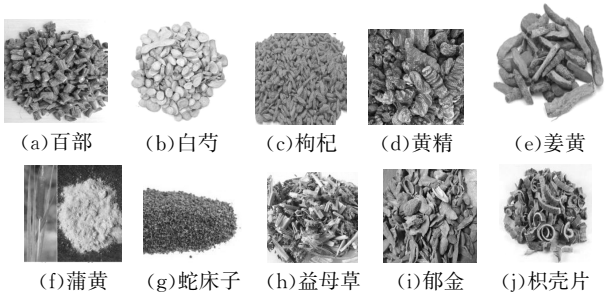


图 1　数据集展示

Fig. 1 Data set presentation

1.2　数据增强

由于数据量较小，为了避免训练过拟合，因此需要对原始数据进行增强处理。本研究使用 Keras(由 Python 编写的开源人工神经网络库)中的 ImageDataGenerator 函数对图片进行随机缩放、随机剪切、水平翻转，对数据进行扩增，再经过数据筛选和数据分布均衡调整，最终构成每类 1 000 张，共 10 000 张图片的新数据集，增强后的数据集命名为 CMM1，按照各占图像总数量的比例进行划分，训练集占比为 80%，即 8 000 张，验证集占比为 10%，即 1 000 张，测试集占比为 10%，即 1 000 张，如表 1 所示。

表 1　实验数据集

Tab.1 Experimental data set

编号	中药材名称	英文名	图像数量/张
A1	百部	Radix stemonae	1 000
A2	白芍	Paeonia lactiflora palls	1 000
A3	枸杞	Wolfberry	1 000
A4	黄精	Rhizoma polygonati	1 000
A5	姜黄	Turmeric	1 000
A6	蒲黄	Cattail polle	1 000
A7	蛇床子	Fructus cnidii	1 000
A8	益母草	Motherwort	1 000
A9	郁金	Curcuma turmeric	1 000
A10	枳壳片	Aurantium aurantium	1 000

1.3　AlexNet 模型

目前，有多种网络模型可用于图像分类，例如 AlexNet 卷积神经网络、VGGNet 卷积神经网络、ResNet 残差卷积神经网络等。查阅相关文献，分析评估参数量、可用性、计算资源需求量、易用性等因素，最终将 AlexNet 网络模型作为本研究的基础模型，其模型结构如图 2 所示。

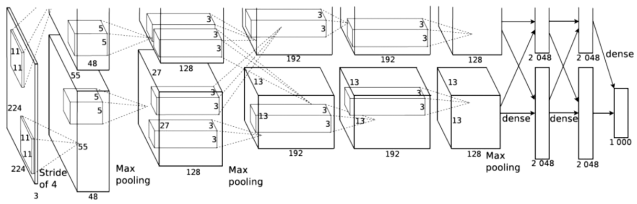


图 2　AlextNet 网络模型

Fig. 2 AlextNet network model

AlexNet 是 2012 年由 Alex Krizhevsky 提出的，它以巨大优势赢得了 ImageNet 2012 图像识别挑战赛，自此以后，卷积神经网络在图像识别领域得到迅猛发展[7]。AlexNet 网络模型由 5 个卷积层(Conv)、3 个最大池化层(Max Pooling)和 3 个全连接层(Dense)构成。卷积层和最大池化层是交替排布的，因为不同卷积层有不同的卷积核，所以提取低层次特征的能力也有所不同。AlexNet 网络模型的优势在于可以利用双 GPU 进行网络加速训练，这跟单 GPU 训练学习相比，学习速度大大提升。AlexNet 网络模型没有使用传统的 Sigmiod 函数而是使用 ReLU 函数，既解决了梯度弥撒问题，也加快了学习速度。LRN 局部响应归一化是在 ReLU 函数后对局部神经元建立竞争机制，抑制了响应较小的神经元，加强了网络的泛化本领。Hinton 在 2012 年的 AlexNet 网络模型中给出其具体的计算如公式(1)所示：

$$b^i_{x,y}=a^i_{x,y}/(k+\alpha\sum_{j=\max(0,i=\frac{n}{2})}^{\min(N-1,i+\frac{n}{2})}(a^j_{x,y})^2)^\beta \tag{1}$$

公式(1)中，a 为卷积层后的输出结果，它是一个四维数组，N 是通道数，n 为相邻的卷积核，k 是偏差，α 和 β 是自定义的值[8]。

1.4　岭回归模型

岭回归(Ridge Regression)分析用于处理共线性的数据，它是对最小二乘估计法的一种改进方程，也是一种有偏估计的回归方法[9]；它不考虑最小二乘法的无偏性，而是通过损失一些信息以此获得一个新的回归系数[10]。岭回归分析虽然降低了数据精度，但是更贴近实际应用，因此是一种值得信任的回归方法，对于中药材图像数据的拟合效果比最小二乘法更好。

对于传统的最小二乘法线性回归，它的代价函数如公式(2)所示：

$$RSS=\sum_{i=1}^{n}(y_i-\beta_0-\sum_{j=1}^{p}\beta_j x_{ij})^2 \tag{2}$$

公式(2)是通过改变拟合系数 β 的大小使 RSS 值最小，而岭回归分析是在传统的最小二乘线性回归方程中加上正则项($L2$ 范数)，用此正则项对参数进行改进。引入正则化项 $L2$ 范数后，代价函数如公式(3)所示：

$$R=RSS+\lambda\sum_{j=1}^{p}\beta_j^2=\sum_{i=1}^{n}(y_i-\beta_0-\sum_{j=1}^{p}\beta_j x_{ij})^2+\lambda\sum_{j=1}^{p}\beta_j^2 \tag{3}$$

公式(3)中，λ 是一个非负的调节参数，当 $\lambda=0$ 时，结果与传统的最小二乘法一致，没有任何改变。当 $\lambda\to+\infty$ 时，RSS 占整个代价函数的比重非常小，正则项变得非常大。当拟合系数 β 的值无限接近于零时，代价函数的值取最小。

www.jgcz.com

1.5 AlexNet 网络的改进设计

为了构建准确率高和收敛快的网络模型，本文提出了改进型 AlexNet 的中药材图像识别的网络模型，具体的改进设计介绍如下。

1.5.1 加入岭回归分析

我国的中药材中有 95%以上是植物药材，它们的图像特征等信息有很多的相似之处，如果使用传统模型训练，就会出现过拟合现象，导致识别准确率较低。为了解决该问题，本文研究引入岭回归分析，它可以在一定程度上弱化过拟合问题，对多重共性信息的数据具有较好的分析能力。

1.5.2 改进 AlexNet 模型

通过文献[11]发现，原 AlexNet 网络模型提取特征能力最强的是第三层和第四层，为了提高分类的精确度，对模型的改进和优化重点集中在这两层。经过多次实验发现，可以在第三层之后插入一层，即新的第三层（如图 3 所示），新的第三层与原来的第三层结构一致，由于增加了卷积操作，因此可以过滤更多的有效特征，因此特征提取能够更精确。新的第三层卷积核的大小为 3 * 3，数量为 192×2＝384 个，步长为 1，使用 ReLU 激活函数。输入特征图像先扩展一个像素，大小为 15×15；输出特征图像大小为(15－3)/1＋1＝13，即 13×13×384。

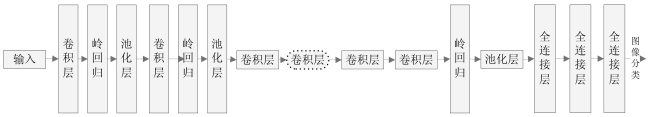


图 3 改进后的 AlextNet 模型网络结构

Fig. 3 Improved network structure of AlextNet model

1.5.3 引入迁移学习

为了进一步加快训练速度，使模型更快收敛，同时为了提高模型识别的精确度，本研究在模型训练时，引入迁移学习的训练方法。迁移学习可以使用其他大批量数据训练所得模型用于训练新数据，由于大部分数据和任务间存在相关性，因此可以减少训练步骤，缩短训练时间，并增加数据量，达到降低过拟合程度的目的。

2 结果验证(Result verification)

2.1 实验平台

本研究基于 PyTorch 平台搭建，PyTorch 主要用于自然语言处理，它是一个基于 Torch 的 Python 开源机器学习库，不但能够实现 GPU 加速，而且能支持动态神经网络。实验所用的软硬件环境如下：Windows 10 操作系统，AMD i5-10210u，16 GB 内存，Python 3.8，JetBrains PyCharm 2018.3.7 x64，GPU 租用 AutoDL AI 算力云，配置的 CPU 为 14 核 2.00 GHz，Intel(R)Xeon(R)Gold 6330，内存为 30 GB。

2.2 实验验证

本文将每个类别 800 张，共 10 个类别，总数量为 8 000 张中药材图像输入改进后的 AlexNet 网络模型中进行训练，如图 4 所示，横坐标迭代次数代表权重更新的次数，左侧纵坐标为训练过程中真实值与预测值之间产生的损失值，右侧纵坐标为 Top1 值。模型训练输入的 *batch_size* 为 128，圆圈曲线代表 Top1 值，倒三角曲线代表损失值。损失值向 0.0 趋近，Top1 值向 1.0 趋近。在训练过程中，就总体趋势而言，随着 Top1 值的不断增大，损失值不断减小，在部分阶段虽然有反弹的趋势，但是随着训练迭代次数的增加，模型最终克服了局部最优的局面，向全局最优的趋势发展。损失值和 Top1 值最终变得平缓并趋于稳定。

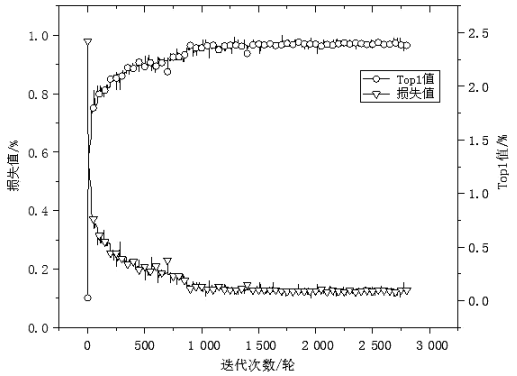
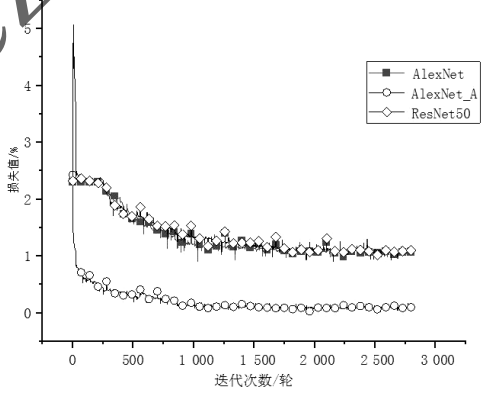


图 4 双轴图

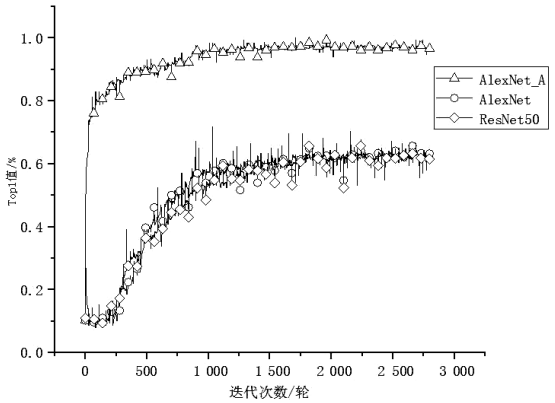
Fig. 4 Two axis diagram

3 网络模型对比实验(Comparison experiment on network models)

为了验证改进型 AlexNet 网络模型性能的优越性，采用损失值图[如图 5(a)所示]和 Top1 值图[如图 5(b)所示]与 AlexNet 原网络模型和 ResNet50 网络模型进行对比。图 5(a)中圆圈代表改进后的 AlexNet，黑色填充正方形代表 AlexNet 原网络模型，菱形代表 ResNet50 网络模型，三个网络模型都采用相同的训练参数。从图 5 可以清晰地看出，本实验提出的改进型 AlexNet 网络的平均识别准确率优于原 AlexNet 网络模型和 ResNet50 网络模型，这证明了改进型 AlexNet 网络模型具有一定的优越性。



(a)损失值对比图



(b)Top1 值对比图

图 5 损失值和 Top1 值对比图

Fig. 5 Loss value and Top1 value comparison chart

WWW.ISCZZ.COM

为了进一步评估改进后的 AlexNet 网络模型、AlexNet 原网络模型和 ResNet50 网络模型三个网络模型的表现，又选取准确率指标对三个网络模型进行评优。如图 6 所示，正方形代表改进后的 AlexNet，三角形代表 AlexNet 原网络模型，圆圈代表 ResNet50 网络模型，选取测试数据集中 10%的图片作为实验评估时选用的数据集，划分数据集时先打乱数据集后再划分，从而确保实验的说服力。在验证时共设置 100 个迭代轮数，每经历 5 个轮数开启一次验证。由图 6 可以看出，改进型 AlexNet 网络模型应用于中药材图像识别较 AlexNet 原网络模型和 ResNet50 网络模型更具优势。

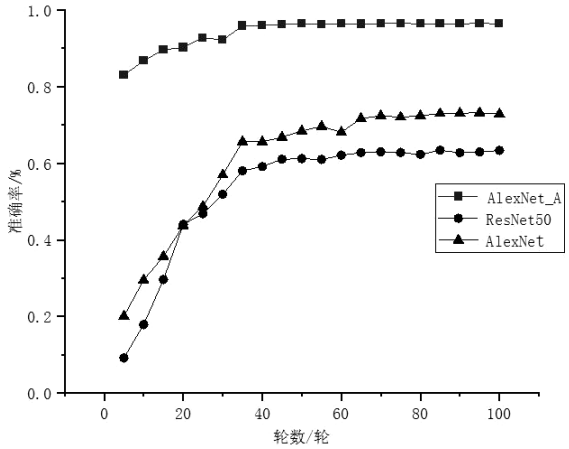


图 6　准确率对比图

Fig. 6 Accuracy comparison chart

完成模型评估后，开始对训练好的模型进行测试，测试集是数据集总数的 10%，测试结果以十行十列混淆矩阵的形式展现，如图 7 所示，横坐标为每一类中药材预测的分类百分比，纵坐标为真实分类百分比。横坐标百分比相加结果应为 1，代表对每一类中药材图像正确预测与错误预测之和。图 7 中横纵坐标重合处对角线代表分类正确的结果，颜色越深，表示分类准确率越高。

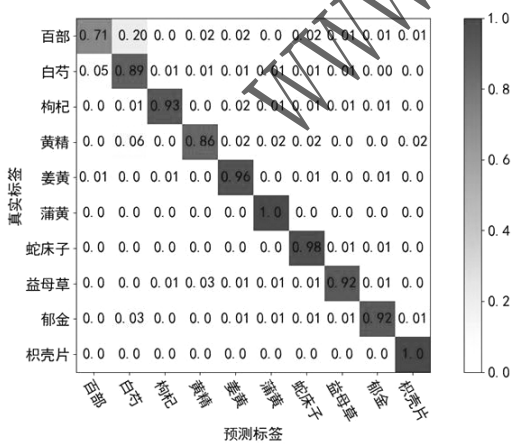


图 7　混淆矩阵

Fig. 7 Matrix of confusion

为了进一步验证训练模型的有效性与准确性，除使用混淆矩阵，还引入 Precision（精度）、Recall（召回率）、F1 Score（平衡 F 分数）三个指标评价模型，结果见表 2。

表2　网络模型评价指标

Tab.2 Network model evaluation index

分类	精度/%	召回率/%	平衡 F 分数/%
百部	94	65	80
白芍	89	87	88
枸杞	97	96	94
黄精	95	92	90
姜黄	95	90	90
蒲黄	100	100	100
蛇床子	98	96	97
益母草	92	96	94
郁金	98	100	98
枳壳片	96	92	90

4　结论(Conclusion)

本研究以百部等 10 种中药材为研究对象，构建了 10 种中药材数据集，通过改进 AlexNet 原网络模型，得到了一种用于百部、白芍、枸杞、黄精、姜黄、蒲黄、蛇床子、益母草、郁金和枳壳片等中药材图像识别的卷积神经网络模型，图像识别准确率为 95.4%，同时解决了模型难以迁移和鉴定结果不可量化的问题。下一步，计划继续提高 AlexNet 网络模型的训练速度和识别精度，主要从两个方面入手，一是增加本研究构建的 AlexNet 网络模型的中药材识别品种，二是继续探索增加新的卷积层数，并优化卷积层与池化层之间的连接。

参考文献(References)

[1] 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.

[2] 屈凌波，相秉仁，安登魁. 人工神经网络在中药模式识别中的应用[J]. 计算机与应用化学，2002，19(4)：428-431.

[3] 刘勃，刘喜平，温志贤，等. 基于图像处理技术的中药材现代检验方法[J]. 自动化与仪器仪表，2012(1)：158-160.

[4] 徐飞，孟沙，吴启南，等. 基于卷积神经网络的人参与西洋参饮片鉴别方法研究[J]. 南京中医药大学学报，2018，34(6)：621-624.

[5] 刘加峰，高子啸，段元民，等. 基于深度学习的中药材饮片图像识别[J]. 北京生物医学工程，2021，40(6)：605-608.

[6] 黄方亮，沈同平，金力. 改进的 AlexNet 卷积神经网络用于中草药叶片分类[J]. 安庆师范大学学报(自然科学版)，2020，26(3)：70-75.

[7] KRIZHEVSKY A，SUTSKEVER I，HINTON G E. ImageNet classification with deep convolutional neural networks[J]. Communications of the ACM，2017，60(6)：84-90.

[8] 郭原东，雷帮军，聂豪，等. 基于深度学习的智能高精度图像识别算法[J]. 现代电子技术，2021，44(4)：173-176.

[9] 胡良平. 岭回归分析[J]. 四川精神卫生，2018，31(3)：193-196.

[10] TIBSHIRANI R. Regression shrinkage and selection via the lasso[J]. Journal of the Royal Statistical Society：Series B (Methodological)，1996，58(1)：267-288.

[11] HINTON G，DENG L，YU D，et al. Deep neural networks for acoustic modeling in speech recognition：the shared views of four research groups[J]. IEEE Signal Processing Magazine，2012，29(6)：82-97.

作者简介：

李万虎(1990-)，男，硕士生。研究领域：图像识别，人工智能与数据挖掘。

吴丽丽(1979-)，女，博士，副教授。研究领域：图像识别，人工智能与数据挖掘。本文通信作者。

www.rjgcz.com