

【引文格式】赵彦辉, 范欣宁, 张建逵, 等. 基于 DeepLearning4J on Spark 深度学习方法在药用植物图像识别中应用初探[J]. 中国中医药图书情报杂志, 2018, 42(5): 18-22.

基于 DeepLearning4J on Spark 深度学习方法 在药用植物图像识别中应用初探

赵彦辉¹, 范欣宁², 张建逵², 谢明^{2*}

1. 辽宁中医药大学图书馆, 辽宁 沈阳 110032; 2. 辽宁中医药大学药学院, 辽宁 大连 116600

摘要: 目的 基于卷积神经网络, 使药用植物的图像识别自动化、智能化, 从而提高药用植物识别的效率。方法 使用中国植物图像库搜集药用植物图像, 进行灰度化、角度随机旋转等批量处理。采用卷积神经网络图像识别技术, 结合 DeepLearning4j 框架, 对模型参数进行优化调试。结果 使用 DL4J-example、AlexNet、VGG16 模型, 小样本单一背景药用植物图像识别正确率分别达到 68.00%、70.00%、24.00%, 大样本复杂背景药用植物图像识别正确率分别达到 48.87%、60.00%、91.35%。结论 图像识别技术的使用提高了复杂背景下药用植物图像识别的正确率和效率, 为药用植物的识别提供了新方法, 识别正确率高低与激励函数、迭代次数、卷积层数及参与训练的图像数量直接相关。

关键词: 图像识别; 卷积神经网络; 深度学习; 迁移学习; 药用植物

DOI: 10.3969/j.issn.2095-5707.2018.05.005

中图分类号: G250.7 文献标识码: A 文章编号: 2095-5707(2018)05-0018-05

Discussion on Application of Deep Learning Method in Medical Plant Image Recognition Based on DeepLearning4J on Spark

ZHAO Yan-hui¹, FAN Xin-ning², ZHANG Jian-kui², XIE Ming^{2*}

(1. Library of Liaoning University of Traditional Chinese Medicine, Shenyang 110032, China; 2. College of Pharmacy, Liaoning University of Traditional Chinese Medicine, Dalian 116600, China)

Abstract: Objective To automate and intelligentize medical plant image recognition based on the convolution neural network. To improve the efficiency of medical plant recognition. **Methods** Chinese plant image database was used to collect medical plant images, and the batch processing of gray scale and angle random rotation were carried out. Convolution neural network image recognition technology was used and DeepLearning4J frame was combined. Model parameters were optimized and debugged. **Results** DL4J-example, AlexNet and VGG16 models were used. The accuracy rates of image recognition for small samples of single background medical plants reached 68.00%, 70.00% and 24.00%, respectively, and the accuracy rates of image recognition of medical plants in complex background reached 48.87%, 60.00% and 91.35%, respectively. **Conclusion** The use of image recognition technology can improve accuracy and efficiency of medical plant image recognition under complex background, and provide a new method for recognition of medical plants. The recognition accuracy rate is directly related to the excitation function, the number of iterations, the number of convolution layers, and the number of images participating in the training.

Key words: image recognition; convolution neural network; deep learning; transfer learning; medical plants

基金项目: 2014 年辽宁省高等学校创新团队 (WT2014004)

第一作者: 赵彦辉, E-mail: 644735344@qq.com

*通讯作者: 谢明, E-mail: x6m6@163.com

传统的药用植物识别主要依靠专业人员通过眼看、手摸、口尝、鼻闻的方式进行^[1],这种方式主要依赖操作者的专业水平和经验,因此不同的操作者识别药用植物有一定的主观性,其正确率有所不同。而图像识别技术能解决人工识别主观性强等问题,可以降低药用植物识别的难度。图像识别是指利用计算机对图像进行处理、分析和理解,以识别各种不同模式的目标和对象的技术,是信息科学和人工智能的重要组成部分^[2]。近年来,深度学习已经成为图像识别领域的热点。为解决人工识别药用植物效率低、主观性强的难题,本研究基于卷积神经网络(Convolutional Neural Network, CNN),探究药用植物图像识别的方法,构建自动化的识别系统,并利用搜集的 5 种药用植物图像对系统进行识别训练,为药用植物的识别提供新方法、新思路。

1 实验材料及开发工具

1.1 操作系统及分布式实时计算环境

Windows 7 sp1 64 位(客户端), CentOS-Linux release 7.4-64 位, Hadoop2.6.0-CDH5.5.2, Spark2.2.1。

1.2 硬件环境

Intel CORE i5-42100-2.4GHZ/6G 内存/1T 硬盘笔记本电脑 1 台;清华同方超越 E550, 2 核 Intel(R) Pentium(R) CPU G645@2.90GHz/8G/500G 2 台; Dell 台式机 2 核 Intel(R) Celeron(R) CPU G1820@2.70GHz/6G/500G 1 台; Dell 虚拟机 Vmware5.5.0-4 核 Intel(R) Xeon(R) CPU E3-1220 V2@3.10GHz/7G/46G 1 台;曙光 Vmware5.5.0-15 核 AMD Opteron(TM) Processor 6212/8G/50G 1 台。

1.3 软件环境

Eclipse Oxygen 集成开发平台、JAVA 语言环境、DeepLearning4j 深度学习框架、Adobe Photoshop CC 软件。

2 实验方法与结果

2.1 卷积神经网络框架设计

本文所构建的 CNN 构架主要由输入层、卷积层、池化层、全连接层和输出层 5 个部分组成(见图 1)。

2.1.1 输入层 若输入对象是图像,那么输入层的原始输入数据为图像的像素值。本研究中原始输入数据即为 256×256 像素大小的图像, JPG 格式, RGB3 通道。

2.1.2 卷积层 卷积层由多个特征面组成,每个特征面由多个神经元组成,每个神经元通过卷积核与

上一层特征面的局部区域相连^[3]。卷积层的主要作用为提取输入数据特征,其过程如图 2 所示。几种滤波(即核)在已被分解为彩色值的图像上滑动,将核内值的滤波乘积加和,并将加和的值返回,得到卷积特征。卷积层中卷积核数量越多,提取的输入数据特征就越多。本实验通过更改核的值来提取不同类型的特征。

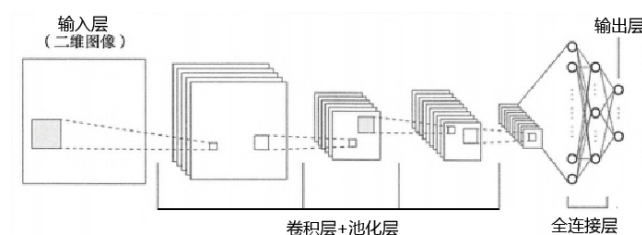


图1 卷积神经网络设计构架示意图

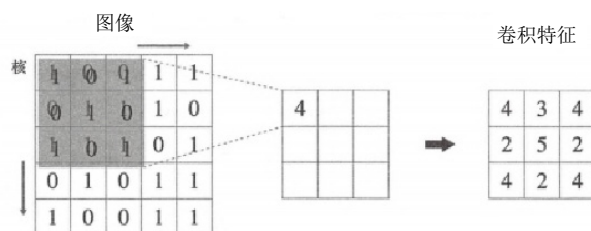


图2 卷积特征提取示意图

2.1.3 池化层 由多个特征面组成,它的每一个特征面均唯一对应于其上一层的一个特征面,且池化层的神经元也与其输入层的局部接受域相连^[3]。池化层自身并不会训练或者学习,仅仅是把卷积层传播过来的图像进行池化(下采样)。其主要作用是在保留有用信息的基础上减少数据的处理量,加快网络训练的速度。

2.1.4 全连接层 后面层的神经元节点和前一层的神经元节点连接,而同一层神经元节点间无连接。每一层神经元节点通过连接线上的权值进行前向传播,加权组合得到下一层神经元节点的输入。其实质为多层感知机的隐含层。本文中,为了提升卷积神经网络性能,降低模型的计算复杂度,在全连接层选用了 Relu(修正线性单元)激活函数。计算公式如式(1)。除此之外,在网络训练的过程中,为了避免过拟合,本文在全连接层使用了正则化方法——dropout 技术,有研究表明,此方法可使隐含层神经元的输出值以 50%的概率变成 0^[4]。因此能有效降低神经元间的复杂性,忽略无效细节,从而提升神经元学习特征的鲁棒性。

$$\text{Relu}(x) = \begin{cases} x & \text{if } x > 0 \\ 0 & \text{if } x \leq 0 \end{cases} \quad (1)$$

2.1.5 输出层 可根据具体应用任务来设定输出层的神经元节点数,本文训练 5 种药用植物图像,因此输出节点数设置为 5。

2.2 DeepLearning4j 深度学习框架搭建

2.2.1 安装集成开发环境 下载并安装 JDK9.0.1、Eclipse Oxygen 版 for windows (64Bit) 软件。设置系统环境变量, JAVA_HOME 为 D:/JAVA/JDK-9.0.1, PATH 增加 %JAVA_HOME%\bin, CLASSPATH 为 .;%JAVA_HOME%\lib。测试运行环境,运行 CMD、java-version、javac 回报版本和编译信息。安装 Eclipse 软件至 D:/ECLIPSE。

2.2.2 安装 DeepLearning4J 访问 <https://deeplearning4j.org/>, 下载 ZIP 文件到 D:/DL4J, 安装 MAVEN 环境, 导入 Maven Existing Maven Projects, 打开 dl4j-examples。

2.2.3 搭建 Hadoop 集群和 Spark 集群 测试任务提交命令, #spark-submit -- class cn.edu.lnntcm.nd4jonSpark.FitFromFeaturized -- master spark://storm6: 7077 -- deploy-mode client -- driver-memory 6g -- executor-memory 4g hdfs://ns1/spark_lib/ Nd4jTestOnMllib-0.0.1-SNAPSHOT.jar。

2.3 图像的搜集与处理

2.3.1 植物图像的搜集 图像特征提取是药用植物图像识别的核心, 直接影响最终的识别效果^[5]。本实验选取银杏、牛蒡、薄荷、知母、芍药 5 种药用植物。

银杏 (*Ginkgo biloba* L.), 幼树树皮近平滑, 浅灰色, 大树之皮灰褐色。叶互生, 在长枝上辐射状散生, 在短枝上 3~5 枚成簇生状, 扇形, 淡绿色, 具多数叉状细脉。

牛蒡 (*Arctium lappa* L.), 茎直立, 带紫色, 上部多分枝。基生叶丛生, 大型, 有长柄; 茎生叶广卵形或心形。头状花序多数, 排成伞状; 总苞球形, 总苞片披针形; 花淡红色, 管状。

芍药 (*Paeonia lactiflora* Pall.), 茎直立, 上部分枝。叶互生, 近革质, 二回三出复叶, 小叶窄卵形。花两性, 数朵生茎顶和叶腋, 白色, 有时基部具深紫色斑块或粉红色。

薄荷 (*Mentha haplocalyx* Briq.), 茎四棱形。单叶对生, 叶片长卵形或长圆形。轮伞花序腋生, 花冠二唇形, 淡紫色或白色, 4 裂。

知母 (*Anemarrhena asphodeloides* Bunge.), 叶基生, 丛出; 线形, 基部扩大成鞘状。花茎直立,

穗状花序稀疏而狭长, 花常 2~3 朵簇生; 花绿色或紫堇色; 花被片 6, 宿存, 排成 2 轮。

在中国植物图像库 (PPBC, <http://www.plantphoto.cn/>) 中搜集具有上述识别特征的植物图像。经筛选, 本文选取了银杏图像 2483 个, 牛蒡图像 1494 个, 薄荷图像 2120 个, 知母图像 591 个, 芍药图像 2318 个, 共计 9006 个图像。以上图像均经过辽宁中医药大学药学院张建逵副教授鉴定。

2.3.2 植物图像的预处理 为加强卷积神经网络对图像特征的学习效果, 利用 Adobe Photoshop CC 软件对这些图像进行批量处理。首先去掉图像中的水印和干扰元素, 只保留识别所需的特征部位, 然后进行灰度化处理, 并归一化像素值为 256×256 大小。为增加图像的数据量、更好地检验卷积神经网络对几何变换、形变等一定程度上的不变性^[6], 本实验将图像进行了随机角度旋转 ($5 \sim 180^\circ$) 变换, 并将随机旋转角度后的图片背景处理为白色。

经过以上处理, 图像数量增长了 1 倍, 变为 18 012 个, 将这些图像构建为大样本复杂背景数据库, 命名为 L (Large) 数据库, 图像见图 3。另外, 为了探究图像数量及背景复杂程度对模型训练效果的影响, 还建立一个小样本简单背景数据库, 即从 5 类植物图像中各随机抽取 25 个图像, 进行单一背景处理, 命名为 S (Small) 数据库, 含单一背景图像 125 个 (见图 4)。



图 3 大样本复杂背景图片示例 (Large 数据库)



图 4 小样本单一背景图片示例 (Small 数据库)

2.3.3 实验模型参数调试 本文分别采用了 DL4J-example、AlexNet 和 VGG16 共 3 种神经网络模型进行有监督学习的训练^[7], 在每次模型训练过程中先将数据库中的图像进行标签, 然后随机选取其中 80% 图片作为训练集, 20% 作为测试集。

为了较好地评价所建立的模型是否具有较强的鲁棒性, 本实验统计了 5 类植物分类的正确率

(accuracy) , 计算了每个类别的查准率 (precision)、查全率 (recall) 及查全率与查准率的加权平均分数 (F_1) , 取其平均值作为模型评价标准。具体计算表达式见式(2)、式(3)、式(4), 其中, n_{ij} 表示类别 i 预测为第 j 类的样本数, n_{ii} 表示类别 i 预测为第 i 类的样本数, n_{ji} 表示类别 j 预测为第 i 类的样本数, n_{cl} 表示样本类别数, 本实验值为 5。

参照查阅的文献, 在 DL4J-examples 源码基础上构建模型, 进行 1~7、10~12 号实验。在第 8、13 号实验引入 AlexNet 网络模型, 在第 9、14 号实验引入 VGG16 网络模型。对模型中的参数进行调

试, 并对其性能进行检测。1~8、10~13 号实验为单机运行所得, 9、14 号实验是在 Spark 集群上运行所得。实验结果见表 1。

$$\text{precision}_{(i)} = \frac{n_{ii}}{\sum_{j=1}^{n_{cl}} n_{ji}}; \quad \text{recall}_{(i)} = \frac{n_{ii}}{\sum_{j=1}^{n_{cl}} n_{ij}} \quad (2)$$

$$F_{1(i)} = 2 \times \frac{\text{precision}_{(i)} \times \text{recall}_{(i)}}{\text{precision}_{(i)} + \text{recall}_{(i)}} \quad (3)$$

$$F_1 = \frac{\sum_{i=1}^{n_{cl}} F_{1(i)}}{n_{cl}} \quad (4)$$

表 1 模型参数设置与模型评估得分

编号	训练集来源	模型参数				评估得分			
		输入图像尺寸	激励函数/网络	迭代次数	卷积层数	正确率 (%)	查准率 (%)	查全率 (%)	加权得分
1	S 数据库	100×100	Relu	1	2	*	*	*	*
2	S 数据库	60× 60	Relu	1	2	22.18	44.43	22.26	12.80
3	S 数据库	50× 50	Relu	1	2	40.91	45.00	42.00	46.13
4	S 数据库	25× 25	Relu	1	2	**	**	**	**
5	S 数据库	50× 50	Relu	10	2	36.36	37.45	38.00	32.76
6	S 数据库	50× 50	Tanh	1	2	31.86	36.11	31.86	31.04
7	S 数据库	50× 50	Relu	50	7	68.00	77.00	68.00	67.87
8	S 数据库	100×100	AlexNet	50	13	70.00	71.00	70.00	69.92
9	S 数据库	244×244	VGG16	50	16	24.00	23.33	24.00	23.64
10	L 数据库	50× 50	Relu	1	7	22.18	44.43	22.26	12.80
11	L 数据库	50× 50	Relu	1	2	32.71	67.42	32.83	34.84
12	L 数据库	50× 50	Relu	15	7	48.87	54.00	48.81	47.70
13	L 数据库	100×100	AlexNet	50	13	60.00	75.00	60.00	69.63
14	L 数据库	244×244	VGG16	3	16	91.35	91.80	91.33	91.33

注: *运行结果出错, 内存溢出。**运行结果出错, 图像尺寸不匹配。

将本文所构建的深度学习模型对单一背景和复杂背景图像识别的正确率分别与文献中查阅到的 CNN+SVM 模型^[8] (8 层 CNN+SVM 分类器, 其中 SVM 选取默认参数 $C=10$, 为 SVM 分类器的惩罚变量值; $\sigma=0.038\ 461\ 54$, 是径向基 RBF 形式核函数的变量值) 及 PCANet 模型^[8] ($L1=10$, $L2=8$, 指层滤波核数分别为 10、8) 的图像识别正确率进行对比, 结果见表 2。

表 2 单一背景和复杂背景下 5 种卷积神经网络模型图像识别正确率 (%)

图片类型	PCANet	CNN+SVM	DL4J-example	AlexNet	VGG16
单一背景图片	90.49	91.11	68.00	70.00	24.00
复杂背景图片	33.80	34.38	48.87	60.00	91.35

3 分析讨论

通过实验数据可看出, 基于卷积神经网络的药

用植物图像模式识别程序的性能受激励函数、迭代次数、卷积层数及图像数量等因素的影响。

3.1 激励函数

Tanh 函数是饱和非线性函数, 能在解决梯度爆炸 (或消失) 问题的同时加快收敛速度^[9]。而 Relu 函数是不饱和非线性函数, 从第 3、6 号实验不难看出, 使用 Relu 函数后, 正确率提升了 9.05%。可见, 在卷积神经网络中选用 Relu 函数, 能够更有效地提升网络模型的性能。

3.2 迭代次数

从第 3、5 号实验的测试结果可以看出, 训练和测试集来自 S 数据库时, 迭代次数为 1 次的准确率比 10 次的高出 4.55%。再对比 10、12 号实验的测试结果可以发现, 训练和测试集来自 L 数据库时,

迭代次数为 15 次时的准确率较 1 次迭代的准确率高出 26.69%。

因此,不能简单地说明迭代次数与准确率成线性关系。对于较小数据库的训练来说,迭代次数少,会使准确率更高;但对于较大数据库的训练来说,迭代次数多会使准确率更高。

3.3 卷积层数

从 10、11、12 号实验可以看出,随着数据量的增大,只增加卷积层的数量是不够的,从 2 层卷积增加至 7 层后查准率反而下降了。因此,我们将卷积层数与迭代次数同时增加,层数加至 7 层(在第二层增加隐含层,在第七层增加池化层),迭代次数加至 15 次,结果表明,正确率有所提高。

3.4 数据量对网络性能的影响

结合上述实验结果来看,数据量越大,所构建的卷积神经网络模型也就越复杂,但是对比 5、12 号实验,不难发现,数据量的增加对于正确率的提升有一定帮助。用于训练的图像数量越多,卷积神经网络能提取、学习到的特征也就越丰富,再加上多次反复的学习,识别的准确度和效率也就越高。

3.5 单机深度学习与分布式迁移学习的差别

因为单机内存不足,利用 VGG16 模型无法运行迁移学习,故 9、14 号实验选择了分布式平台,搭建了由 5 台计算机组成的 Spark 集群环境。由单机勉强运行的 13 号实验(13 层卷积,迭代 50 次),训练了 24 747 397 个参数,运行了 7.70 h,才得出结果。而在 Spark 集群环境下,冻结 134 260 544 个参数,训练 20 485 个参数,仅运行 6.95 min,得出了 91.35%的正确率。

3.6 与其他模型对比分析

从表 2 中可以看出,与其他模型相比,本文所构建的模型在复杂背景图像的识别应用中将现有研究中的最佳正确率提高到 91.35%,使用迁移学习方

法,训练时间大大缩短。但与学习识别单一背景图像的模型相比,由于数据量较小,正确率较低。

4 小结

图像识别技术通过对图像特征提取、处理和分析,能够自动完成对图像进行描述、辨认、分类和解释,已经成为信息科学和人工智能的重要组成部分。将卷积神经网络应用到药用植物识别中,降低了药用植物识别的难度,避免了手工提取特征而导致误差累积的缺点。使用相对较少的预处理,更适用于复杂的图像特征,具有适用性强、特征提取与分类同时进行、泛化能力强、全局优化训练参数少等优点,为药用植物的识别提供了新方法、新思路。

参考文献

- [1] 叶锋,蔡光东,郑子华,等.基于多特征融合的药用植物标本识别[C]//2011 年中国智能自动化学术会议论文集(第一分册).北京,2011:5.
- [2] 龚丁禧,曹长荣.基于卷积神经网络的植物叶片分类[J].计算机与现代化,2014(4):12-19.
- [3] 周飞燕,金林鹏,董军.卷积神经网络研究综述[J].计算机学报,2017,40(6):1229-1251.
- [4] 孙俊,谭文军,毛罕平,等.基于改进卷积神经网络的多种植物叶片病害识别[J].农业工程学报,2017,33(19):209-215.
- [5] 谭亮.基于内容的中草药图像检索关键技术研究[D].杭州:浙江大学,2016.
- [6] 马珍玉,王树森,聂磊,等.基于 Matlab 的植物叶片识别研究与实现[J].内蒙古科技与经济,2016(9):48-49,51.
- [7] 张善文,黄德双.一种鲁棒的监督流形学习算法及其在植物叶片分类中的应用[J].模式识别与人工智能,2010,23(6):836-841.
- [8] 张帅,淮永建.基于分层卷积深度学习系统的植物叶片识别研究[J].北京林业大学学报,2016,38(9):108-115.
- [9] 罗晓曙.人工神经网络理论·模型·算法与应用[M].桂林:广西师范大学出版社,2005.

(收稿日期:2018-03-23)

(修回日期:2018-05-02;编辑:魏氏)