

# 基于一种轻量级卷积神经网络的植物叶片图像识别研究

李文逵, 韩俊英

(甘肃农业大学信息科学技术学院, 甘肃 兰州 730070)

✉448671882@qq.com; 3243727977@qq.com



**摘要:** 对轻量级卷积神经网络MobileNet V2的模型结构进行改进, 将深度可分离卷积中的激活函数ReLU替换成Leaky ReLU, 从而避免神经元死亡问题, 倒置残差卷积中的跨越连接添加Dropout层, 增大模型的泛化能力。实验结果表明, 预测结果的总体准确率达到91.41%, 最高精确率为95.12%, 最高召回率为97.39%, 取得较好的预测结果。这说明将MobileNet V2卷积神经网络用于植物叶片图像识别是实际可行的, 为移动端植物叶片图像识别提供了实现方法和技术支撑。

**关键词:** 植物叶片; 图像识别; MobileNet V2; 卷积神经网络; 深度学习

**中图分类号:** TP520.40 **文献标识码:** A

## Research on Plant Leaf Image Recognition based on a Lightweight Convolutional Neural Network

LI Wenkui, HAN Junying

(College of Information Science and Technology, Gansu Agricultural University, Lanzhou 730070, China)

✉448671882@qq.com; 3243727977@qq.com

**Abstract:** This paper proposes to improve the model structure of the lightweight convolutional neural network MobileNet V2. The activation function ReLU in the deep separable convolution is replaced with Leaky ReLU, thereby avoiding the problem of neuron death. A Dropout layer is added across connections in inverted residual convolution to increase the generalization of the model. The experimental results show that the overall accuracy rate of the prediction results reaches 91.41%, the highest accuracy rate is 95.12%, and the highest recall rate is 97.39%, achieving good prediction results. It shows that it is practical to use the MobileNet V2 convolutional neural network for plant leaf image recognition, and it provides an implementation method and technical support for mobile terminal realization of plant leaf image recognition.

**Keywords:** plant leaf; image recognition; MobileNet V2; convolutional neural network; deep learning

### 1 引言(Introduction)

植物是自然界的基本组成部分<sup>[1]</sup>, 人类的许多生产活动与植物密切相关, 比如农业中以植物类农作物为主<sup>[2]</sup>, 生活中的中草药、护肤品和化妆品<sup>[3-4]</sup>等都是对植物进行加工得到的。随着人们对大自然不断进行不合理的开采, 人类的许多生活生产活动对植物生态环境构成威胁, 植物物种不断灭绝。为

了保护植物多样性, 建立植物物种数据库是十分必要的, 这就需要人类进行植物分类与识别, 确认新的或者罕见的植物品种。但是由于植物科学研究的普及性不高, 通常人们能看到一些植物, 却不能快速准确地获取该植物的学名、性状、品类和价值, 使得植物保护和植物认知研究产生一定的困难, 因此植物识别是进行植物学习过程中的一项基本任务。

相较于植物的其他器官，大部分植物的叶片是一个二维的平面，具有形状结构稳定、特征性状多、区分度大、生存周期长、采集便利等特点，从叶片出发能更加快速准确地分辨植物种类。

随着计算机视觉技术的不断成熟，尤其是深度学习技术在图像识别方面的不断发展，人们已经能对植物叶片图像实现不错的识别效果。2021年，韩斌等<sup>[5]</sup>将LBP特征和Gabor特征融入AlexNet卷积神经网络，对189种叶片识别的平均正确识别率是96.37%；王维<sup>[6]</sup>提出改进的VGGNet神经网络，植物叶片识别准确率达到99.39%。2020年，边缘等<sup>[7]</sup>利用Flavia数据集，运用深度学习算法，识别准确率超过90%；朱良宽等<sup>[8]</sup>利用迁移学习获得96.57%的植物叶片识别精度。2019年，张露<sup>[9]</sup>使用AlexNet、Inception-V3、VGG-16预训练模型，在测试集上得到的准确率分别为95.31%、93.86%、95.40%。可见，卷积神经网络已经普遍运用于植物叶片的图像识别，其识别准确率可满足实用要求。

卷积神经网络(CNN)是一种深度学习模型或类似于人工神经网络的多层感知器<sup>[10]</sup>，常用来分析视觉图像，可以自动从大规模矩阵数据中学习特征，并把结果向同类型未知数据泛化。MobileNet V2是一种轻量级CNN网络，由Google团队在2018年提出，属于MobileNets系列，相比于传统卷积神经网络，极大减少了运算所需的参数量，同时也确保了准确率不会降低。它体积小特点适用于低内存的设备，反应延迟低的特点符合嵌入式设备的要求。本文设计了基于MobileNet V2的卷积神经网络，利用深度学习算法，对11种植物叶片进行识别，实现了较好的识别效果。

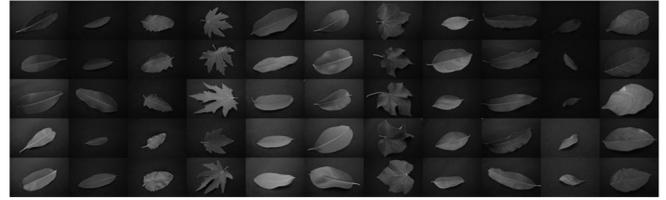
### 2 建立植物叶片样本数据库(Establishing plant leaf sample database)

植物叶片样本库选自公共数据集，包括糖胶树、阿江榄仁、罗勒、悬铃木、番石榴、蒲桃、麻风树、柠檬、芒果、石榴、水黄皮共计11类植物品种。植物叶片图像原始分辨率为6000像素×4000像素，受显存大小限制，所有图像的分辨率缩小为600像素×400像素，从中选取没有病虫害和形状相对完整的植物叶片图像作为研究对象，共选取的图像数量为2,277张，按照4:1的比例<sup>[5-11]</sup>随机抽取图片来划分成训练集和测试集，相同类别的叶片图像置于同一文件夹中。表1为不同植物种类训练集和测试集具体的叶片图像数量，图1为植物叶片样本图像数据集的示例，每个种类展示5张图片。

表1 不同种类植物叶片图像数量

Tab.1 The number of leaf images of different plants

数据集	糖胶树	阿江榄仁	罗勒	悬铃木	番石榴	蒲桃	麻风树	柠檬	芒果	石榴	水黄皮
训练集	143	176	118	82	232	223	106	127	136	230	258
测试集	36	44	30	21	45	56	27	32	34	57	64



注：从左到右依次为糖胶树、阿江榄仁、罗勒、悬铃木、番石榴、蒲桃、麻风树、柠檬、芒果、石榴、水黄皮

图1 植物叶片图像数据集样本示例  
Fig.1 Sample plant leaf image dataset

### 3 卷积神经网络(Convolutional neural network)

#### 3.1 卷积神经网络原理

图2表示卷积神经网络的一般结构，由输入层、卷积层、池化层和全连接层组成<sup>[12-14]</sup>。最左侧输入层的矩阵代表一张输入图片。卷积层<sup>[15]</sup>中每一个节点的输入是上一层图像矩阵中的一小块，该小块大小通常是3×3或者5×5，图3表示具体卷积操作。池化层的作用是缩小上一层图像矩阵的大小，将一张分辨率高的图像通过算法转化成分率低低的图像。通过池化层<sup>[15]</sup>，可以进一步缩小最后全连接层中节点的个数，从而达到减少整个神经网络参数的目的，图4表示具体池化操作。经过重复多次的卷积和池化，图像矩阵中的像素信息就被泛化成高信息含量的特征向量，在特征提取完成后，接下来由全连接层及其激活函数进行分类，全连接层常用的激活函数有ReLU、Softmax、Sigmoid等。

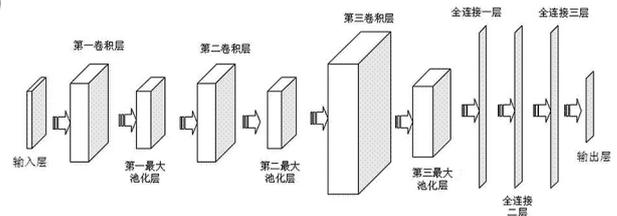


图2 卷积神经网络运算流程

Fig.2 Operation flow of convolutional neural network

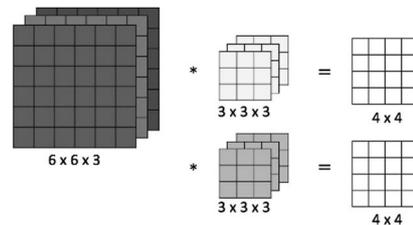
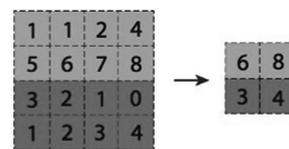


图3 卷积操作

Fig.3 Convolution process



最大池化

图4 池化操作

Fig.4 Pooling process

### 3.2 MobileNet V2卷积神经网络

MobileNet V2是一个轻量级的卷积神经网络<sup>[16]</sup>。大部分传统卷积神经网络结构单元是标准卷积(Standard Convolutions)，产生的参数量庞大，进而导致对内存资源要求高和运算量巨大，无法在嵌入式等微型电子设备上运行。而MobileNet V2网络结构是针对移动设备和嵌入式微机等而开发设计的，用深度可分离卷积(Depthwise Sparable Convolutions)<sup>[17]</sup>取代以往的标准卷积，还使用了两个超参数 $\alpha$ 和 $\beta$ 。 $\alpha$ 参数是一个倍率因子，用来调整卷积核的个数， $\beta$ 是控制输入网络的图像尺寸参数，从而大大减少了运算量和参数数量，同时有效缩减了模型体积大小。MobileNet V2的创新之处是在深度可分离卷积的基础上添加了残差连接(Residual Connection)，形成倒置残差模块(Inverted Residual Block)。传统的残差结构中是 $1 \times 1$ 卷积降维 $\rightarrow 3 \times 3$ 卷积 $\rightarrow 1 \times 1$ 卷积升维，特征图(Feature Map)的通道数是先减少再增加；而在倒置残差结构中正好相反，是 $1 \times 1$ 卷积升维 $\rightarrow 3 \times 3$ 卷积 $\rightarrow 1 \times 1$ 卷积降维，特征图的通道数是先增加再减少。如图5所示的是倒置残差结构模块<sup>[18-20]</sup>，模块的激活函数前两层使用ReLU6，最后一层使用线性激活函数Linear，这样可减少特征损失。MobileNet V2的网络层次不深，核心由17个倒置残差模块组成<sup>[15]</sup>，其模型的网络层次参数配置如表2所示。

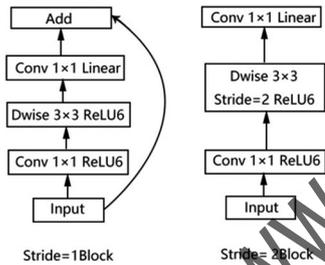


图5 倒置残差模块

Fig.5 Inverted residual block

表2 MobileNet V2网络结构参数配置

Tab.2 Parameter configuration of MobileNet V2

输入图像大小	操作	扩展系数	重复次数	输出通道数	步长
$224^2 \times 3$	卷积	—	32	1	2
$112^2 \times 32$	倒置残差卷积	1	16	1	1
$112^2 \times 16$	倒置残差卷积	6	24	2	2
$56^2 \times 24$	倒置残差卷积	6	32	3	2
$28^2 \times 32$	倒置残差卷积	6	64	4	2
$14^2 \times 64$	倒置残差卷积	6	96	3	1
$14^2 \times 96$	倒置残差卷积	6	160	3	2
$7^2 \times 160$	倒置残差卷积	6	320	1	1
$7^2 \times 320$	卷积	—	1,280	1	1
$7^2 \times 1280$	平均池化	—	—	1	—
$1 \times 1 \times 1280$	卷积	—	$k$	—	—

注：“—”表示该处值不存在，“ $k$ ”为任意正整数

### 3.3 改进的MobileNet V2卷积神经网络

MobileNet V2使用的激活函数全部是ReLU(Rectified

Linear Unit)函数，其定义为： $F(x)=\max(0,x)$ 。它是常用的卷积神经网络激活函数，在 $x>0$ 区域上，不会出现梯度饱和与梯度消失，计算复杂度低，不需要进行指数运算，只要一个阈值就可以得到激活值。但在 $x<0$ 时，梯度为0，这个神经元及之后的神经元梯度永远为0，不再对任何数据有所响应，导致相应参数永远不会被更新，即神经元坏死。Leaky ReLU函数在ReLU函数的基础上，当 $x<0$ 时引入一个非常小的 $\gamma$ 值作为梯度，能够避免神经元坏死，同时补充了梯度。倒置残差卷积中使用跨越连接，增加了模型的复杂程度，训练模型时更容易发生过拟合，因此对跨越连接添加Dropout层，随机丢弃部分输入数据，降低过拟合的可能性，同时增大模型泛化性和鲁棒性。

## 4 训练模型(Training model)

### 4.1 实验环境搭建

实验的操作系统为Windows 10，第三方包管理与环境管理使用Anaconda，深度学习框架采用TensorFlow Gpu 1.8版本，编程语言为Python 3.6版本，脚本编辑器为Spyder，显卡为(GPU)NVIDIA GeForce GTX 1050，显存大小为4 GB，中央处理器(CPU)为Intel Pentium G4600 3.6 GHz，固态硬盘大小为8 GB。

### 4.2 数据增强

采用动态的数据增强方法，在每轮训练开始前，重新随机选择训练集中的叶片图像进行左右翻转或上下翻转，以及在一定范围内随机设置图片亮度、对比度和饱和度。

### 4.3 训练MobileNet V2模型

使用训练集的样本图像训练MobileNet V2模型，模型训练只使用显卡运算，不使用中央处理器，每一轮参与训练的样本总数为1,831。每轮训练采用分批的方法，受到显卡性能的限制，一个批次训练所使用的样本数(Batch Size)不宜太大，设为4，因此每轮的迭代次数(Epoch)等于457。初始学习率(Learning Rate)设为0.01，采用动态学习率的策略，神经网络的容量(Capacity)设为100。图6表示准确率(Accuracy)和损失值(Loss)随轮次的变化，可见随着轮次数值接近40次时，训练集损失值收敛于0，准确率趋于1，表明模型训练成功。

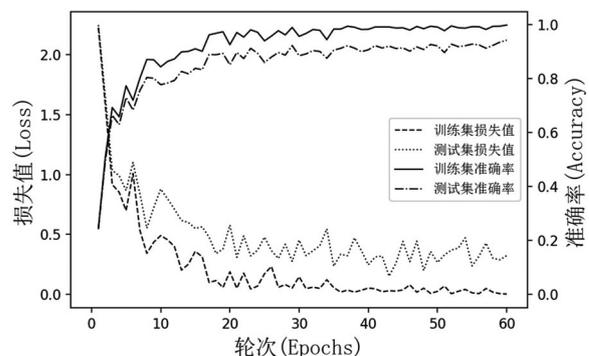


图6 准确率和损失值随轮次的变化

Fig.6 Accuracy and loss value change with epochs

### 5 评估模型预测效果(Evaluate the prediction effect of the model)

#### 5.1 预测结果描述

使用测试集样本图像测试已经完成训练的MobileNet V2深度学习模型，测试集所有叶片图像均未用于训练模型，总计446张，其中各类植物叶片图像具体数量如表1所示。模型预测结果的描述方法使用混淆矩阵，图7表示具体的预测结果，横轴为预测值，纵轴为真实值，对角线上的数字表示某一样本被正确预测的个数，非对角线上的数字表示某一样本被错误预测为其他类别的个数，对角线上方块的颜色越深，表示该类别叶片图像的预测效果越好。

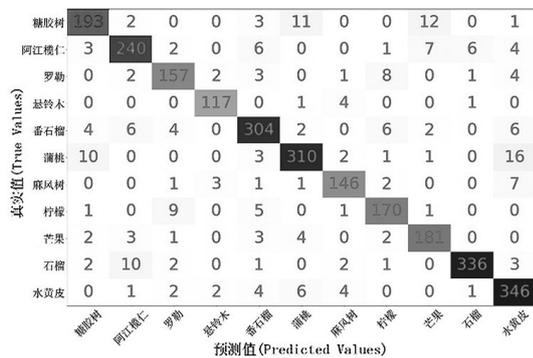


图7 预测结果  
Fig.7 Prediction results

#### 5.2 预测结果评价

经过模型预测，正确预测的样本数为358张，总体准确率为91.41%，总体准确率定义的是预测正确的个数与所有样本的个数百分比。模型预测效果的评价指标使用精确率和召回率<sup>[21-23]</sup>，精确率定义的是预测正确的个数与预测为正例的个数百分比，召回率定义的是预测正确的个数与真实为正例的个数百分比，表3表示预测结果中各个类别的精确率和召回率。精确率最高的前5个类别依次是悬铃木、水黄皮、石榴、芒果、番石榴，分别是95.12%、94.54%、94.12%、92.35%、91.02%；召回率最高的前5个类别依次是石榴、悬铃木、蒲桃、番石榴、麻风树，分别是97.39%、94.35%、92.54%、91.29%、91.25%。

表3 预测结果的精确率和召回率

Tab.3 Accuracy and recall rate of prediction results

测试结果	糖胶树	阿江榄仁	罗勒	悬铃木	番石榴	蒲桃	麻风树	柠檬	芒果	石榴	水黄皮
精确率/%	86.94	89.22	88.20	95.12	91.02	90.38	90.68	90.91	92.35	94.12	94.54
召回率/%	89.77	90.91	88.20	94.35	91.29	92.54	91.25	89.01	88.73	97.39	89.41

### 6 结论(Conclusion)

MobileNet V2卷积神经网络是针对移动端和嵌入式设备而提出的轻量级模型，具有体积小、参数量少、计算量低等优点，适用于手机等各类移动设备。本文基于MobileNet V2

卷积神经网络设计出多类别的植物叶片图像识别方法，利用公共数据集建立包含11类植物叶片图像的数据集。模型训练开始时，首先根据硬件情况对初始学习率、轮次、步长等参数进行合理设置，创新使用动态数据增强方法，极大扩增了样本训练量，并在接下来的测试集预测中，总体准确率高达91.41%，取得较好的识别效果，表明使用该算法对植物叶片进行分类是实际可行的，同时也为实现叶片图像识别的移动端提供了理论技术和实用方法。

### 参考文献(References)

- [1] 郭志芳.植物在维护生态平衡中的作用[J].广东蚕业,2020,54(05):18-19.
- [2] 雷星梅.如何使植物保护新技术在生态农业中发挥作用[J].农家参谋,2020(14):92.
- [3] 王升,蒋待泉,康传志,等.药用植物次生代谢在中药材生态种植中的作用及利用[J].中国中药杂志,2020,45(09):2002-2008.
- [4] 邓艳芳.植物生长物质对农产品生产的作用和影响[J].农业开发与装备,2017(10):62-63.
- [5] 韩斌,曾松伟.基于多特征融合和卷积神经网络的植物叶片识别[J].计算机科学,2021,48(S1):113-117.
- [6] 王维.外来入侵植物叶片图像识别与分类方法研究[D].沈阳:沈阳大学,2021.
- [7] 边缘,孔小莹,张莉,等.基于卷积神经网络的植物叶片树种识别研究与实现[J].智能计算机与应用,2020,10(10):23-26.
- [8] 朱良宽,晏铭,黄建平.一种新型卷积神经网络植物叶片识别方法[J].东北林业大学学报,2020,48(04):50-53.
- [9] 张露.基于深度学习的植物叶片图像识别方法研究[D].北京:北京林业大学,2019.
- [10] 盖荣丽,蔡建荣,王诗宇,等.卷积神经网络在图像识别中的应用研究综述[J].小型微型计算机系统,2021,42(09):1980-1984.
- [11] 孙颖异,李健,时天,等.基于改进的AlexNet卷积神经网络的植物叶片识别[J].种子,2020,39(02):77-81.
- [12] 洪奇峰,施伟斌,吴迪,等.深度卷积神经网络模型发展综述[J].软件导刊,2020,19(04):84-88.
- [13] 章琳,袁非牛,张文睿,等.全卷积神经网络研究综述[J].计算机工程与应用,2020,56(01):25-37.
- [14] 林景栋,吴欣怡,柴毅,等.卷积神经网络结构优化综述[J].自动化学报,2020,46(01):24-37.
- [15] ILESANMI A E, ILESANMI T O. Methods for image denoising using convolutional neural network: A review[J]. Complex & Intelligent Systems, 2021, 7(5):2179-2198.