

## 基于 Enhanced VGG16 的油茶品种分类

孟志超<sup>1</sup>, 贺磊盈<sup>1,2</sup>, 杜小强<sup>1,2\*</sup>, 张国凤<sup>1,2</sup>, 姚小华<sup>3</sup>, 吴顺凯<sup>1</sup>, 郭豪鉴<sup>1</sup>

(1. 浙江理工大学机械与自动控制学院, 杭州 310018; 2. 浙江省种植装备技术重点实验室, 杭州 310018; 3. 中国林业科学研究院亚热带林业研究所, 杭州 311400)

**摘要:** 随着油茶产业不断壮大, 市场上也出现了油茶幼苗品系混乱、以假乱真、以次充好的现象, 因此急需开发一种专门的分类识别算法实现不同油茶品种的准确识别。农业领域常用 VGG、ResNet 网络模型进行分类工作, 但存在权重空间过大和准确率不高等问题。该研究对 VGG16 网络模型进行层间删减以及结构调整, 提出了 Enhanced VGG16 网络模型, 在油茶叶数据集上完成模型训练与测试, 并与现有经典卷积神经网络 (AlexNet、VGG16、Resnet50、InceptionV3、Xception) 进行对比。结果表明, Enhanced VGG16 网络模型的训练集准确率和测试集准确率分别为 98.98% 和 98.44%, 权重空间为 90.6 MB。与原始 VGG16 模型相比, 训练集准确率和测试集准确率分别提高 3.08 和 2.05 个百分点, 权重空间下降 165.4 MB, 模型性能显著提升。Enhanced VGG16 网络模型与经典卷积神经网络相对比, 模型综合性能更优。该研究为通过油茶叶进行品种分类识别提供了依据, 同时可为其他农作物品种识别提供参考。

**关键词:** 深度学习; 油茶叶; 分类; Enhanced VGG16; hard-Swish; ReLU6

doi: 10.11975/j.issn.1002-6819.2022.10.021

中图分类号: S223.71

文献标志码: A

文章编号: 1002-6819(2022)-10-0176-06

孟志超, 贺磊盈, 杜小强, 等. 基于 Enhanced VGG16 的油茶品种分类[J]. 农业工程学报, 2022, 38(10): 176-181.

doi: 10.11975/j.issn.1002-6819.2022.10.021 <http://www.tcsae.org>

Meng Zhichao, He Leiying, Du Xiaoqiang, et al. Classification of *Camellia oleifera* based on Enhanced VGG16 network[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2022, 38(10): 176-181. (in Chinese with English abstract) doi: 10.11975/j.issn.1002-6819.2022.10.021 <http://www.tcsae.org>

## 0 引言

油茶是中国南方特有的木本油料树种, 是世界四大木本油料之一<sup>[1-2]</sup>, 种植面积和产量位于中国木本油料之首。油茶品种依靠肉眼进行判断, 费时费力, 错误率高, 难以防止以次充好的事件发生, 急需开发一种专门的分类识别算法实现不同油茶品种的准确识别, 以便于农户放心购买高产油茶及油茶苗的推广和售卖, 这对促进油茶产业可持续发展具有重要意义。油茶的叶子不像花和果子一样具有季节性, 它的性状更容易被用来区分油茶品种。Liu 等<sup>[3]</sup>利用叶子切片对山茶属植物进行了研究。Pi 等<sup>[4]</sup>通过叶片形态解剖特征对山茶科植物进行分类。

油茶类内变异性低, 在品种级别对叶子分类仍然是一项具有挑战的任务<sup>[5]</sup>。目前依靠叶子实现农作物病虫害分类较为常见<sup>[6-10]</sup>, 针对品种级别的分类研究较少。Lu 等<sup>[11]</sup>运用聚类、支持向量机等方法依靠叶形态和脉络特征实现山茶的分类。陈舒迪等<sup>[12]</sup>使用形状、颜色和纹理

特征对金线莲叶片图像进行提取和融合, 运用 LightGBM 进行分类。但传统机器学习 (KNN<sup>[13]</sup>、朴素贝叶斯<sup>[14]</sup>、决策树<sup>[15]</sup>、SVM<sup>[16]</sup>、随机森林<sup>[17]</sup>等) 进行物体分类在特征提取过程中需要人工提取特征, 主观性较强, 表达能力有限, 模型鲁棒性差。近年来深度学习<sup>[18]</sup>在目标分类和检测的方面展现出了巨大优势, 被广泛应用在农业分类领域<sup>[19-23]</sup>。Krishnamoorthy 等<sup>[24]</sup>基于迁移学习技术运用 InceptionResNetV2 对水稻图像中的病虫害进行分类, 获得了 95.67% 的准确率。Nandhini 等<sup>[25]</sup>利用 CNN 方法对不同类型的叶病进行分类获得了整体 99% 的成功率。

以上研究的分类目标均具有较为明显的特征差异, 而油茶叶片的特征差异不明显, 通过油茶叶进行品种级分类难度大, 非专业人员难以区分。本文用扫描仪获取纹理清晰、平整的油茶叶数据集图像, 考虑到嵌入式设备和移动端中储存资源有限, 选择模型深度较浅的 VGG16 进行层间删减以及结构调整, 在保证准确率的同时降低模型权重空间。在油茶叶数据集上完成模型训练与测试, 并与现有经典卷积神经网络 (AlexNet<sup>[26]</sup>、VGG16<sup>[27]</sup>、Resnet50<sup>[28]</sup>、InceptionV3<sup>[29]</sup>、Xception<sup>[30]</sup>) 进行对比, 以期为油茶叶进行品种分类识别提供依据, 并为其其他农作物品种识别提供参考。

## 1 材料与方法

### 1.1 油茶叶图像数据集构建

本文所用“长林”3号、4号、40号和53号油茶叶均来自浙江省金华市东方红林场国家油茶良种基地。由

收稿日期: 2021-09-16 修订日期: 2022-05-11

基金项目: 国家自然科学基金项目 (31971798); 国家重点研发计划课题油茶生态经济型品种筛选及配套栽培技术 (2019YFD1001602); 浙江省“领雁”研发攻关计划项目 (2022C02057)

作者简介: 孟志超, 博士生, 研究方向为智能农业装备。

Email: 942002162@qq.com

\*通信作者: 杜小强, 博士, 教授, 博士生导师, 研究方向为智能农业装备。

Email: xqiangdu@zstu.edu.cn

中国农业工程学会会员: 杜小强 (E043500018M)

于油茶叶具有叶面卷曲和尖端自然下垂等特点，自然摆放图像采集时会受姿态的影响，因此本文采用数字扫描仪（Epson Perfection V30）获取压平后的油茶叶图像。采集后的图像如图 1 所示，图中油茶叶脉络纹理更清晰、平整，有利于卷积神经网络模型的训练。获取 4 个品种油茶叶图像各 1 800 张，以 3:2 的比例划分训练集和测试集。训练集 4 320 张、测试集 2 880 张。

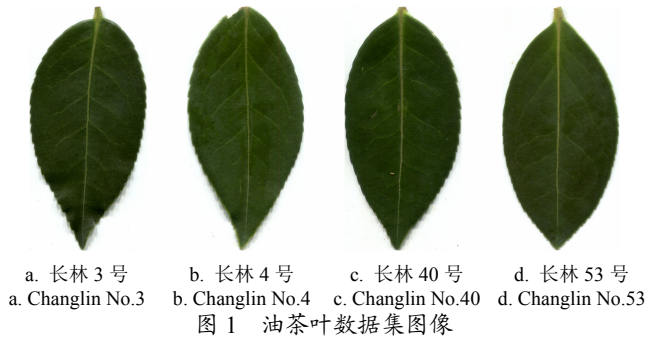


Fig.1 Examples of *Camellia oleifera* leaf image data set

## 1.2 数据增强

深度学习具有强表达能力，容易牺牲对未来数据的解释能力，导致模型过拟合。为了进一步增加数据的多样性，进而提高模型性能，防止过拟合现象发生，需要对训练集进行数据增强。本文运用 Keras 中的 Image Data Generator 进行数据增强，对图像进行亮度调整、20° 内的随机旋转、水平和垂直方向的镜像等数据增强方式，并将所有图像调整为 512 像素×512 像素×3 通道。

## 1.3 VGG16 网络模型改进策略

### 1.3.1 VGG16 网络模型

VGG16 网络模型由牛津大学视觉几何组（Visual Geometry Group）提出，是经典的卷积神经网络之一，它由 5 组卷积层（convolution layer）、3 个全连接层（full connected layer）和一个 Softmax 输出层构成，每组卷积层之间通过最大池化层（maxpooling layer）分开，所有隐层的激活单元都采用 ReLU 非线性激活函数。VGG16 网络使用 3×3 小卷积核代替大卷积核，不仅减少了网络的训练参数，而且增加了网络中的非线性单元，使得网络学习能力增强。VGG16 前两个全连接层神经元数量为 4 096 个，产生大量网络参数，因此对全连接层神经

个数进行调整，有利于防止模型过拟合和减小模型权重空间。

### 1.3.2 ReLU6、Swish 和 hard-Swish 激活函数

激活函数引入的非线性特征对神经网络的训练起着至关重要的作用。在 VGG16 网络模型中，运用 ReLU 作为激活函数，其数学定义为

$$f(x)=\max(0,x) \quad (1)$$

图像数据经过卷积计算后，会产生大量的负值，ReLU 激活函数将其变为 0 后，会丢失大量的特征信息。

ReLU6 是在 ReLU 激活函数的基础上将大于 6 的数据部分变为 6，为了在移动端 float16 低精度的时候，也能有很好的数值分辨率，如果对 ReLU 的输出值不加限制，那么输出范围就是 0 到正无穷，而低精度的 float16 无法精确描述其数值，带来精度损失。其数学定义为

$$f(x)=\min(\max(0,x),6) \quad (2)$$

2017 年和 2019 年 Google 相继提出了 Swish 与 hard-Swish 激活函数，其数学定义为

$$\text{Swish} = x \cdot \text{Sigmoid}(\beta x) \quad (3)$$

$$\text{hard-Swish}[x] = x \cdot \frac{\text{ReLU6}(x+3)}{6} \quad (4)$$

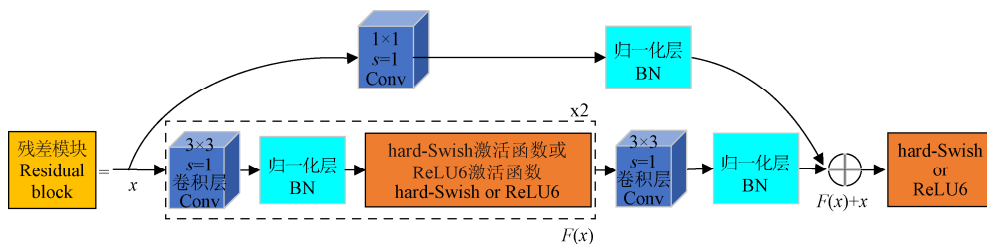
Swish 具有无上界有下界、平滑、非单调的特性，在模型效果上优于 ReLU6 激活函数，但计算更为复杂。hard-Swish 函数形状与 Swish 函数类似，能避免神经元死亡现象，不但连续可微，而且计算量比 Swish 函数小。因此本文利用 ReLU6 和 hard-Swish 激活函数代替 ReLU 激活函数，提高模型准确率。

### 1.3.3 残差模块

本文通过在输入通道和输出通道间运用 1×1 卷积层进行升维增加跳跃连接，并在 3×3 卷积层后加入归一化层和 hard-Swish 激活函数或 ReLU6 激活函数构建残差模块，如图 2 所示。在一定程度上避免了卷积层在进行信息传递时的特征损失，可以在输入特征的基础上学习到新的特征，防止网络过深导致的梯度消失<sup>[31]</sup>。

### 1.3.4 Dropout 与 L2 正则化优化

Dropout<sup>[32]</sup>是 Srivastava 等在 2012 年提出的，通过在前向传导的过程中，随机选择部分神经元，使其激活值按照特定的概率暂停工作，从而增强模型的泛化能力，防止过拟合现象的发生。



注：s 代表步长，1×1 与 3×3 代表卷积核大小，Conv 代表卷积层，BN 代表归一化层，F(x)代表残差函数，x 代表输入。

Note: s represents stride, 1×1 and 3×3 represent size of convolution kernel, Conv represents convolution layer, BN represents Batch Normalization, F(x) represents residual function, x represents input.

图 2 残差模块

Fig.2 Residual block

L2 正则化是另一项有效防止过拟合的手段，其形式

是在原先损失函数的基础上多加一项正则项，加上 L2 正

则项的损失函数  $J$  可以表示为

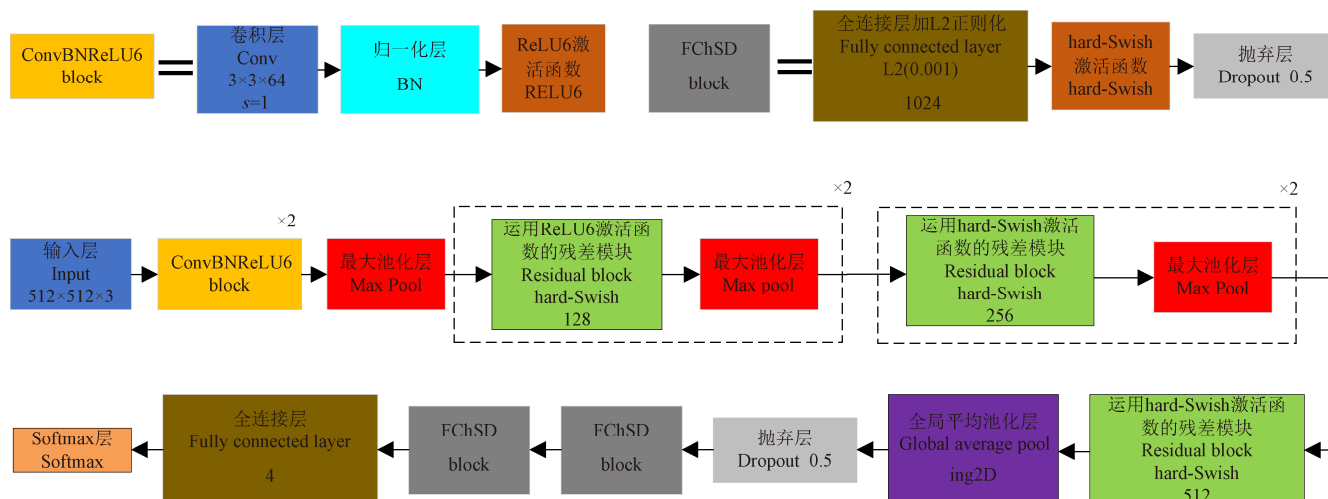
$$J = J_0 + \lambda \sum_i^n \theta_i^2 \quad (5)$$

式中  $J_0$  是原始的损失函数,  $\lambda \sum_i^n \theta_i^2$  是 L2 正则项,  $\theta$  是待学习参数,  $\lambda$  控制正则项的大小, 对稀疏的有尖峰的权重向量施加大的惩罚, 偏好于均匀的参数。

本文通过在全连接层中运用 Dropout 与 L2 正则化, 防止模型过拟合。

### 1.3.5 Enhanced VGG16 网络模型

本文运用 hard-Swish 和 ReLU6 激活函数、残差模块、Dropout、L2 正则化等技术, 对 VGG16 网络模型进行层间删减以及结构调整。Enhanced VGG16 网络模型架构如图 3 所示, ConvBNReLU6 模块由卷积层、归一化层和 ReLU6 激活函数组成, FChSD 模块由采用 L2 正则化的全连接层、hard-Swish 激活函数和 Dropout 层组成。整体架构由 2 个 ConvBNReLU6 模块、5 个残差模块和 2 个 FChSD 模块组成, 通过 Max pool 层减小图像尺寸, 全连接层输出维度设为 4, 利用 Softmax 进行分类。



注: ConvBNReLU6 block 代表由卷积层、归一化层和 ReLU6 激活函数组成的模块, FChSD block 代表由全连接层、hard-Swish 激活函数和抛弃层组成的模块,  $512 \times 512 \times 3$  代表输入图像大小,  $3 \times 3 \times 64$  代表卷积核大小与数量,  $\times 2$  代表模块重复次数, 128、256 和 512 代表卷积核个数, 1024 和 4 代表神经元个数, 0.001 代表正则化系数, 0.5 代表丢弃概率。

Note: ConvBNReLU6 block represents a block consisting of a convolutional layer, a batch normalization layer and a ReLU6 activation function, FChSD block represents a block consisting of a fully connected layer, a hard-Swish activation function, and a dropout layer,  $512 \times 512 \times 3$  represents input image size,  $3 \times 3 \times 64$  represents size and number of convolution kernels,  $\times 2$  represents block repetition times, 128, 256 and 512 represent the number of convolution kernels, 1024 and 4 represent the number of neurons, 0.001 represents regularization coefficient, 0.5 represents the dropout probability.

图 3 Enhanced VGG16 网络模型  
Fig.3 Enhanced VGG16 network model

## 2 结果与分析

### 2.1 激活函数对模型性能的影响

为了证明在 Enhanced VGG16 网络中运用 hard-Swish 和 ReLU6 激活函数代替 ReLU 激活函数对改进模型性能提升具有实质性贡献, 对 4 种不同的激活函数实施策略进行试验。试验结果如表 1 所示, 在相同条件下, 运用 ReLU6 激活函数替换 ReLU 激活函数, 测试集准确率提高 2.81 个百分点; 运用 hard-Swish 激活函数代替 ReLU 激活函数, 测试集准确率提高 3.23 个百分点; 在最后一个残差模块和全连接层运用 hard-Swish 激活函数, 其余均采用 ReLU6 激活函数, 测试集准确率提高 3.65 个百分点。通过本试验的结果来看, 联合采用 hard-Swish 和

该模型卷积层通过连续堆叠 3 个  $3 \times 3$  小卷积核代替 1 个  $7 \times 7$  大卷积核, 在感受野相同前提下, 堆叠小卷积核不仅可以减小网络参数量, 且增加了网络中的非线性单元, 并吸收 ResNet 网络残差学习思想, 利用跳跃连接, 避免了卷积层在进行信息传递时的特征损失。

根据 Howard 在 MobileNet V3 网络中提出的 hard-Swish 激活函数在层数较深时才能发挥出自身优势<sup>[33]</sup>, 因此 Enhanced VGG16 在最后一个残差模块和全连接层运用 hard-Swish 代替 ReLU 激活函数, 其余均用 ReLU6 代替 ReLU 激活函数, 提高模型准确率。

原始 VGG16 网络的参数绝大部分集中于全连接网络之中, 将前两个全连接层神经元个数从 4 096 改为 1 024, 有利于减小模型的网络参数。在全连接层加入 L2 正则化和 Dropout 层, 有利于提高分类精度, 加速网络收敛, 防止模型过拟合。该模型比 AlexNet、VGG16、Resnet50、InceptionV3、Xception 网络模型训练参数更少, 比 Resnet50、InceptionV3、Xception 模型深度更浅, 更适宜部署在移动端和嵌入式设备进行油茶茶叶分类识别。

ReLU6 激活函数可以显著提高模型准确率, 增强了模型的可靠性, 且不会引入多余可训练参数。

### 2.2 残差、Dropout 和 L2 正则化对模型性能影响

为了证明增加残差、Dropout 和 L2 正则化对提高模型准确率具有重要作用, 将 Enhanced VGG16 网络与 Enhanced VGG16 中不使用残差结构、Dropout 层或 L2 正则化的网络进行对比。试验结果如表 2 所示, 残差结构使训练集准确率和测试集准确率分别提高 2.38 和 1.7 个百分点, 使权重空间上升仅 2 MB。使用 Dropout 层和 L2 正则化, 训练集准确率和测试集准确率分别提高 0.86、3.33 个百分点和 2.26、3.41 个百分点。残差结构虽然使权重空间上升了 2 MB, 但是使模型准确率大幅度提高。Dropout 层和 L2 正则化的加入, 加强了模型的泛化能力,



防止模型过拟合，使模型可靠性增强。

表 1 4 种不同激活函数策略的模型性能对比

Table 1 Performance comparison of four different activation function strategies %

激活函数策略 Activation function strategies	训练集准确率 Training set accuracy	测试集准确率 Test set accuracy
全部采用 ReLU 激活函数 All by ReLU activation function	97.34	94.79
全部采用 ReLU6 激活函数 All by ReLU6 activation function	96.37	97.60
全部采用 hard-Swish 激活函数 All by hard-Swish activation function	98.12	98.02
联合采用 ReLU6 和 hard-Swish 激活函数 Combination ReLU6 and hard-Swish activation functions	98.98	98.44

表 2 去除残差、Dropout 和 L2 正则化方法对模型性能影响

Table 2 Effects of removing residual, dropout and L2 regularization on model performance

方法 Methods	权重空间 Space of weight /MB	可训练参数数量 Number of trainable parameters/ $10^6$	训练集准确率 Training set accuracy /%	测试集准确率 Test set accuracy /%
无残差 No residual	88.6	11.6	96.60	96.74
无 Dropout 层 No Dropout	90.6	11.8	98.12	96.18
无 L2 正则化 No L2 Regularization	90.6	11.8	95.65	95.03
增强的 VGG16 Enhanced VGG16	90.6	11.8	98.98	98.44

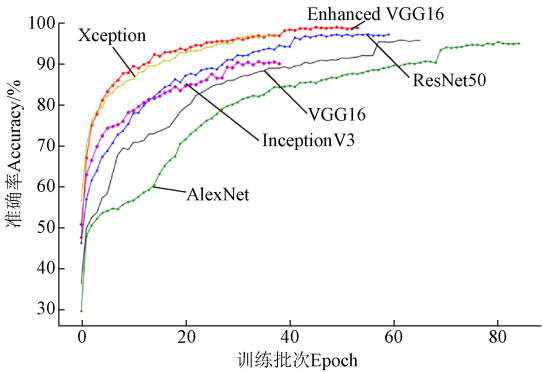
2.3 模型性能对比试验

为进一步验证所改进模型分类油茶叶的效果，在相同试验条件下与其他 5 种经典卷积神经网络进行对比。根据图 4 所示，在模型准确率收敛方面，Enhanced VGG16 网络模型大约在 40 个 epoch 后趋于收敛，在所有模型中收敛速度处于前列，能在较短的时间训练出更加优秀的模型。因为 Enhanced VGG16 网络模型运用了 Dropout、L2 正则化技术，训练过程中测试集准确率波动最小。

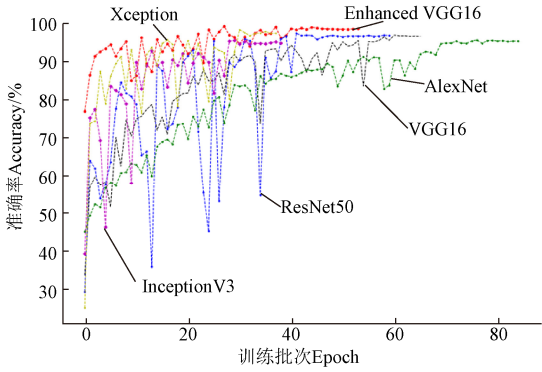
根据表 3 可知，因为 Enhanced VGG16 网络模型运用了残差结构和非线性更强的，性能更好的 hard-Swish 和 ReLU6 激活函数，所以在训练集和测试集上准确率分别达到 98.98%和 98.44%，训练集准确率与 AlexNet、VGG16、ResNet50、InceptionV3、Xception 相对比分别提高 3.77、3.08、1.57、8.79 和 2.13 个百分点，测试集准确率与 AlexNet、VGG16、ResNet50、InceptionV3、Xception 相对比分别提高 3.27、2.05、1.77、3.75 和 1.22 个百分点，取得了最高的训练集和测试集准确率，分类效果优于其他模型。在权重空间与可训练参数方面，Enhanced VGG16 网络模型没有采用通过堆叠网络层数增强模型效果的方法，因此权重空间为 90.6 MB，可训练参数为  $11.8 \times 10^6$ ，均小于其他对比模型，易于部署在移动端和嵌入式设备中。Enhanced VGG16 网络模型平均测试耗时 55.32 ms。选择测试优化器为 RMSprop，损失函数为 categorical\_crossentropy，Batch Size 和学习率分别为 4 和 0.000 01，随着迭代次数的增加，学习率降低到原来的 10%。与 Xception 模型相比，权重空间下降 68.4 MB；与原始

VGG16 模型相比，训练集准确率和测试集准确率分别提高 3.08 和 2.05 个百分点，权重空间下降 165.4 MB，平均测试损耗下降 2.18 ms。

综上所述，本文提出的 Enhanced VGG16 网络模型综合性能最优，相对于其他经典的卷积神经网络，最适合用于油茶叶品种识别任务，不仅分类准确率最高，且权重空间最小，能部署在更多设备中，适用范围更广，模型可靠性和适应性更佳。虽然速度稍慢于 ResNet50 和 Xception，但满足实时性要求，而且油茶分类识别任务对时间要求不高。



a. 训练集准确率变化曲线  
a. Accuracy curve of the training set



b. 测试集准确率变化曲线  
b. Accuracy curve of the test set  
图 4 模型准确率对比

Fig.4 Model accuracy comparison

表 3 模型性能对比表

Table 3 Performance comparison of different models

模型 Models	权重空间 Space of weight /MB	可训练参数数量 Number of trainable parameters/ $10^6$	训练集准确率 Training set accuracy /%	测试集准确率 Test set accuracy /%	平均测试耗时 Average detection time/ms
AlexNet	164	21.6	95.21	95.17	62.15
VGG16	256	33.6	95.90	96.39	57.5
ResNet50	180	23.5	97.41	96.67	52.81
InceptionV3	167	21.8	90.19	94.69	57.25
Xception	159	20.8	96.85	97.22	53.86
Enhanced VGG16	90.6	11.8	98.98	98.44	55.32

3 结 论

1) 本文运用 hard-Swish 和 ReLU6 激活函数、残差模块、Dropout、L2 正则化等技术，对 VGG16 网络模型

进行层间删减以及结构调整,提出了 Enhanced VGG16 网络模型,用于油茶叶品种识别,从而避免传统方法依赖于耗时费力的人工特征设计,提高模型的检测分类能力,降低了网络模型权重空间,并在油茶叶数据集进行验证。结果表明,Enhanced VGG16 网络模型在训练集和测试集准确率上比 VGG16 分别提高了 3.08 和 2.05 个百分点,权重空间下降 165.4 MB,平均测试损耗下降 2.18 ms,模型性能显著提升。

2) 在相同测试条件下,与 AlexNet、ResNet50、InceptionV3、Xception 进行对比试验,本文提出的 Enhanced VGG16 网络模型测试集准确率相比于效果最好的 Xception 提高 1.22 个百分点,权重空间下降 68.4 MB,训练过程中准确率波动较小。因此其综合性能最优。

本文基于 VGG16 网络模型,充分吸收现有算法优点,获得了准确率高、权重空间小,易于部署在移动端和嵌入式设备中的油茶叶品种分类模型。该研究方法为油茶叶进行品种分类识别提供了依据,同时可为其他农作物品种识别提供参考,但由于油茶的多样性和复杂性,还存在一定的改进之处。一是油茶品种多样性,本研究仅对长林系 4 个品种进行分类研究,为加强模型的适用性,后期将增加训练油茶叶品种,完成国家审定的 10 个长林系品种的分类识别任务,为更大范围内的油茶叶品种分类提供支持。二是油茶叶纹理和颜色随着生育时期的变化有一定量的改变,为加强模型的可靠性,增加不同生育时期的油茶叶以扩大训练样本量,增强模型鲁棒性。三是非结构环境下光照等因素影响,田间环境的油茶叶分类仍是一个具有挑战的任务。

#### 参考文献

- [1] 谢凤,钟海雁.油茶多酚的提取及功能研究进展[J].经济林研究,2015,33(2):158-162.  
Xie Feng, Zhong Haiyan. Advance in research on extraction and function of polyphenols from *camellia oleifera*[J]. Non-wood Forest Research, 2015, 33(2): 158-162. (in Chinese with English abstract)
- [2] 谢世平,江贤华,黄秋兰.2014—2018 年将乐县气象因子对油茶产量的影响[J].粮食科技与经济,2019,44(12):152-153.
- [3] Lin X, Peng Q, Lu H, et al. Leaf anatomy of *camellia sect. Oleifera* and sect. *Paracamellia* (Theaceae) with reference to their taxonomic significance[J]. Journal of Systematics & Evolution, 2008, 46(2): 183-193.
- [4] Pi E, Peng Q, Lu H, et al. Leaf morphology and anatomy of *Camellia* section *camellia* (Theaceae)[J]. Botanical Journal of the Linnean Society, 2009, 159(3): 456-476.
- [5] 湖南省油茶品种研究协作组.湖南油茶品种分类及优良类型的研究[J].经济林研究,1985,3(1):1-26.  
The Cooperative Group for the Study of the *Camellia Oleifera* Types in Hunan. The classification of camellia species in huinan and the study of its classic types[J]. Non-wood Forest Research, 1985, 3(1): 1-26. (in Chinese with English abstract)
- [6] 刘志勇,张丽秀,钟婷婷,等.基于改进 LeNet-5 的番茄病虫害识别的研究[J].赣南师范大学学报,2020,41(6):76-80.  
Liu Zhiyong, Zhang Lixiu, Zhong Tingting, et al. Study on tomato pest identification based on improved LeNet-5[J]. Journal of Gannan Normal University, 2020, 41(6): 76-80. (in Chinese with English abstract)
- [7] 王东方,汪军.基于迁移学习和残差网络的农作物病害分类[J].农业工程学报,2021,37(4):199-207.  
Wang Dongfang, Wang Jun. Crop disease classification with transfer learning and residual networks[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2021, 37(4): 199-207. (in Chinese with English abstract)
- [8] Selvam L, Kavitha P. Classification of ladies finger plant leaf using deep learning[J]. Journal of Ambient Intelligence and Humanized Computing, 2020(5): 1-9.
- [9] 樊湘鹏,许燕,周建平,等.基于迁移学习和改进 CNN 的葡萄叶部病害检测系统[J].农业工程学报,2021,37(6):151-159.  
Fan Xiangpeng, Xu Yan, Zhou Jianping, et al. Detection system for grape leaf diseases based on transfer learning and updated CNN[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2021, 37(6): 151-159. (in Chinese with English abstract)
- [10] Waheed A, Goyal M, Gupta D, et al. An optimized dense convolutional neural network model for disease recognition and classification in corn leaf[J]. Computers and Electronics in Agriculture, 2020, 175: 105456.
- [11] Lu H, Jiang W, Ghiassi M, et al. Classification of camellia (Theaceae) species using leaf architecture variations and pattern recognition techniques[J]. Plos One, 2012, 7(1): e29704.
- [12] 陈舒迪,柴琴琴,张勋,等.基于多特征融合和 LightGBM 的金线莲品系识别[J].江苏农业学报,2021,37(1):155-162.  
Chen Shudi, Cai Qinqin, Zhang Xun, et al. Identification of *anoectochilus roxburghii* strains based on multi feature fusion and LightGBM[J]. Jiangsu Journal of Agricultural Sciences, 2021, 37(1): 155-162. (in Chinese with English abstract)
- [13] Cover T. Estimation by the nearest neighbor rule[J]. IEEE Transactions on Information Theory, 1966, 14(1): 50-55.
- [14] Rish I. An empirical study of the naive bayes classifier[J]. Journal of Universal Computer Science, 2001, 3(22): 41-46.
- [15] Shlien S. Multiple binary decision tree classifiers[J]. Pattern Recognition, 1990, 23(7): 757-763.
- [16] Platt J. Sequential minimal optimization: A fast algorithm for training support vector machines[J]. Advances in Kernel Methods-Support Vector Learning, 1999, 208(1): 212-223.
- [17] Breiman L. Random forests[J]. Machine learning, 2001, 45(1): 5-32.
- [18] Arel I, Rose D, Karnowski T. Deep machine learning-A new frontier in artificial intelligence research[J]. IEEE Computational Intelligence Magazine, 2010, 5(4): 13-18.
- [19] Tavakoli H, Alirezazadeh P, Hedayatipour A, et al. Leaf image-based classification of some common bean cultivars using discriminative convolutional neural networks[J]. Computers and Electronics in Agriculture, 2021, 181: 105935.
- [20] Zhou H, Zhuang Z, Liu Y, et al. Defect classification of green plums based on deep learning[J]. Sensors, 2020, 20(23): 6993.
- [21] Majeed Y, Zhang J, Zhang X, et al. Deep learning based segmentation for automated training of apple trees on trellis wires[J]. Computers and Electronics in Agriculture, 2020, 170: 105277.
- [22] Jia W, Tian Y, Luo R, et al. Detection and segmentation of overlapped fruits based on optimized mask R-CNN application in apple harvesting robot[J]. Computers and Electronics in Agriculture, 2020, 172: 105380.
- [23] Liang C, Xiong J, Zheng Z, et al. A visual detection method for nighttime litchi fruits and fruiting stems[J]. Computers and Electronics in Agriculture, 2020, 169: 105192.
- [24] Krishnamoorthy N, Prasad L, Kumar C, et al. Rice leaf diseases prediction using deep neural networks with transfer

- learning[J]. Environmental Research, 2021, 198(11): 111275.
- [25] Nandhini S, Ashokkumar K. Improved crossover based monarch butterfly optimization for tomato leaf disease classification using convolutional neural network[J]. Multimedia Tools and Applications, 2021, 80(12): 18583-18610.
- [26] Krizhevsky A, Sutskever I, Hinton G. Imagenet classification with deep convolutional neural networks[J]. Advances in Neural Information Processing Systems, 2012, 25(2): 1097-1105.
- [27] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition[J]. Computer Science, 2014, 34(2): 14099-1422.
- [28] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition(CVPR). Las Vegas: IEEE, 2016: 770-778.
- [29] Szegedy C, Vanhoucke V, Ioffe S, et al. Rethinking the inception architecture for computer vision[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition(CVPR). Las Vegas: IEEE, 2016: 2818-2826.
- [30] Chollet F. Xception: Deep learning with depthwise separable convolutions[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition(CVPR). Hawaii: IEEE, 2017: 1251-1258.
- [31] 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): 92-98. Lü Mengqi, Zhang Ruixiang, Jia Hao, et al. Research on seed classification based on improved ResNet[J]. Journal of Chinese Agricultural Mechanization, 2021, 42(4): 92-98. (in Chinese with English abstract)
- [32] Srivastava N, Hinton G, Krizhevsky A, et al. Dropout: A simple way to prevent neural networks from overfitting[J]. Journal of Machine Learning Research, 2014, 15: 1929-1958.
- [33] Howard A, Sandler M, Chu G, et al. Searching for mobileNetV3[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision(ICCV). Seoul: IEEE, 2019: 1314-1324.

## Classification of *Camellia oleifera* based on Enhanced VGG16 network

Meng Zhichao<sup>1</sup>, He Leiying<sup>1,2</sup>, Du Xiaoqiang<sup>1,2\*</sup>, Zhang Guofeng<sup>1,2</sup>, Yao Xiaohua<sup>3</sup>, Wu Shunkai<sup>1</sup>, Guo Haojian<sup>1</sup>

(1. Faculty of Mechanical Engineering & Automation, Zhejiang Sci-Tech University, Hangzhou 310018, China; 2. Key Laboratory of Transplanting Equipment and Technology of Zhejiang Province, Hangzhou 310018, China; 3. Research Institute of Subtropical Forestry, Chinese Academy of Forestry, Hangzhou 311400, China)

**Abstract:** *Camellia oleifera* is one of the four largest woody oil plants in the world. There are the largest planting area and yield among all woody oil plants, most of which are distributed in more than 1 100 counties and cities in 18 provinces of southern China. Various kinds of *Camellia oleifera* cultivars have emerged in the market, particularly with the continuous development of *Camellia oleifera* industry in recent years. Therefore, a rapid and accurate classification can be urgent to identify the different varieties of *Camellia oleifera*. Deep learning can also be expected to serve as a promising way for classification, due mainly to the strong performance in many operations, such as classification, detection, and segmentation. Although the VGG and ResNet network models have been commonly used for the classification in agriculture, some limitations still remain, such as too large weight space and low accuracy. In this study, a systematic classification was performed on the image of *Camellia oleifera* plant leaf using the enhanced VGG16 network. Four cultivars of *Camellia oleifera* were selected to test, including the Changlin No. 3, No. 4, No. 40, and No. 53 taken from the National *Camellia oleifera* Seed Base of Dongfanghong Forest Farm, Jinhua City, Zhejiang Province, China. An Epson Perfection V30 scanner was used to collect the data set of the *Camellia oleifera* leaf images with a clear texture. 1800 images were obtained for the four *Camellia oleifera* varieties each. The training set and test set were then divided into the proportion of 3:2, where the training set was 4 320, and the test set was 2 880. Some operations of data enhancement were performed on the image during training, such as brightness adjustment, and random enhancement. The Enhanced VGG16 network was constructed using the hard-Swish and ReLU6 activation function, Residual block, Dropout, L2 regularization, the inter-layer deletion, and structural adjustment for the VGG16 network model. The performance of Enhanced VGG16 network model was then evaluated to compare with the classical convolutional neural networks (AlexNet, VGG16, Resnet50, InceptionV3, Xception). The model was also trained and tested on the sampled *Camellia oleifera* data set. More importantly, the hyperparameters dominated the model training and performance. The test optimizer was selected as the RMSprop optimization, while the loss function was categorical\_crossentropy, as well as the Batch Size and the learning rate were 4, and 0.000 01, respectively. Furthermore, the learning rate was reduced to 10% of the original, if the accuracy of two epochs training sets remained the constant, as the number of iterations increased. The epochs were set to 100, while the training stopped in advance, if the accuracy of the four epochs training set remained. The results show that the accuracy of the training and test set of the Enhanced VGG16 network model were 98.98% and 98.44%, respectively. The average detection time of a single image was 55.32 ms, and the space of weight was 90.6 MB. The accuracies of validation and test set were improved by 3.08 and 2.05 percentage points, respectively, compared with the original. The space of weight and the average detection time were reduced by 165.4 MB and 2.18 ms, respectively, indicating the better performance of the enhanced model. Additionally, the enhanced VGG16 network model performed better in the classification of the *Camellia oleifera* leaf, and was much easier to deploy in the mobile terminals and embedded devices, compared with AlexNet, VGG16, Resnet50, InceptionV3, and Xception networks. The *Camellia oleifera* leaf data set can be further expanded to enhance the images at different growth stages, in order to overcome the interference of light in the unstructured and field environment. This finding can also provide promising technical support for crop species identification.

**Keywords:** deep learning; *Camellia oleifera* leaf; classification; Enhanced VGG16; hard-Swish; ReLU6