

# 基于深度学习的芒果病虫害分类识别

曹欢,方睿

(成都信息工程大学 计算机学院,四川 成都 610225)

**摘要:**传统芒果病虫害防治,需要人工进行识别,现引入深度学习技术,可快速准确地对芒果病虫害进行识别。以攀西地区芒果的12种病虫害为研究对象,采用的数据集一部分来自公开数据集MangoLeafBD,另一部分由爬虫技术获得的网络图片组成,共获取图片6769张,其中4879张为训练集,1220张为验证集,670张为测试集。为迎合实际应用的需要,选择了MobileNetV3、MobileViT等4种不同规模的轻量级深度学习网络模型,结合迁移学习训练策略进行对比实验,比较了各个模型的参数量、精确率、召回率等参数。实验结果显示,MobileViT模型用于芒果病虫害分类识别效果最佳,该模型的精确率为96.31%,召回率为96.12%,F1为96.20%,均优于其他模型。由此表明,模型具有较好的鲁棒性和识别性能,可为芒果病虫害分类识别提供技术参考。

**关键词:**芒果病虫害识别;轻量级卷积神经网络;MobileViT;迁移学习;MangoLeafBD

中图分类号:TP391.41;S41

文献标识码:A

文章编号:1673-629X(2023)10-0115-05

doi:10.3969/j.issn.1673-629X.2023.10.018

## Classification and Identification of Mango Diseases and Pests Based on Deep Learning

CAO Huan, FANG Rui

(School of Computer, Chengdu University of Information Technology, Chengdu 610225, China)

**Abstract:** Traditional mango pest control requires manual identification. Now deep learning technology is introduced to quickly and accurately identify mango pests and diseases. In this paper, 12 diseases and insect pests of mango in Panxi area are taken as the research object. One part of the data set is from the public data set MangoLeafBD, and the other part is composed of network pictures obtained by crawler technology. A total of 6 769 pictures were obtained, of which 4 879 were training sets, 1 220 were validation sets, and 670 were test sets. In order to meet the needs of practical applications, four lightweight deep learning networks with different scales, such as MobileNetV3 and MobileViT, were selected. Combined with the transfer learning training strategy, comparative experiments were carried out to compare the parameters, accuracy and recall rate of each model. The experimental results show that the MobileViT model has the best effect on the classification and recognition of mango diseases and insect pests. The accuracy rate of the model is 96.31%, the recall rate is 96.12%, and the F1 is 96.20%, which are better than that of other models. It is showed that such model has excellent robustness and recognition performance, which can provide technical reference for mango pest classification and recognition.

**Key words:** mango diseases and pests identification; lightweight convolutional neural network; MobileViT; transfer learning; MangoLeafBD

### 0 引言

芒果属漆树科热带常绿大乔木,早结丰产,种植效益高,经过多年培育发展,中国现已成为世界第二大芒果生产国,仅次于印度,年产量约为435.1593万吨,占世界芒果总产量的11.2%。芒果可直接食用,也可以入药,随着近年来芒果种植面积的不断扩大,已在中国形成三大芒果主产区:海南昌江,广西百色,攀枝花

盐边。受当前温室效应的影响,芒果培育期也随温度增高而延长,利于产量的增加。芒果大面积种植后,若有病虫害发生,需要大量人力物力进行监管,还需通过人工进行识别,引入深度学习技术,可加快智慧农业的步伐,在发病早期监测到病害,可减少农药的用量,定点施药,将对芒果产量、质量的影响降到最低,还能减少对环境的污染。

收稿日期:2022-11-09

修回日期:2023-03-09

基金项目:国家重点研发计划(2020YFA0608000)

作者简介:曹欢(1999-),女,硕士研究生,通讯作者,研究方向为图形图像处理;方睿(1974-),男,教授,硕导,研究方向为基于网络的计算机应用。

该文以攀西地区芒果常见的 12 种病虫害图片建立数据集,在 Pytorch 框架下搭建 MobileNetV2、MobileNetV3、ShuffleNetV2、MobileViT 这 4 种不同规模的卷积神经网络模型,并结合当前常用的训练策略迁移学习,对比不同模型的参数量、精确率、召回率等,最终选择了 MobileViT 为芒果病虫害识别模型,该模型的精确率为 96.31%,召回率为 96.12%,F1 为 96.20%,表现均优于其他模型。

## 1 相关研究

当前已有大量深度学习技术用于农业病虫害识别的研究。陈伟文等<sup>[1]</sup>利用改进后的网络 AlexNet-Improved 对番茄叶片病害图片进行识别,原模型的过拟合问题得到改善,模型准确率为 95.8%,比改进前高 2%;蒋力顺等<sup>[2]</sup>用 AlexNet、VGG-16 和 ResNet50 三种不同深度的网络对黄瓜病害识别进行研究,最终选择 ResNet50 进行病害识别,识别率为 99.06%;王超等<sup>[3]</sup>用 Resnet50 对玉米叶片病害进行识别,分类准确率为 92.82%,实验结果明显优于传统机器学习方法。

上述文章采用的网络,参数庞大,不利于实际应用,于是很多研究者引入轻量级模型进行识别。叶鸣炆等<sup>[4]</sup>利用 MobileNetV2 对玉米病害进行识别,相比于经典卷积神经网络 LeNet 等,计算量少,准确率高;李好等<sup>[5]</sup>引入 ECA 模块改进 ShuffleNetV2 模型识别农作物病害,平均准确率达到 99.24%,还使用了 H-Swish 激活函数以使网络参数更小;林建吾等<sup>[6]</sup>引入 Focal Loss 代替交叉熵函数在 MobileNetV2 模型上对番茄病害图像进行识别,最终准确率为 99.87%,模型大小仅 8.74 M。上述研究引入轻量级模型,在保证识别准确率的同时,也能用于移动端,满足了实际需求。

## 2 用于病虫害识别的轻量级模型

卷积神经网络由一开始的 AlexNet<sup>[7]</sup>、VGGNet<sup>[8]</sup>,再到现在广泛使用的 ResNet<sup>[9]</sup>、Vision Transformer<sup>[10]</sup>等,这些模型随着硬件的快速发展,参数量十分庞大,训练模型也需耗费巨大的资源,不利于移植到移动端广泛应用,因此出现了轻量级网络模型,如 ShuffleNet<sup>[11]</sup>、MobileNet<sup>[12]</sup>等。

### 2.1 MobileNet 系列

MobileNet 系列是由 Google 团队于 2017 年提出的。MobileNetV1 的一大创新是提出了深度可分离卷积,常规的卷积块作用于图像后,会同时改变特征图的大小和个数,而深度可分离卷积就是将改变通道大小的操作(PW 卷积)和改变特征图大小的操作(DW 卷积)分离开,极大地减少了计算量。如图 1 所示,左边为传统卷积,右边为深度可分离卷积。Mobile-

NetV2<sup>[13]</sup>在此基础上提出了倒残差结构,模块与 ResNet 类似,都由三个模块组成,先是  $1 \times 1$  卷积块,再是  $3 \times 3$  卷积块,最后是  $1 \times 1$  卷积块。ResNet 是利用这三个模块先降维后升维,而 MobileNetV2 的作用结果正好相反,是先升维后降维,所以称为倒残差结构,这样能使特征提取在高维进行,还使用了 ReLU6 激活函数。MobileNetV3<sup>[14]</sup>在前两个版本的基础上再进行了优化。

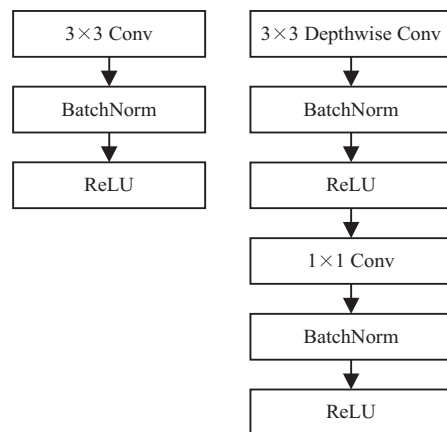


图 1 传统卷积(左)与深度可分离卷积(右)

### 2.2 ShuffleNet 系列

ShuffleNetV1 于 2018 年提出,当时轻量级神经网络模型研究已经趋于成熟,该作者主要用了分组 DW 卷积和 Channel Shuffle 技术,在保证模型精度的前提下,极大地减少了计算开销。卷积中大量使用的  $1 \times 1$  卷积块,在实际计算过程中复杂度较高,消耗了大量计算资源,由此想到进行分组 DW 卷积,在分组卷积之后利用 Channel Shuffle 技术,获取不同通道间的信息。ShuffleNetV2<sup>[15]</sup>在 V1 的基础上进行了进一步的改进,提出了四个设计轻量化网络的原则:当输入和输出通道相同时,内存访问成本最小;分组数过大会导致内存访问成本增加;碎片化程度高,不利于并行运算;元素操作如 ReLU 太多,会增加内存的消耗。

### 2.3 MobileViT

苹果公司于 2021 年提出 MobileViT<sup>[16]</sup>,该模型是基于轻量级模型 MobileNets,并加入当前的研究热门 ViT 模型,以此结合 CNN 和 ViT 各自的优势而提出的。CNN 的卷积层可获得图片的局部信息,以及有归纳偏置特性。ViT 适用于大尺度图像识别,需要在较大的数据集上进行预训练,才能达到 CNN 的识别精度,而且也没有归纳偏置特性,但可以很好地获取图像的全局特征。MobileViT 很好地结合了两种模型的特性,通过深度可分离卷积和 ViT 分别学习图像的局部信息和全局信息,该模型有更好的泛化性能和鲁棒性。图 2 为 MobileViT 的模型架构,其中的 MV2 模块就是 MobileNetV2 中的倒残差结构,↓2 为两倍下采样。

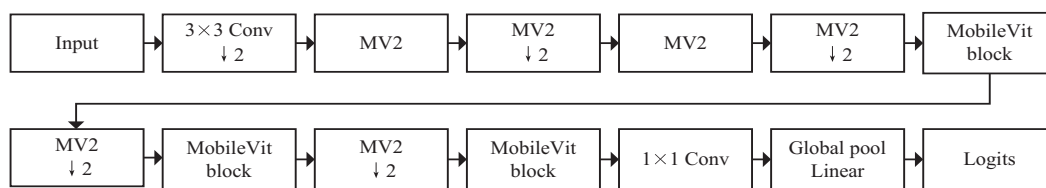


图 2 MobileViT 模型

续表 1

| 病虫害名称  | 图片数量 | 训练集 | 验证集 | 测试集 |
|--------|------|-----|-----|-----|
| 橘小实蝇   | 573  | 413 | 103 | 57  |
| 流胶病    | 586  | 422 | 106 | 58  |
| 切叶象甲   | 516  | 372 | 93  | 51  |
| 炭疽病    | 595  | 429 | 107 | 59  |
| 细菌性黑斑病 | 586  | 422 | 106 | 58  |
| 蚜虫     | 535  | 386 | 96  | 53  |
| 叶瘿蚊    | 581  | 418 | 105 | 58  |

### 3 实验与结果分析

#### 3.1 数据集制作

图像根据何平等<sup>[17]</sup>总结出的四川攀西地区芒果主要病虫害,选取了其中的白粉病、疮痂病、横线尾夜蛾、蓟马、蚧壳虫、橘小实蝇、流胶病、切叶象甲、炭疽病、细菌性黑斑病、蚜虫、叶瘿蚊 12 种病虫害,以此选取图片建立数据集。

数据集一部分图片来自网络公开的 MangoLeafBD 数据集<sup>[18]</sup>,该数据集包含 8 个类别,选取其中的白粉病、切叶象甲、炭疽病、细菌性黑斑病 4 种病虫害图片。还有一部分通过爬虫技术,在百度搜索引擎获取图片。网络获取的图片质量参差不齐,先通过人工筛选为芒果病虫害相关图片,再对图片进行数据增强,改变图像亮度或进行图片翻转,如图 3 所示,以此扩充数据集,增加数据的鲁棒性。

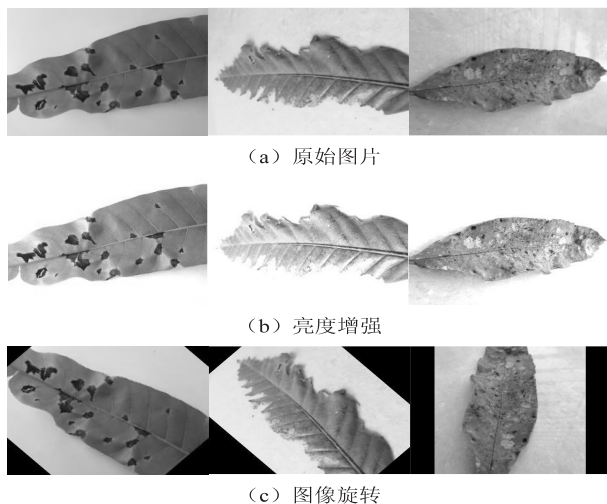


图 3 数据增强

最后共获取 6 769 张图片,具体数量见表 1,其中 4 879 张图片为训练集,1 220 张图片为验证集,670 张图片为测试集。

表 1 芒果病虫害图片数量

| 病虫害名称 | 图片数量 | 训练集 | 验证集 | 测试集 |
|-------|------|-----|-----|-----|
| 白粉病   | 597  | 430 | 108 | 59  |
| 疮痂病   | 579  | 418 | 104 | 57  |
| 横线尾夜蛾 | 517  | 373 | 93  | 51  |
| 蓟马    | 539  | 389 | 97  | 53  |
| 蚧壳虫   | 565  | 407 | 102 | 56  |

#### 3.2 实验环境与网络训练

实验软件环境为 Windows10 操作系统,以 Pytorch 为框架,使用 python3.10 版本的编程语言,CUDA11.6 编程平台。实验硬件环境为 16 GB 内存,NVIDIA GeForce RTX3050 Ti 显卡,4 GB 显存,搭载 AMD Ryzen 7 5800H 处理器。训练参数 epoch 为 100, batchsize 为 16,学习率为 0.001,优化器为 AdamW,激活函数为 ReLU6。

先使模型在大型数据集 ImageNet-1k 上进行预训练,将训练后得到的模型分类器有 1 000 个类别,修改为与数据集相对应的 12 个类别,以此用于芒果病虫害识别。再将数据集 6 769 张图片分为三部分,其中训练集有 4 879 张图片,验证集有 1 220 张图片,测试集有 670 张图片。将图片以  $3 \times 224 \times 224$  的大小,分别输入到 MobileNetV2、MobileNetV3、ShuffleNetV2、MobileViT 等 4 个模型中,其中 MobileViT\_0 为 MobileViT 未使用迁移学习时的模型。实验运用了现在深度学习的常用方法迁移学习,可以使模型具有较好的泛化性能,减少了模型因为数据集较小而产生的过拟合现象。图 4 为部分模型的训练损失曲线。

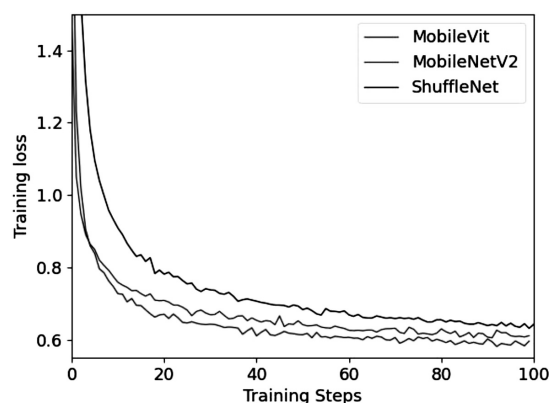


图 4 部分模型的训练损失曲线

### 3.3 评价指标

说明评价指标前需要先了解四个基本概念。真阳性(True Positive, TP)指预测结果为正,实际结果也为正;假阳性(False Positive, FP)指预测结果为正,实际结果为负;假阴性(False Negative, FN)指预测结果为负,实际结果为正。精确率、召回率、F1 分数三个指标就是在这三个参数的基础上进行计算,如下式。精确率是针对预测结果而言,计算预测出来的某类样本中有多少是被正确预测的。召回率是针对测试集而言,有多少样本被正确地预测出来。而 F1 分数同时考虑了精确率和召回率,是精确率和召回率的一种加权平均。

$$\text{精确率} = \frac{TP}{TP + FP} \quad (1)$$

$$\text{召回率} = \frac{TP}{TP + FN} \quad (2)$$

$$\text{F1 分数} = 2 \times \frac{\text{精确率} \times \text{召回率}}{\text{精确率} + \text{召回率}} \quad (3)$$

表 2 各个模型的性能对比

| 模型                | 参数量/Million | FLOPs/M | 精确率/% | 召回率/% | F1 分数/% |
|-------------------|-------------|---------|-------|-------|---------|
| MobileNetV2       | 3.50        | 307.45  | 96.06 | 95.87 | 95.88   |
| MobileNetV3_small | 2.54        | 58.40   | 93.49 | 93.48 | 93.45   |
| MobileNetV3_large | 5.48        | 223.31  | 95.60 | 95.57 | 95.55   |
| ShuffleNetV2      | 5.37        | 588.44  | 95.93 | 95.83 | 95.84   |
| MobileViT_0       | 5.58        | 848.43  | 89.94 | 89.66 | 89.55   |
| MobileViT         | 5.58        | 848.43  | 96.31 | 96.12 | 96.20   |

表 2 为各个模型的性能实验结果。同种模型之间进行比较,如 MobileNetV3\_small 和 MobileNetV3\_large,同种模型参数量越大,模型精确度越高,可见模型大小也会影响到模型的识别效果。当模型较小时,学习能力就会受限制,影响模型的泛化性能,准确率也会受影响。但是模型相差不大的 ShuffleNetV2 和 MobileViT, FLOPs 却相差很多,得益于 ShuffleNet 的 channel shuffle 技术,减少了 1×1 模块的数量,极大地减少了计算量。

表 3 为 MobileViT 模型对各个病虫害图片实验后的数据。实验结果较优的为横线尾夜蛾和切叶象甲两个类别,因为这两个类别的图片数量相对较少,可能实验数据相较其它类别会更单一。蚧壳虫类别的精确率和召回率相差 6.73 百分点,召回率在各个类别中最低为 87.50%,证明此类最容易误判为其它类别,可能与其他类别存在相似性。

MobileViT 模型的精确率为 96.31%,召回率为 96.27%,F1 分数为 96.27%,均优于其他模型,参数量为 5.58 M, FLOPs 为 848.43 M,通过更大的模型和更多的时间,换取更高的识别精度。综上所述,选择

对各个分类类别计算精确率、召回率、F1 分数后,使用宏平均(表示所有类别对应指标的平均值)求取各模型对应的精确率、召回率、F1 分数三个指标。

### 3.4 结果分析

从图 4 可知,随着迭代次数的不断增加,准确率在开始的前 30 个 epoch 增长较快,后面的 70 个 epoch 增长较慢,逐渐趋于稳定,而损失值在前 20 个 epoch 下降较快,后面 80 个 epoch 下降较慢,逐渐趋于稳定。表 2 为各个模型的参数量、FLOPs、精确率、召回率、F1 分数。

MobileViT\_0 相较于使用了迁移学习的 MobileViT 模型,精确率相差 6.37 百分点,召回率相差 6.46 百分点,F1 分数相差 6.65 百分点,证明迁移学习对于模型的准确率增长有明显的作用,可极大地提升模型的泛化能力,当数据量较少的时候,可以通过迁移学习提高模型的准确率,增加模型的识别效果。

MobileViT 模型进行芒果病虫害识别,模型具有较好的识别性能和鲁棒性。

表 3 MobileViT 实验数据

| 病虫害名称  | 精确率/%  | 召回率/%  | F1 分数/% | 测试集大小 |
|--------|--------|--------|---------|-------|
| 白粉病    | 96.72  | 98.31  | 97.51   | 59    |
| 疮痂病    | 98.25  | 98.25  | 98.25   | 57    |
| 横线尾夜蛾  | 100.00 | 100.00 | 100.00  | 51    |
| 蓟马     | 94.12  | 90.57  | 92.31   | 53    |
| 蚧壳虫    | 94.23  | 87.50  | 90.74   | 56    |
| 橘小实蝇   | 91.38  | 92.98  | 92.17   | 57    |
| 流胶病    | 95.00  | 98.28  | 96.61   | 58    |
| 切叶象甲   | 100.00 | 100.00 | 100.00  | 51    |
| 炭疽病    | 96.67  | 98.25  | 97.45   | 59    |
| 细菌性黑斑病 | 98.28  | 98.28  | 98.28   | 58    |
| 蚜虫     | 96.23  | 96.23  | 96.23   | 53    |
| 叶瘿蚊    | 94.83  | 94.83  | 94.83   | 58    |

## 4 结束语

基于迁移学习,从参数量、FLOPs、精确率、召回



率、F1 分数五个方面,比较了 MobileNetV2、MobileNetV3、ShuffleNetV2、MobileViT 这 4 种不同规模的轻量级深度学习网络模型,最终选择了 MobileViT 为芒果病虫害识别模型,该模型的精确率为 96.31%,召回率为 96.12%,F1 分数为 96.20%,表现均优于其他模型。并且 MobileViT 模型的参数量为 5.58 M,有助于解决农业实际部署时,计算资源有限和实时快速准确识别的问题。

#### 参考文献:

- [1] 陈伟文,邝祝芳,王忠伟.基于卷积神经网络的种苗病害识别方法[J].中南林业科技大学学报,2022,42(7):35-43.
- [2] 蒋力顺,董志学,胡 潇,等.基于卷积神经网络的黄瓜病害识别研究[J].计算技术与自动化,2022,41(2):153-157.
- [3] 王 超,王春圻,刘金明.基于深度学习的玉米叶片病害识别方法研究[J].现代农业研究,2022,28(6):102-106.
- [4] 叶名炆,张杰强.基于轻量化网络 MobileNetV2 的玉米病害识别研究[J].现代计算机,2022,28(11):46-50.
- [5] 李 好,邱卫根,张立臣.改进 ShuffleNet V2 的轻量级农作物病害识别方法[J].计算机工程与应用,2022,58(12):260-268.
- [6] 林建吾,张 欣,陈孝玉龙,等.基于轻量化卷积神经网络的番茄病害图像识别[J].无线电工程,2022,52(8):1347-1353.
- [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] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[J]. arXiv:1409.1556, 2014.
- [9] HE Kaiming, ZHANG Xiangyu, REN Shaoqing, et al. Deep residual learning for image recognition[J]. arXiv:1512.03385, 2015.
- [10] DOSOVITSKIY A, BEYER L, KOLESNIKOV A, et al. An image is worth 16x16 words; transformers for image recognition at scale[J]. arXiv:2010.11929, 2020.
- [11] ZHANG Xiangyu, ZHOU Xinyu, LIN Mengxiao, et al. ShuffleNet: an extremely efficient convolutional neural network for mobile devices[J]. arXiv:1707.01083, 2017.
- [12] HOWARD A G, ZHU M, CHEN B, et al. Mobilenets: efficient convolutional neural networks for mobile vision applications[J]. arXiv:1704.04861, 2017.
- [13] SANDLER M, HOWARD A, ZHU M, et al. MobileNetV2: inverted residuals and linear bottlenecks[J]. arXiv:1801.04381, 2018.
- [14] HOWARD A, SANDLER M, CHU G, et al. Searching for MobileNetV3[J]. arXiv:1905.02244, 2019.
- [15] MA N, ZHANG X, ZHEN H, et al. ShuffleNet V2: practical guidelines for efficient CNN architecture design[J]. arXiv:1807.11164, 2018.
- [16] MEHTA S, RASTEGARI M. Mobile vit: light-weight, general-purpose, and mobile-friendly vision transformer[EB/OL]. (2021)[2021-12-18]. <https://arxiv.org/abs/2110.02178>. pd.
- [17] 何 平,余 爽,刘大章,等.四川攀西地区芒果病虫害绿色防控技术[C]//绿色植保与乡村振兴——中国植物保护学会 2018 年学术年会论文集. 西安:中国植物保护学会,2018:286-289.
- [18] AHMED S I, IBRAHIM M, NADIM M, et al. MangoLeaf-BD: a comprehensive image dataset to classify diseased and healthy mango leaves[DB/OL]. <https://arxiv.org/abs/2209.02377>.