

基于深度可分离卷积的苹果叶病理识别^①



王 健¹, 刘雪花²

¹广州大学 华软软件学院 网络技术系, 广州 510900)

²广州大学 华软软件学院 计算机系, 广州 510900)

通讯作者: 王 健, E-mail: 38607607@qq.com

摘 要: 本文以斑点落叶病等 5 种苹果叶病害为研究对象, 设计了一种基于深度可分离卷积的苹果叶病理识别方法. 该方法采用概率数据增强对原始数据集进行扩增, 使用迁移学习探索了深度可分离卷积神经网络在农作物病理识别中的应用; 设计了一种适用于受限设备的深度学习模型以实现苹果叶病害的识别分类, 并将该模型进行压缩和转换, 移植到某嵌入式系统上进行了验证. 实验结果表明该方法在受限设备上的识别率最高仍可达 85.96%, 具有较好的识别效果.

关键词: 迁移学习; 图像处理; 数据增强; 深度卷积网络

引用格式: 王健, 刘雪花. 基于深度可分离卷积的苹果叶病理识别. 计算机系统应用, 2020, 29(11): 190-195. <http://www.c-s-a.org.cn/1003-3254/7710.html>

Pathological Recognition of Apple Leaves Based on Deeply Separable Convolution

WANG Jian¹, LIU Xue-Hua²

¹(Department of Network Technology, South China Institute of Software Engineering, Guangzhou University, Guangzhou 510990, China)

²(Department of Computer, South China Institute of Software Engineering, Guangzhou University, Guangzhou 510990, China)

Abstract: In this study, we take a few kinds of leaf diseases of apple tree, such as *Alternaria mali* Roberts, as research objects, and a pathological identification method for apple tree leaf diseases based on depth-separable convolution is designed. The probability data enhancement is used to amplify the original dataset, a deep separable convolutional neural network is explored by using transductive transfer learning, and is applied to crop pathological recognition. An in-depth learning model for restricted equipment is designed to recognize and classify the apple tree leaf diseases, and the model is compressed, transformed, and transplanted to an embedded system for verification. The experimental results show that the proposed method has a good recognition effect, the recognition rate is up to 85.96% in the restricted equipment.

Key words: transductive transfer learning; image processing; data enhancement; deep convolutional network

随着机器学习的发展与计算机数据处理能力的不断增强, 结合图像处理的机器学习技术在植物病害智能识别上的应用越来越广泛, 并取得了一定的成果^[1,2]. 相关研究表明作物的病害一般都会在叶片上表现出来, 不同的病害种类引起的叶部病理症状也不同^[3]. 由于病

害叶片图像及其病斑的复杂多样性, 一般的特征选择方法因其识别率低, 泛化能力不强, 使得人们很难从病斑图像中提取出有效的分类特征^[4]. 深度学习方法能够利用多个隐含层从大规模数据中逐层、逐级学习到更有用的深层判别特征, 以提高分类的准确性^[5], 因此利

① 基金项目: 2018 年广东省普通高校特色创新项目 (2018KTSCX341); 2018 年广东省创强青年人才项目 (2018KQNCX389); 2018 广州大学华软软件学院创新强校工程 (JXTD201804)

Foundation item: Year 2018, Characteristic Innovation Project of Higher Education of Guangdong Province(2018KTSCX341); Year 2018, Young Talent Program for Innovation Empowering School of Guangdong Province (2018KQNCX389)

收稿时间: 2020-03-12; 修改时间: 2020-04-14, 2020-05-13, 2020-05-26; 采用时间: 2020-06-15; csa 在线出版时间: 2020-10-29

用现代图像处理技术和深度学习方法对作物叶片病害进行诊断和识别已成为目前一个主流研究方向^[6]。深度学习中的卷积神经网络 (Convolutional Neural Network, CNN) 具有较强的特征提取能力^[7], 目前被广泛应用于各个领域的图像分类任务当中。

将 CNN 应用于苹果叶的病理检测需要解决以下两个问题: 第一, 农作物的病理检测普遍由物联网中的边缘计算节点 (Edge Computing Node, ECN) 完成, 虽然传统的 CNN 模型 (如: IDCNNs、GoolgeNet) 已经达到了 95% 的识别率^[4], 但其模型参数规模较大, 无法满足 ECN 对于计算资源有限和低功耗运行的要求。第二, 传统的 CNN 需要大量的训练样本进行学习才能提高模型提取图像特征的能力, 而目前关于农作物叶片病理的公开数据集较少, 无法满足 CNN 的训练要求。针对以上问题, 本文作了如下工作: ① 采用概率数据增强对原始数据集进行了扩增; 在使用小规模的数据集进行训练时仍然可以达到较好的识别精度; ② 通过迁移学习方法对深度可分离卷积网络进行训练, 实现了一种适用于受限设备部署的轻量级模型; ③ 将模型进行压缩和转换, 部署到以 PaddlePi-K210 终端运算模块为核心的嵌入式系统上, 在受限设备上验证了该模型的可行性。

1 系统结构

MobileNets 是为受限设备提出的高效模型, 是一种使用深度可分离卷积来有效构建轻量级深度神经网络低延迟网络架构模型, 可用于移动应用和嵌入式视觉应用的设计^[8]; 迁移学习使 CNN 能获得在大数据集上已经训练得到的较好的特征提取能力, 减小网络模型训练所需的样本数量^[9]。基于上述研究成果, 本文采用了 MobileNets V2 为基础模型, 使用迁移学习的方法进行重新建模。系统流程如图 1 所示, ① 采用 ImageNet 对 MobileNets 进行预训练, 得到预训练模型; ② 将经

过概率数据增强的数据集送入预训练模型进行训练, 并调整网络的全连接层与相应超参数, 得到训练完成模型; ③ 将测试数据集直接送入训练完成后的模型, 输出识别结果; ④ 通过模型压缩与转换, 将转换后的模型移植到 PaddlePi-K210 嵌入式开发板上, 亦可通过摄像头采集作物的信息进行实时的病理叶检测识别。

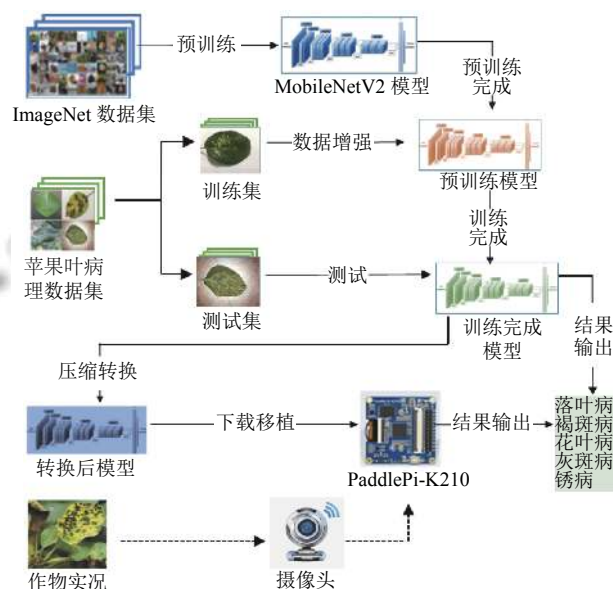


图 1 系统流程图

2 数据集与数据增强

本文采用的公开数据集由西北农林科技大学制作, 分别在西北农林科技大学白水苹果试验站、洛川苹果试验站和庆城苹果试验站进行的采集^[10]。数据集主要在晴天光线良好的条件下获取, 部分图像在阴雨天进行采集, 不同的采集条件进一步增强了数据集的多样性。数据集包括斑点落叶病、褐斑病、花叶病、灰斑病以及锈病等五种常见的苹果叶面病理图像数据, 如图 2 所示, 数量分别为: 斑点落叶病 5343 张, 褐斑病 5655 张, 灰斑病 4810 张, 花叶病 4875 张, 锈病 5694 张。



图 2 苹果页面病理图示

数据增强是扩充数据样本规模的一种有效方法. 当进行神经网络训练时, 数据的规模越大, 模型的质量越高, 泛化能力越强. 本文采取的数据增强思路如下: 在获得一个批次的训练数据集时随机对数据进行裁

剪、翻转, 对亮度、对比度、饱和度、色度和色彩进行调整. 其中, 色彩随机调整了定义两种处理顺序, 按照 3:7 的概率实现. 经过以上处理后, 每种样本数据集的数量扩增 6 倍, 数据增强的效果如图 3 所示.

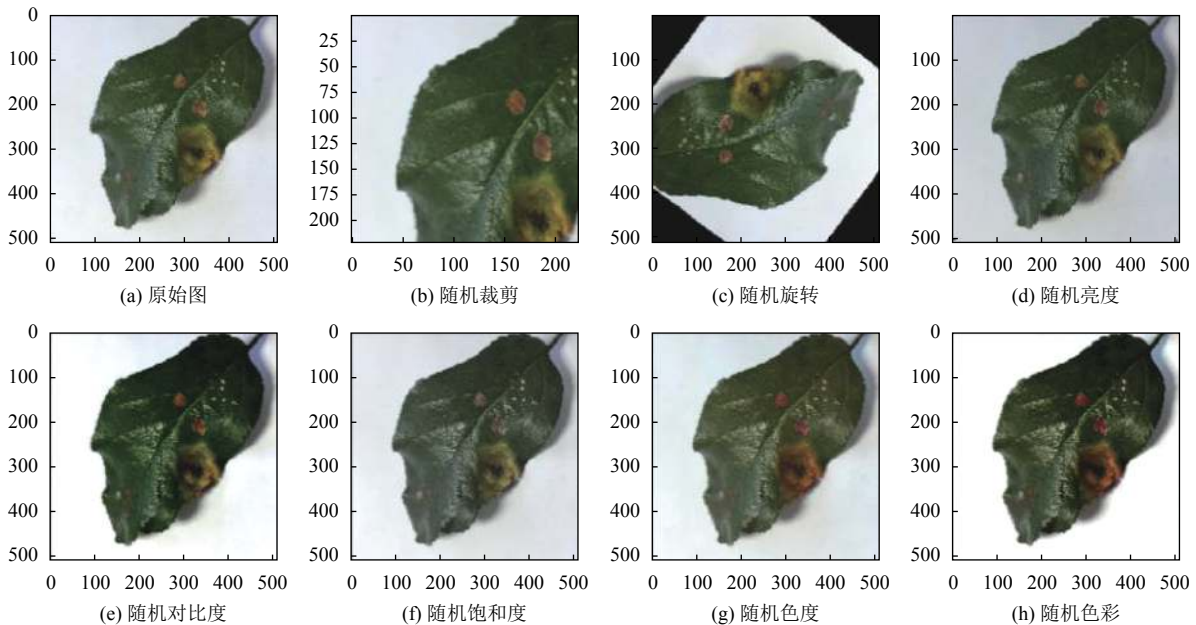


图3 数据增强效果对比图

3 基于迁移学习的深度可分离卷积网络模型构建

迁移学习是机器学习中解决训练数据不足问题的重要工具. 它试图通过放宽训练数据和测试数据, 将知识从源域迁移到目标域^[11]. 本文采用的是基于网络的深度迁移学习方法, 该方法基于以下假设: 神经网络类似于人类大脑的处理机制, 是一个迭代且连续的抽象过程, 网络的前面层可被视为特征提取器, 提取的特征是通用的, 后层为分类功能.

3.1 迁移学习策略

本文通过复用在源域 (ImageNet 图片分类) 中预先训练好的部分网络, 包括其网络结构和连接参数, 再使用网络的一部分将其迁移到目标域 (苹果叶病理分类). 卷积神经网络的迁移学习一般针对卷积基和分类器两个部分来展开工作的. 卷积基由卷积层和池化层的堆栈所组成, 其主要目的是提取图像中的特征; 分类器通常是由多个全连接层组成的, 其主要目标是通过特征对图像进行分类. 卷积基的较低层适用的是通用特征, 而较高层适用的是特殊特征. 本文采用了冻结部分卷积基的迁移学习策略, 其主要思想是将卷积基底

层保持在原始形式, 对高层进行重新训练, 然后使用其输出提供给分类器, 并重新训练分类器. 这种策略适用缺少计算资源, 数据集较小的应用场景.

3.2 深度可分离卷积模型

MobileNet 的核心层为深度可分离卷积层, 研究表明深度可分离卷积和常规卷积是等效的^[12]. 其最初在纹理分类任务中被引入^[13], 随后在 Inception 模型中使用^[14]. 深度可分离卷积由两层组成: 深度卷积 (depthwise convolution) 和 1×1 点卷积 (pointwise convolution).

假设一个普通卷积层的输入特征图为 X , 大小为 $H \times W \times C$, 其中 N 表示样本数, C 表示输入特征图通道数, H 和 W 分别是输入特征图的高和宽. Depthwise 是指将 $N \times H \times W \times C$ 的输入分为 $group=C$ 组, 然后每一组做 3×3 卷积, 即相当于收集了每个 Channel 的空间特征. 深度卷积可表示为式 (1). 其中 \widehat{K} 是大小为 $D_k \times D_k \times M$ 的深度卷积核, 其中, 将 \widehat{K} 中的第 m 个滤波器应用于 F 中的第 m 个通道以产生滤波后的输出特征图 \widehat{G} 的第 m 个通道.

$$\widehat{G}_{k,l,m} = \sum_{i,j} \widehat{K}_{i,j,m} \cdot F_{k+i-1,l+j-1,m} \quad (1)$$

Pointwise 是指对 $N \times H \times W \times C$ 的输入做 k 个普通的 1×1 卷积, 相当于收集了每个点的特征. 与常规卷积网络一样, Depthwise+Pointwise 最终输出任为 $N \times H \times W \times k$, 但深度可分离卷积的计算量缩减为常规卷积的 $1/9$ 左右. 图 4 是常规卷积 (左) 与 MobileNet 中深度可分离卷积 (右) 的结构对比. Depthwise 卷积和 1×1 卷积后都增加了 BN (Batch Normalization, BN) 层和 ReLU6 (Rectified Linear Unit, ReLU). BN 主要作用是加快网络的训练和收敛的速度、控制梯度爆炸防止梯度消失和防止过拟合; BN 可表示为式 (2). $B = \{x_1, \dots, x_m\}$ 为输入数据集合, y_i 为 BN 输出, β 是可训练参数, $\varepsilon(1e-8)$ 是平滑项, 用于在方差极小的情况下, 避免除 0 错. ReLU 是一种非线性激活函数, 其作用是给神经元引入了非线性因素, 而 ReLU6 函数在低精度计算下具有更强的鲁棒性, ReLU6 可表示为式 (3).

$$y_i = \gamma \frac{x_i - \frac{1}{m} \sum_{i=1}^m x_i}{\sqrt{\frac{1}{m} \sum_{i=1}^m \left(x_i - \frac{1}{m} \sum_{i=1}^m x_i\right)^2 + \varepsilon}} + \beta \quad (2)$$

$$F(x) = \min(\max(0, x), 6) \quad (3)$$

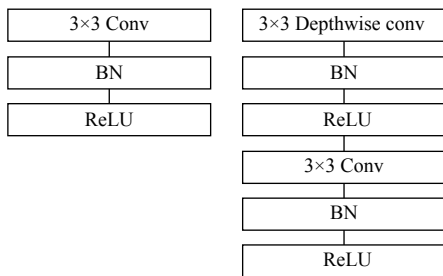


图 4 常规卷积与深度卷积对比

3.3 模型构建

本文使用 MobileNetV2 基本网络通过迁移学习来构建模型. MobileNetV2 结构基于 inverted residual, 保留 Depthwise Separable Convolutions. MobileNetV2 与 V1 的 block 结构如图 5 所示, 图 5(a) 是 V1 的 block, 没有 Shortcut 并带有最后的 ReLU6; 图 5(b) 是 V2 的加入了 1×1 升维, 引入 Shortcut 并且去掉了最后的 ReLU, 改为 Linear. 步长为 1 时, 先进行 1×1 卷积升维, 再进行深度卷积提取特征, 最后通过 Linear 的逐点卷积降维, 将 input 与 output 相加, 形成残差瓶颈结构. MobileNetV2 网络输入的图片大小为 $224 \times 224 \times 3$, 经过 32 个卷积核的初始全卷积层和后续的 19 个残差瓶颈层结构后输出大小为 $1 \times 1 \times 1280$ 的特征向量. 整体结构如表 1 所示.

4 模型训练与结果分析

本文实验所用神经网络训练环境的 GPU 配置为 Tesla P100, 在 AI Studio 平台上使用 paddlePaddle 框架完成, Python 版本为 Python3.5; ECN 采用的是迦南科技的 PaddlePi-K210. 对于系统所采用的主要评价指标为在验证集上的平均精度均值 mAP.

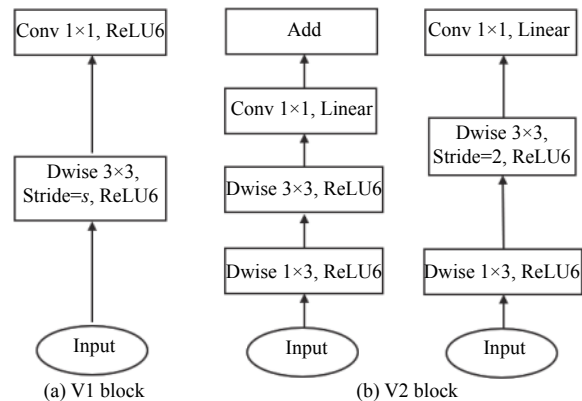


图 5 MobileNetV1 与 V2 block 对比图

表 1 MobileNetV2 网络结构

Input	Operator	t	c	n	s
$224^2 \times 3$	conv2d	-	32	1	2
$112^2 \times 32$	bottleneck	1	16	1	1
$112^2 \times 16$	bottleneck	6	24	2	2
$56^2 \times 24$	bottleneck	6	32	3	2
$28^2 \times 32$	bottleneck	6	64	4	2
$14^2 \times 64$	bottleneck	6	96	3	1
$14^2 \times 96$	bottleneck	6	160	3	2
$7^2 \times 160$	bottleneck	6	320	1	1
$7^2 \times 320$	conv2d 1×1	-	1280	1	1
$7^2 \times 1280$	avgpool 7×7	-	-	1	-
$1 \times 1 \times 1280$	conv2d 1×1	-	k	-	-

4.1 实验参数设置

本文中, 模型参数初始设置是根据已有的 MobileNet 模型应用实例^[15,16], 实验中进行了微调. 调整参数包括优化器、训练批量大小 (batchsize) 和学习率. 按照收敛速度的快慢将常见的 4 种优化器 (SGD、Momentum、RMSProp、Adam) 分成两组, 分别设置两种学习率: SGD 和 Momentum 设为 0.001, RMSProp 和 Adam 设为 0.0001. 设定模型训练停止目标值为 0.95. Batchsize 设置为 10 至 120, 以 10 为步长分别对每个优化器共进行 12 次训练, 最后根据 mAP 和 speed 选取出模型的最优参数设置.

4.2 实验结果分析

如图 6 所示, 从优化器对 mAP 的影响来分析, RMS

优化器的效果最好, 当 *batchsize* 为 50 时, *mAP* 为 0.8596. 虽然其他优化器的最高 *mAP* 均可以达到 0.80 以上, 但存在 ± 0.05 的差距, 可见优化器对于 *mAP* 存在一定的影响; 从训练批量大小对 *mAP* 的影响来看, 没有明显的线性相关; 从训练速度 *epochs* 来分析, 取各优化

器在验证集上的最优结果: $SGD_{mAP}=0.8083$, $Momentum_{mAP}=0.8305$, $RMSProp_{mAP}=0.8596$, $Adam_{mAP}=0.8206$. 绘制 *Loss* 与 *Acc* 曲线, 如图 7 所示, *SGD* 所需迭代次数最多, *Adam* 所需要的迭代次数最少, 表明 *epochs* 与 *mAP* 没有明显的线性相关.

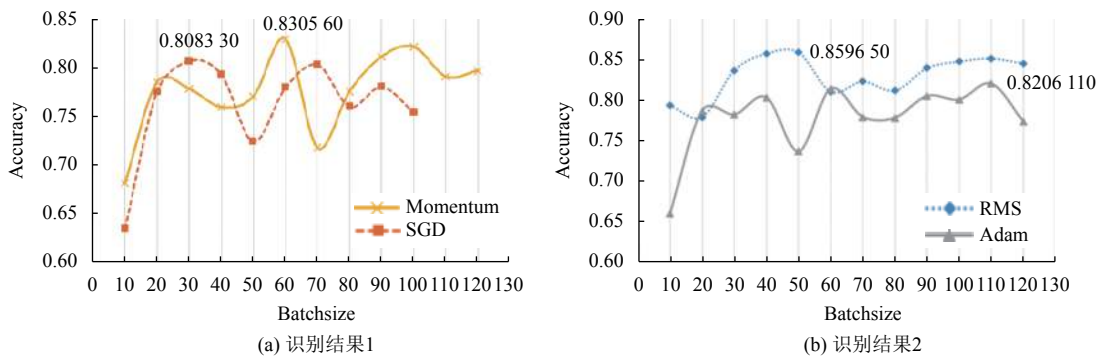


图 6 不同优化器的识别结果

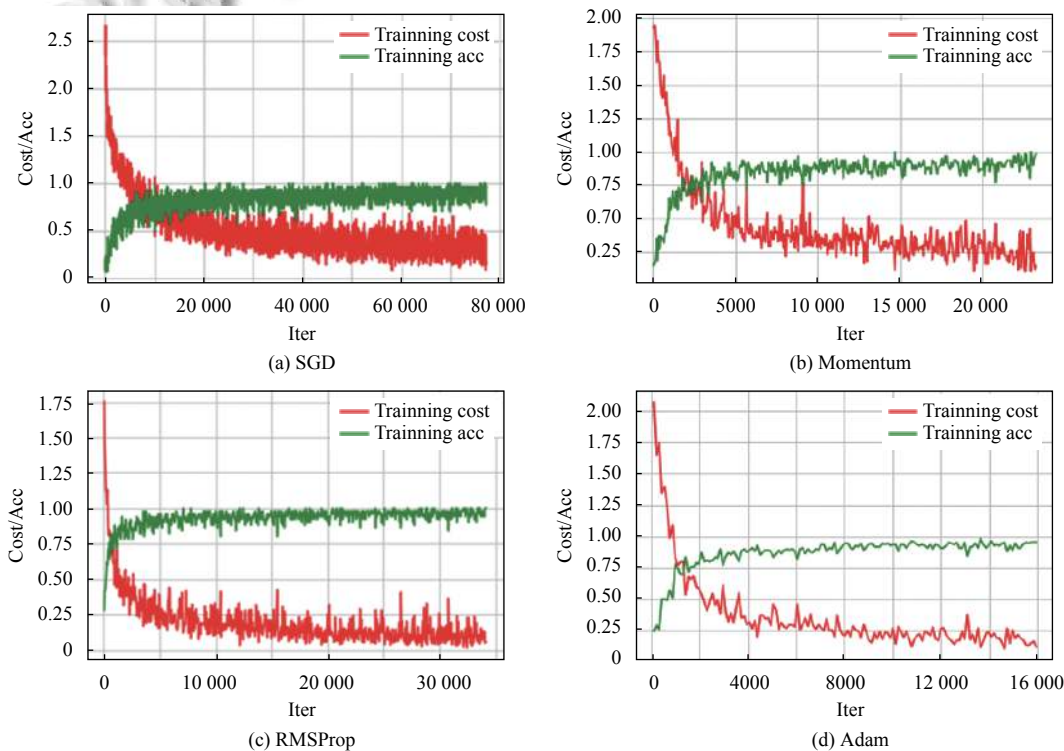


图 7 *Loss* 与 *Acc* 训练曲线

取 *MobileNet* 模型最优结果 $mAP=0.8596$ 与其他识别算法的 *mAP* 进行对比, 如表 2 所示, 该模型的准确率高于大多常规卷积模型, 虽然 *GoogleNet* 的 *mAP* 高于 *MobileNe*, 但是 *GoogleNet* 由于所需计算资源更多, 并不适合部署到嵌入式设备上. 因此, 综合评价表

明该方法优于其他已有方法. 模型在验证集上的各病理的识别结果如图 8 所示 (0-斑点落叶病, 1-褐斑病, 2-灰斑病, 3-花叶病, 4-锈病). 除去灰斑病较低外, 其他病理的识别效果均为良好. 其中少量的灰斑病和一定量的褐斑病被识别成了斑点落叶病, 这是因为这 3 种病

理的图像特征具有一定的相似性,即使在人工分类的情况下也会出现一定的误判。最后对模型进行压缩和量化。为了保证量化后的精度,再次使用训练图片对模型进行调整。将最终模型移植到 PaddlePi-K210 终端运算模块上,并连接 IR 摄像头,系统能实时对苹果叶病理进行检测,且运行稳定,验证了模型的有效性,效果如图 9 所示。

表 2 模型准确值对比

模型	准确率(mAP)	模型	准确率(mAP)
MobileNetV2	0.8596	VGG-16	0.7655
GoogleNet	0.9771	LeNet	0.7315
Res-Net	0.8124	AlexNet	0.7123

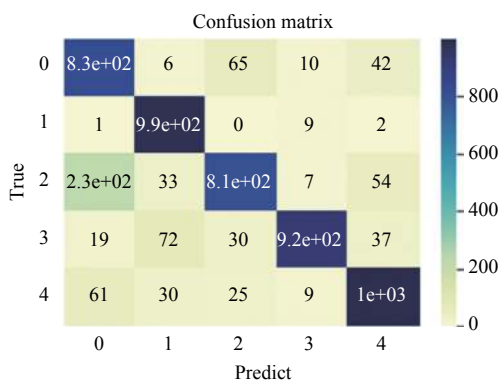


图 8 混淆矩阵图



图 9 原型效果图

5 结语

本文提出了一种基于深度可分离卷积的苹果叶病理识别方法,在 5 种苹果病害叶片图像数据集上进行了数据集扩增和一系列实验,对网络模型的相关参数进行了优化,该模型能有效的对苹果病害叶片病理进行识别。实验结果表明该方法对苹果病理识别具有较好的应用效果。

参考文献

1 Fang Y, Ramasamy RP. Current and prospective methods for plant disease detection. *Biosensors*, 2015, 5(3): 537–561.

[doi: 10.3390/bios5030537]

- Martinelli F, Scalenghe R, Davino S, *et al.* Advanced methods of plant disease detection. A review. *Agronomy for Sustainable Development*, 2015, 35(1): 1–25. [doi: 10.1007/s13593-014-0246-1]
- 车腾腾, 冯益明, 吴春争. “3S”技术在精准林业中的应用. *绿色科技*, 2010, (10): 158–162. [doi: 10.3969/j.issn.1674-9944.2010.10.067]
- 张善文, 张晴晴, 李萍. 基于改进深度卷积神经网络的苹果病害识别. *林业工程学报*, 2019, 4(4): 107–112.
- Schmidhuber J. Deep learning in neural networks: An overview. *Neural Networks*, 2015, 61: 85–117. [doi: 10.1016/j.neunet.2014.09.003]
- 贾庆节, 齐国红, 忽晓伟. 一种基于模糊 C 均值聚类的多阈值苹果病害图像分割方法. *智能计算机与应用*, 2019, 9(2): 63–66, 72. [doi: 10.3969/j.issn.2095-2163.2019.02.014]
- Brahimi M, Arsenovic M, Laraba S, *et al.* Deep learning for plant diseases: Detection and saliency map visualisation. Zhou JL, Chen F. *Human and Machine Learning*. Cham: Springer, 2018.
- Howard AG, Zhu ML, Chen B, *et al.* MobileNets: Efficient convolutional neural networks for mobile vision applications. arXiv: 1704.04861, 2017.
- 许景辉, 邵明焱, 王一琛, 等. 基于迁移学习的卷积神经网络玉米大斑病与锈病的识别. *农业机械学报*, 2020, 51(02): 230–236, 253.
- 周敏敏. 基于迁移学习的苹果叶面病害 Android 检测系统研究 [硕士学位论文]. 杨凌: 西北农林科技大学, 2019.
- Tan CQ, Sun FC, Kong T, *et al.* A survey on deep transfer learning. *Proceedings of the 27th International Conference on Artificial Neural Networks*. Rhodes, Greece. 2018.
- Guo JB, Li YX, Lin WY, *et al.* Network decoupling: From regular to depthwise separable convolutions. *Proceedings of British Machine Vision Conference*. Newcastle, UK. 2018. 248.
- Sifre L, Mallat S. Rigid-motion scattering for texture classification. arXiv: 1403.1687, 2014.
- Chollet F. Xception: Deep learning with depthwise separable convolutions. arXiv: 1610.02357, 2016.
- Hwang S, Yu G, Jung C, *et al.* Attention-aware linear Depthwise convolution for single 27-7710 super-resolution. arXiv: 1908.02648, 2019.
- Chollet F. Xception: Deep learning with Depthwise separable convolutions. *Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Honolulu, HI, USA. 2017. 1800–1807.