

doi: 10.12133/j.smartag.2021.3.1.202009-SA004

面向植物病害识别的卷积神经网络精简结构 Distilled-MobileNet模型

邱文杰¹, 叶进^{1*}, 胡亮青¹, 杨娟², 李其利³, 莫贱友³, 易万茂¹

(1. 广西大学 计算机与电子信息学院, 广西南宁 530004; 2. 广西大学 农学院, 广西南宁 530004;

3. 广西壮族自治区农业科学院 植物保护研究所, 广西南宁 530007)

摘要: 卷积神经网络(CNN)的发展带来了大量的网络参数和庞大的模型体积,极大地限制了其在小规模计算资源设备上的应用。为将CNN应用在各种小型设备上,研究了一种基于知识蒸馏的结构化模型压缩方法。该方法首先利用VGG16训练了一个识别率较高的教师模型,再将该模型中的知识通过蒸馏的方法迁移到MobileNet,从而大幅减少了模型的参数量。将知识蒸馏后的Distilled-MobileNet模型应用在14种作物的38种常见病害分类中。进行了知识蒸馏在VGG16、AlexNet、GoogleNet和ResNet 4种不同网络结构上的表现测试,结果表明,当VGG16作为教师模型时,模型准确率提升了97.54%;使用单个病害识别率、平均准确率、模型内存、平均识别时间4个指标对训练好的Distilled-MobileNet模型进行真实环境下准确性评估,经测试,模型的平均准确率达到97.62%,平均识别时间缩短至0.218 s,仅占VGG16模型的13.20%,模型大小压缩仅为19.83 MB,相比于VGG16缩小了93.60%,使其具备了较高的准确性和实时性要求。本方法模型在压缩内存体积和缩短识别时间上较传统神经网络有了明显提高,为内存和计算资源受限设备上的病害识别提供了新的思路。

关键词: 病害识别; 深度学习; 模型压缩; 知识蒸馏; 卷积神经网络

中图分类号: TP3-0

文献标志码: A

文章编号: 202009-SA004

引用格式: 邱文杰, 叶进, 胡亮青, 杨娟, 李其利, 莫贱友, 易万茂. 面向植物病害识别的卷积神经网络精简结构Distilled-MobileNet模型[J]. 智慧农业(中英文), 2021, 3 (1): 109-117.

QIU Wenjie, YE Jin, HU Liangqing, YANG Juan, LI Qili, MO Jianyou, YI Wanmao. Distilled-mobilenet model of convolutional neural network simplified structure for plant disease recognition[J]. Smart Agriculture, 2021, 3 (1): 109-117. (in Chinese with English abstract)

1 引言

卷积神经网络(Convolutional Neural Networks, CNN)由于具有自动提取特征和无压力处理高维数据等优势,已经在各个领域都有了非

常广泛的应用,在农业病虫害识别方面的应用已经成为了热门的研究方向^[1]。例如, Karleka和Seal^[2]提出了一种识别大豆病害的方法,将大豆的叶片从复杂的背景中分割出后,通过设计CNN模型SoyNet进行分类; Tetila等^[3]对模型中

收稿日期: 2020-09-29 修订日期: 2020-12-03

基金项目: 国家重点研发计划(2017YFD020106-3); 广西创新驱动发展专项项目(桂科AA17204059-9); 地方综合性研究型大学“五有”领军型人才培养研究与实践(2018JGZ103); 2020年教育部新农科研究与改革实践项目

作者简介: 邱文杰(1997—),男,硕士,研究方向为计算机视觉与农业病虫害识别。E-mail: qiuwenjie1997@163.com。

*通讯作者: 叶进(1970—),女,博士,教授,研究方向为计算机软件与系统开发。电话: 13877192800。E-mail: yejin@gxu.edu.cn。

的网络权重进行分析,并使用简单线性迭代聚类(Simple Linear Iterative Clustering, SLIC)方法对图片分割出大豆叶片进行了识别,准确率高达99.04%。Adeel等^[4]提出了一种新的识别框架对葡萄病害进行早期检测,通过提取 AlexNet 和 ResNet101 经过迁移学习后的特征,使用 YEAk 技术选择出最好的特征,之后使用并行方法融合这些强特征,最后输入最小二乘支持向量机(Least Squared Support Vector Machine, LSSVM)中进行分类,对从 PlantVillage 数据集得到的受感染葡萄叶片进行模拟,获得了99%的准确率。Zhang等^[5]使用基于深度学习的改进的 GoogleLeNet 和 Cifar-10 模型,用于玉米叶片病害的识别,平均识别准确率为98.9%。陈桂芬等^[6]使用迁移学习的方法在预训练模型的基础上训练了卷积神经网络,识别准确率达到96.6%。吴华瑞^[7]使用深度残差网络模型,对番茄四种常见病害进行了识别,准确率达到了95%以上。上述病害识别方法的准确率均达到了90%以上,由此可见,CNN已经在农业病虫害识别领域有了广泛的应用,并取得了不错的效果。

然而,CNN经过多年的发展,网络结构越来越复杂,网络参数不计其数,需要耗费大量的存储资源跟计算资源。经典的VGG-16图像分类网络包含1.38亿个参数,而最新的ResNet-200仍然包含超过6000万个参数。网络越复杂,消耗的磁盘空间和内存就越大,并且识别速度就越慢。研究表明,在计算资源受限的ARM平台上,单张图片的识别时间达到13.2 s以上^[8],远远不能满足用户对于实时性的要求。即使将模型部署在云服务器,也无法从本质上解决复杂神经网络模型内存占用大、识别时间长的问题,且部署在服务器上的模型识别时间易受到网络环境波动的影响,无法在网络受限的环境中使用。因此,如何对复杂的神经网络模型进行压缩,使其满足在终端设备使用时的实时性要求和更低的内存占用限制,是目前神经网络亟待解决的问题。

针对上述情况,近年来有学者提出了不同的

模型精简方法。Han等^[9]提出了ClipDL模型,通过同时减少模型参数的数量以及减少仅对关键数据的计算来加速深度学习系统运行,将模型训练过程平均加速了2.32倍,模型大小基本保持不变,而准确性损失仅为1.86%。Rahman等^[10]针对移动设备提出了一种具有两个阶段的小型CNN模型,该模型在准确率上达到了期望准确率的93.3%,并且具有更小的结构规模,跟VGG16对比,其结构规模减少了99%。Shih等^[11]提出了具有扩张卷积和多尺度特征串联的简化区域提议网络(Reduced Region Proposal Network, RRPN),模型的检测准确率相比ZF-Net提高了2.2%~60.2%,相对压缩了ZF-Net网络的81.3%的规模大小。Hc等^[12]提出了一种称为块更改学习(Block Change Learning, BCL)的方法,该方法通过更改由“层”组成的“块”来执行局部和全局知识提炼。在专门针对特征提取任务当中,BCL仅表现出约5%的性能下降,而其他方法的模型性能下降约为17%。Wu等^[13]提出了用于压缩视频动作识别的框架来对模型进行压缩,压缩后的模型准确率下降了0.35%,但参数量上的压缩率达到2.4倍,计算量减少1.2倍。

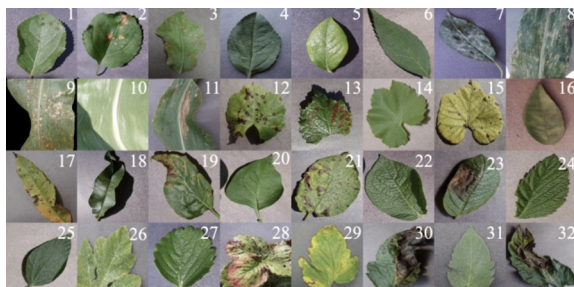
目前,上述模型压缩方法存在压缩率不高和压缩后模型精度下降的问题,因此,本研究结合PlantVillage^[14]植物病害数据集,提出了一种基于知识蒸馏^[15]的轻量级卷积神经网络精简方法,在广西壮族自治区农业科学院提供的含有复杂背景的数据集上做了真实环境下的模型性能测试,结果表明,该方法不以损失模型精度为代价,能够大幅降低模型内存的要求。本方法具有通用性,无论是将模型部署在云服务器还是本地设备中,都可以在少量的教师机内存占用和训练开销降低的同时,降低模型参数量、提高模型精度。

2 数据样本集

本研究使用的数据集来自PlantVillage标准数据集和自建数据集两个部分。

2.1 PlantVillage 数据集

PlantVillage 数据集包含来自 14 种植物的 54,305 张植物叶片图，共 38 种常见病害。叶片图像均为实验室环境下拍摄的单一背景图像，每类的图像数量从 152 张到 5507 张不等。将数据集按照 80% 和 20% 的比例划分为训练集和验证集。数据集中部分病害的为害状如图 1 所示。由于病原体感染农作物的方式相似，不同的农作物可能会感染相同的病原体，因此导致图 1 中不同农作物会患有同种疾病。



注：1. 苹果黑星病 2. 苹果黑腐病 3. 苹果锈病 4. 苹果健康叶片 5. 蓝莓健康叶片 6. 樱桃健康叶片 7. 樱桃白粉病 8. 玉米褐斑病 9. 玉米锈病 10. 玉米健康叶片 11. 玉米大斑病 12. 葡萄黑腐病 13. 葡萄埃斯卡真菌病 14. 葡萄健康叶片 15. 葡萄叶枯 16. 桔子黄龙病 17. 桃子菌斑 18. 桃子健康叶 19. 辣椒菌斑 20. 辣椒健康叶片 21. 马铃薯早疫病 22. 马铃薯健康叶片 23. 马铃薯晚疫病 24. 覆盆子健康叶片 25. 大豆健康叶片 26. 南瓜白粉病 27. 草莓健康叶片 28. 草莓叶焦病 29. 番茄菌斑病 30. 番茄早疫病 31. 番茄健康叶片 32. 番茄晚疫病

图 1 PlantVillage 数据集部分展示

Fig. 1 Part of the samples of the PlantVillage dataset

为使识别模型具有更好的鲁棒性^[16]，本研究还使用了基于深度卷积生成对抗网络（Deep Convolutional Generative Adversarial Networks, DCGAN）的数据增强方法^[17]模拟产生了 5861 张新图像参与建模，部分样本如图 2 所示。经过感知哈希算法^[18]测试，生成模型产生的新样本和真实样本之间的平均相似度达到了 98.75% 以上，表明产生的图像和真实图片相类似，能够作为扩充样本。

2.2 自建数据集

在使用 PlantVillage 数据集的基础上，本研

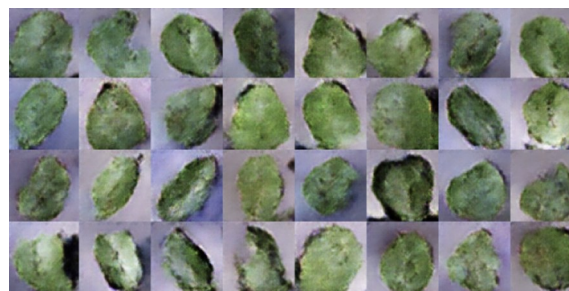
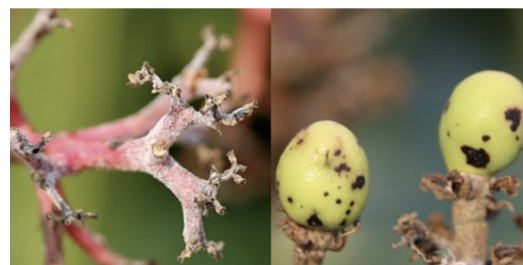


图 2 深度卷积对抗生成网络生成样本

Fig. 2 Samples of deep convolutional generative adversarial network

究还使用了一个自建数据集，其中包括 4835 张芒果炭疽病图片和 4810 张芒果白粉病图片，如图 3 所示。该数据集由广西壮族自治区农业科学研究院植保所提供，图片使用佳能 5D mark III 型数码相机拍摄，像素为 4475×3840，图片中病斑点轮廓所占面积大于叶片面积的 10%。自建数据集中的图片均为真实环境下拍摄的具有复杂背景的照片，主要用于模型的测试。表 1 给出了本研究所用数据集的详细分布情况。



(a) 芒果白粉病

(b) 芒果炭疽病

图 3 自建数据集样本

Fig. 3 Samples of self-built dataset

表 1 试验数据集结构

Table 1 Experimental data set

试验样本	图片数量/张
PlantVillage 数据集	54,305
DCGAN 样本扩充	5861
数据增强	76,639
自建数据集	9645

3 识别模型训练及其精简结构设计

本研究提出了一种结构化的病害识别方法及轻量级神经网络精简方法，图 4 为本研究的整体设计，大致分为以下三个步骤。

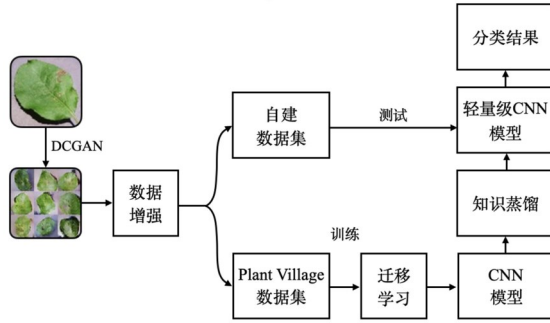


图4 研究整体结构图

Fig. 4 Overall structure of diagram

(1) 采用DCGAN产生大量数据样本, 实现数据扩充。此外, 本研究对扩充后的数据使用了几种常见数据增强方法, 包括平移、旋转、缩放、剪裁和翻转等, 从而提高了模型的平移、缩放等不变性, 有效地防止了模型过拟合。

(2) 为了加快模型在训练过程中的收敛速度, 本研究使用迁移学习^[19]的方法, 将不同网络结构在ImageNet数据集上预先训练好的权值迁移到对应的不同网络结构中, 在此基础上训练不同神经网络模型。

(3) 在训练好的模型上使用知识蒸馏方法, 将知识从冗余的集成模型转移到轻量级模型中, 并保留复杂网络的性能, 从而训练得到精简模型, 通过这种方法有效地降低模型体积, 提高模型精度。

3.1 识别模型 MobileNet 的训练

MobileNet 是 Google 提出的轻量级神经网络, 用于解决复杂模型在移动端和嵌入式终端上无法使用的问题^[20]。相比于其他神经网络, MobileNet 的计算复杂度和模型内存都很小, 网络参数总量为 132 万, 内存仅占 VGG16 的 3%。因此, 本研究选择 MobileNet 作为学生模型进行训练。MobileNet 模型的核心是深度可分离卷积 (Depthwise Separable Convolution, DSC), 原理如图 5 所示, DSC 将标准卷积分解成深度卷积和逐点卷积。

假设神经网络的输入图片维度为 $D_F \times D_F \times M$, 深度卷积的卷积核大小为 $D_K \times D_K \times 1$, 个数为

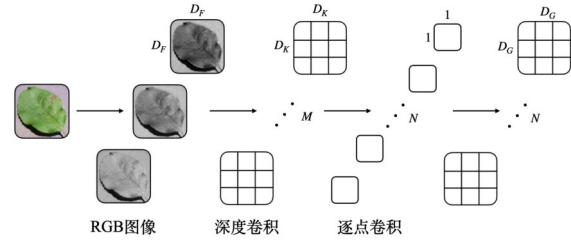


图5 深度可分离卷积原理

Fig. 5 Depthwise separable convolution

M , 逐点卷积的卷积核大小为 $1 \times 1 \times M$, 个数为 N , 经过卷积后输出维度为 $D_G \times D_G \times N$, 则深度可分离卷积的计算成本与常规卷积操作的比值如公式 (1) 所示。

$$\frac{DSC}{Std\ conv} = \frac{1}{N} + \frac{1}{D_K^2} \quad (1)$$

其中, DSC 和 $Std\ conv$ 分别表示深度可分离卷积和常规卷积操作的参数量; N 为逐点卷积核的个数, 个; D_K 为深度卷积核的大小。

由公式 (1) 可知, 当卷积核个数越少、尺寸越小时, 深度可分离卷积相对于常规卷积操作就可以减少更多的参数量。

3.2 精简模型 Distilled-MobileNet 设计

在知识蒸馏的学习过程中, 学生模型是通过模仿教师模型在相同样本下的输出来进行训练的, 在传统 *Softmax* 分类器中, 给定任何输入图片 x , 模型生成一个向量 $\mathbf{S}'(x) = [S'_1(x), S'_2(x), \dots, S'_k(x)]$, 其中, $S'_k(x)$ 代表第 k 种病害所对应的得分, 在神经网络的最后, 使用 *Softmax* 作为分类器将神经网络的输出 $\mathbf{S}'(x)$ 转化成概率分布 $\mathbf{p}'(x)$, 即 $\mathbf{p}'(x) = [p'_1(x), p'_2(x), \dots, p'_k(x)]$, 对于 $\mathbf{p}'(x)$ 中的每个元素 $p'_k(x)$, 其计算方法如公式 2 所示。

$$p'_k(x) = \frac{e^{s'_k(x)}}{\sum_j e^{s'_j(x)}} \quad (2)$$

Hinton 等^[15]提出, 训练有素的教师模型的输出会无限接近于 One-Hot 编码的真实输出, 这就造成有用的类间信息在训练过程中被忽视, 进而直接导致学生模型的训练效果不理想, 因

此, 需要使用温度标度“软化”这些概率, 如公式 (3) 所示。

$$\tilde{p}_k'(x) = \frac{e^{s_k'(x)/T}}{\sum_j e^{s_j'(x)/T}} \quad (3)$$

其中, T 是可调的超参数, $T > 1$ 。通过在 *Softmax* 分类器中加入 T 这一参数, 学生会类似地产生更“柔和”的分类概率分布 $\tilde{p}'(x)$, 从而保留了样本不同类别之间的概率关系。与使用传统 One-Hot 编码^[15] 的硬标签作为训练目标相比, 由于教师模型经过 *Softer-Softmax* 分类器后输出的软目标很好地保留了样本不同类别之间的概率关系, 所以通常会带来更好的性能。为了获得更好的分类结果, 将知识蒸馏中的损失函数由交叉熵替换为巴氏距离^[21], 令 BC 定义为巴氏系数, 满足: $0 \leq BC \leq 1$, 其关系定义如公式 (4) 所示。

$$BC(P, Q) = \sum \sqrt{P(x)Q(x)} \quad (4)$$

其中, P 、 Q 分别表示教师模型和学生模型; $P(x)$ 和 $Q(x)$ 分别代表教师模型和学生模型的输出。

由公式 (4) 可知, 当 P 、 Q 两个标签趋同时, 巴氏系数值 BC 会随之变大; 而当两个标签完全没有重叠时, BC 将会趋于零。

采用巴氏距离后的 L_{KD} 、 L_{cls} 和学生损失函数 L 分别如公式 (5) ~ (7) 所示:

$$L_{KD} = -\ln(BC(r_k(x), \tilde{p}_k'(x))) \quad (5)$$

$$L_{cls} = -\ln(BC(q_k(x), p_k'(x))) \quad (6)$$

$$L = \alpha L_{cls} + (1 - \alpha) L_{KD} \quad (7)$$

其中, $q_k(x)$ 为样本的真实标签; $r_k(x)$ 为教师模型的输出; α 为可调的超参数; 通常 $T \in \{1, 3, 4, 5\}$, $\alpha \in [0.5, 0.9]$ 。

本研究通过知识蒸馏的方法, 结合公式 (5) ~ (7) 中的损失函数, 训练了精简模型 Distilled-MobileNet。首先, 使用图像的真实标签来预训练教师模型, 之后再使用相同的训练样本作为学生模型的输入并进行训练, 所建立的 Distilled-MobileNet 模型与教师模型相比, 可以达到轻量级的体积。图 6 给出了知识蒸馏的具体结构。

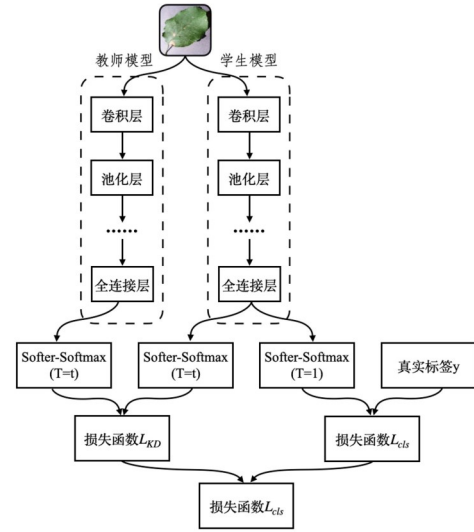


图 6 知识蒸馏结构图

Fig. 6 Diagram of knowledge distillation structure

同时, 在损失函数的设计上, 使用两个不同目标函数的加权平均值, 其中, 损失函数 L_{KD} 用来表示当教师模型和学生模型共享相同的温度 T 时, 教师模型和学生模型输出之间的巴氏距离; 损失函数 L_{cls} 为学生模型和真实标签之间的交叉熵。最后, 通过线性加权的方式计算损失函数 L , 如公式 (8) 所示。

$$L = \alpha L_{cls} + (1 - \alpha) L_{KD} \quad (8)$$

表 2 给出了模型在不同 α 和 T 下的试验结果。经试验验证, 当超参数 α 的值设为 0.6、蒸馏温度 T 的值为 3 时, 模型验证集准确率达到最高 90.62%, 表明模型获得最佳性能。

表 2 参数 α 和 T 对模型验证集准确率的影响

Table 2 Influence of the parameters α and T on the accuracy of the model validation set

单位: %				
T	$\alpha=0.2$	$\alpha=0.4$	$\alpha=0.6$	$\alpha=0.8$
1	83.59	84.93	89.06	79.86
2	84.37	82.81	87.50	77.34
3	81.25	85.15	90.62	80.46
4	75.40	82.13	85.49	83.21

4 试验结果及分析

本试验硬件环境包括具有 Intel i9-10900X (3.20 GHz) 10 核 20 线程 CPU, NVIDIA GeForce RTX2080Ti 11GBx2 的服务器。软件环境为

64位 Windows10 系统, CUDA 9.0, CUDNN 7.0, PyCharm 2018.2。用于训练模型的试验框架的前端和后端分别是 Keras 和 Tensorflow。

为确保试验结果的公平性和有效性,本研究对超参数进行了统一化。将学习率设置为 0.001, batch size 设置为 128, L2 正则化参数设置为 0.0005。

其次,在模型训练中使用了 SGD、Adam 和 Nadam 三种不同的优化器,该类优化算法应用在非凸优化问题中有计算高效、所需内存少等诸多优势,三种优化算法超参数的具体参数如表 3 所示。

表 3 使用三种优化器的 CNN 训练超参数

Table 3 CNN training hyperparameters using three optimizers

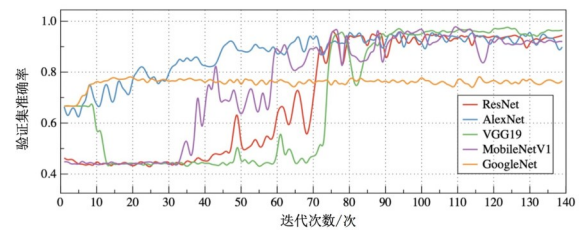
超参数	SGD	Adam	Nadam
Batch Size	32	32	32
Learning Rate	0.001	0.001	0.001
Momentum	0.9	—	—
Decay	None	None	None
β_1	—	0.9	0.9
β_2	—	0.999	0.999
ϵ	—	$1e^{-7}$	$1e^{-7}$

4.1 学习模型性能测试

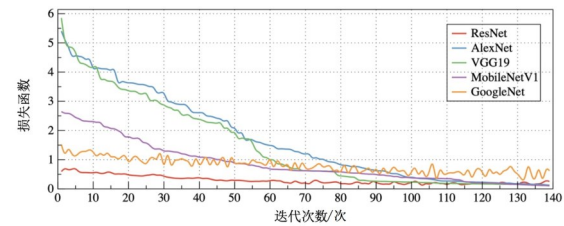
为确定教师和学生网络最佳模型,本研究在 128×128 大小的 3 通道 RGB 图像下对比了 VGG16、AlexNet、GoogLeNet、MobileNet 和 ResNet 5 种主流神经网络结构的验证集准确率、损失函数以及参数量的大小。图 7 给出了迭代过程中验证集准确率与损失函数变化情况。

由图 7 (a) 可以看出,不同模型在迭代过程中验证集准确率变化呈现不同特点。具体来讲,除 GoogleNet 外,5 种模型在迭代过程中其准确率均出现抖动现象,对于 MobileNet 模型,虽在初始迭代时验证集准确率较低,但随着训练的深入,准确率一直呈现上升趋势,且在 75 轮后超越了其余 5 种模型的准确率,说明 MobileNet 模型相较于其他对比模型具有较高稳定性。

表 4 给出了 5 种模型最终验证集准确率的对比较情况。可知, VGG16 和 AlexNet 虽准确率较高,但模型内存在 128×128 的 3 通道输入下内存



(a) 验证集准确率变化曲线



(b) 损失函数变化曲线

图 7 模型验证集准确率与损失函数变化曲线图

Fig. 7 Model validation accuracy rates and loss function change curves

占用分别达到了 1022.6 和 443.0 MB, 这必然会降低某些设备(如智能手机, 物联网设备甚至是计算能力较低的计算机)的运行速度, 同时占用大量内存空间。因此, 本研究将在上述方法的基础上, 使用知识蒸馏方法对模型进行压缩。

表 4 不同网络结构在 PlantVillage 数据集上的表现

Table 4 Performances of different network structures in the PlantVillage data set

网络模型	验证集准确率/%	内存占用/MB
ResNet	94.44	181.3
AlexNet	89.97	443.0
VGG16	96.52	1022.0
MobileNet	92.05	25.1
GoogleNet	76.58	58.1

结合不同模型的内存占用情况, 本研究选用 MobileNet 作为知识蒸馏的学生网络。此结果与本研究之前的设计相一致。

4.2 知识蒸馏在不同模型上的表现

表 5 给出了知识蒸馏在 VGG16、AlexNet、GoogleNet 和 ResNet 4 种不同网络结构上的表现。根据表 5 可知, 由于学生模型不仅学习样本的真实标签, 同时还学习了教师模型产生的保留了类间关系的软化标签, 因而不同网络结构经过知识蒸馏后, 模型准确率均出现了明显提升^[22,23], 平

均提升了1.42%。具体地说，当VGG16作为教师模型时，模型准确率提升了2.04%，达到97.54%，相比于ResNet作为教师模型，准确率

提高了2.38%，因此，经过综合考虑识别准确率和识别时间两个因素，本研究最终选择将VGG16作为教师模型。

表5 知识蒸馏在不同教师/学生模型上的表现

Table 5 The performance of knowledge distillation on different teacher/student models

学生模型	教师模型	准确率/%		模型内存/MB		体积降低/%	识别时间/ ms	
		教师	学生	教师	学生		教师	学生
MobileNet	VGG16	95.50	97.54	304.0	20.8	93.16	1635	261
	AlexNet	93.56	94.89	356.6	21.0	94.11	490	254
	GoogleNet	92.19	94.70	36.6	20.9	42.90	487	260
	ResNet	95.36	95.16	179.8	21.0	88.32	928	249
	平均	—	—	—	—	79.62	—	—

在内存方面，由于学生模型均采用 MobileNet，导致不同网络模型经过知识蒸馏后的内存相差不大，但相较于原教师模型内存则出现明显降低，平均降低79.62%，其中AlexNet模型压缩率最高，达到94.11%。总的来说，试验结果表明，知识蒸馏不仅可以加快模型的运行速度，使其满足实时性要求，同时还能减少模型内存。

4.3 真实环境下的性能测试

为评估 Distilled-MobileNet 模型在真实环境下的表现，本研究在由广西壮族自治区农业科学院提供的真实数据集上用3.2节所述方法对经过知识蒸馏改进的 Distilled-MobileNet 模型进行了性能测试。模型训练过程中的准确率和损失函数变化曲线如图8所示。

从图8中可以看出，在训练过程中，损失函数下降平稳，同时，模型的验证集准确率呈整体上升趋势，且模型的收敛速度很快，在60次迭代后即达到了较好的收敛状态。

使用单个病害识别率、平均准确率、模型内存、平均识别时间4个指标对训练好的 Distilled-

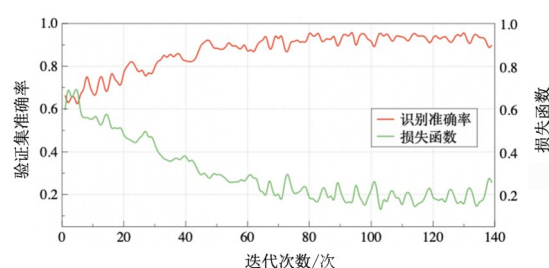


图8 迭代次数对模型识别准确率的影响

Fig. 8 Effects of iteration times on model recognition accuracy and loss function

MobileNet模型进行准确性评估。模型在真实环境下的评估结果见表6。结果表明，模型的平均准确率达到97.62%，平均识别时间缩短至0.218 s，仅占VGG16模型的13.20%，模型大小压缩仅为19.83 MB，相比于VGG16缩小了93.60%。综上，Distilled-MobileNet模型通过将复杂模型转化为简单的小模型，使其能够部署在单片机、智能手机中，从而能够不受不同应用场景对于模型的内存、时间限制，更能满足不同应用场景的要求，由此证明了本方法的有效性和可行性。

表6 真实环境下的模型性能测试

Table 6 Model performance testing in real environment

识别模型	单个病害识别率/%		平均准确率/%	内存占用/MB	平均识别时间/s	总识别时间/s
	芒果白粉病	芒果炭疽病				
VGG16	95.68	94.55	95.15	309.98	1.652	15,636
AlexNet	94.51	94.77	94.64	359.43	0.494	4675
GoogleNet	94.03	95.14	94.58	37.49	0.487	4609
ResNet	94.17	93.20	93.68	183.47	0.943	8925
Distilled-MobileNet	97.59	97.65	97.62	19.83	0.218	2063

5 总结

本研究针对常规的植物病害识别方法存在的识别效率较低的问题,提出了基于知识蒸馏的结构化压缩模型 Distilled-MobileNet,与经典的卷积神经网络模型相比,本方法的优越性是能够在略微提高准确率的同时大幅降低参数内存占用量、缩短识别时间,从而使模型满足更高的实时性要求。本研究对不同模型的性能进行了对比分析,取得的主要试验结果及结论如下。

(1) 对比了 VGG16, AlexNet, GoogLeNet, MobileNet 和 ResNet 共 5 种网络结构的准确率及参数内存占用量,结果表明,虽然 MobileNet 准确率略低于 VGG16 和 ResNet,但网络参数最少,因此,本研究选择 MobileNet 作为知识蒸馏中的学生模型。

(2) 对比了知识蒸馏在不同教师/学生模型上的表现,结果表明,Distilled-MobileNet 模型可以在小幅提高准确率同时明显降低模型的参数量内存,加快模型识别速度。

参考文献:

- [1] KAMILARIS A, PRENAFETA-BOLDÚ F X. A review of the use of convolutional neural networks in agriculture[J]. The Journal of Agricultural Science, 2018, 156(3): 312-322.
- [2] KARLEKAR A, SEAL A. SoyNet: Soybean leaf diseases classification[J]. Computers and Electronics in Agriculture, 2020, 172: ID 105342.
- [3] TETILA E C, MACHADO B B, MENEZES G K, et al. Automatic recognition of soybean leaf diseases Using UAV images and deep convolutional neural networks[J]. IEEE Geoscience and Remote Sensing Letters, 2019, 17(5): 903-907.
- [4] ADEEL A, KHAN M A, AKRAM T, et al. Entropy controlled deep features selection framework for grape leaf diseases recognition[J]. Expert Systems, 2020(1): DOI: 10.1111/exsy.12569.
- [5] ZHANG X, QIAO Y, MENG F, et al. Identification of maize leaf diseases using improved deep convolutional neural networks[J]. IEEE Access, 2018(6): 30370-30377.
- [6] 陈桂芬, 赵姗, 曹丽英, 等. 基于迁移学习与卷积神经网络的玉米植株病害识别[J]. 智慧农业, 2019, 1(2): 34-44.
- CHEN G, ZHAO S, CAO L, et al. Corn plant disease recognition based on migration learning and convolutional neural network[J]. Smart Agriculture, 2019, 1(2): 34-44.
- [7] 吴华瑞. 基于深度残差网络的番茄叶片病害识别方法[J]. 智慧农业, 2019, 1(4): 42-49.
- WU H. Method of tomato leaf diseases recognition method based on deep residual network[J]. Smart Agriculture, 2019, 1(4): 42-49.
- [8] 雷小康, 尹志刚, 赵瑞莲. 基于 FPGA 的卷积神经网络定点加速[J]. 计算机应用, 2020, 40(10): 2811-2816.
- LEI X, YIN Z, ZHAO R. FPGA-based convolutional neural network fixed-point acceleration[J]. Computer Applications, 2020, 40(10): 2811-2816.
- [9] HAN R, LIU C, LI S, et al. Accelerating deep learning systems via critical set identification and model compression[J]. IEEE Transactions on Computers, 2020, 69(7): 1059-1070.
- [10] RAHMAN C R, ARKO P S, ALI M E, et al. Identification and recognition of rice diseases and pests using convolutional neural networks[EB/OL]. arXiv:1812.01043v3 [cs.CV] 4 Mar 2020.
- [11] SHIH K H, CHIU C, LIN J A, et al. Real-time object detection with reduced region proposal network via multi-feature concatenation[J]. IEEE Transactions on Neural Networks and Learning Systems, 2020, 31: 2164-2173.
- [12] HC A, YL A, KCY B, et al. Block change learning for knowledge distillation[J]. Information Sciences, 2020, 513: 360-371.
- [13] WU M, CHIU C, WU K. Multi-teacher knowledge distillation for compressed video action recognition based on deep learning[C]// ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Piscataway, New York, USA: IEEE, 2019.
- [14] HUGHES D P, SALATHE M. An open access repository of images on plant health to enable the development of mobile disease diagnostics[EB/OL]. 2015. arXiv:1511.08060 [cs.CV].
- [15] HINTON G, VINYALS O, DEAN J. Distilling the knowledge in a neural network[EB/OL]. Computer Science, 2015. arXiv:1503.02531 [stat.ML].
- [16] SHORTEN C, KHOSHGOFTAAR T M. A survey on image data augmentation for deep learning[J]. Journal of Big Data, 2019, 6(1): ID 60.
- [17] RADFORD A, METZ L, CHINTALA S. Unsupervised

- representation learning with deep convolutional generative adversarial networks[EB/OL]. 2015. arXiv: 1511.06434v1 [cs.LG] 19 Nov 2015.
- [18] WENG L, PRENEEL B. A secure perceptual hash algorithm for image content authentication[C]// 12th IFIP TC 6/TC 11 International Conference on Communications and Multimedia Security. Berlin Heidelberg, Germany: Springer Verlag, 2011: 108-121.
- [19] PAN S J, YANG Q. A survey on transfer learning[J]. IEEE Transactions on Knowledge and Data Engineering, 2009, 22(10): 1345-1359.
- [20] HOWARD A G, ZHU M, CHEN B, et al. MobileNets: Efficient convolutional neural networks for mobile vision applications[EB/OL]. 2017. arXiv: 1704.04861 [cs.CV].
- [21] KAILATH T. The divergence and Bhattacharyya distance measures in signal selection[J]. IEEE Transactions on Communication Technology, 1967, 15(1): 52-60.
- [22] ASHOK A, RHINEHART N, BEAUNY F, et al. N2N learning: Network to network compression via policy Gradient Reinforcement Learning[C]// 6th International Conference on Learning Representations, ICLR 2018. Vancouver, BC, Canada: International Conference on Learning Representations, ICLR, 2018.
- [23] CHEBOTAR Y, WATERS A. Distilling knowledge from ensembles of neural networks for speech recognition[C]// Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH. San Francisco, CA, United States: International Speech Communication Association, 2016: 3439-3433.

Distilled-MobileNet Model of Convolutional Neural Network Simplified Structure for Plant Disease Recognition

QIU Wenjie¹, YE Jin^{1*}, HU Liangqing¹, YANG Juan², LI Qili³, MO Jianyou³, YI Wanmao¹

(1. School of Computer and Electronic Information, Guangxi University, Nanning 530004, China;

2. School of Agriculture, Guangxi University, Nanning 530004, China;

3. Institute of Plant Protection, Guangxi Academy of Agricultural Sciences, Nanning 530007, China)

Abstract: The development of convolutional neural networks(CNN) has brought a large number of network parameters and huge model volumes, which greatly limits the application on devices with small computing resources, such as single-chip microcomputers and mobile devices. In order to solve the problem, a structured model compression method was studied in this research. Its core idea was using knowledge distillation to transfer the knowledge from the complex integrated model to a lightweight small-scale neural network. Firstly, VGG16 was used to train a teacher model with a higher recognition rate, whose volume was much larger than the student model. Then the knowledge in the model was transferred to MobileNet by using distillation. The parameters number of the VGG16 model was greatly reduced. The knowledge-distilled model was named Distilled-MobileNet, and was applied to the classification task of 38 common diseases (powdery mildew, Huanglong disease, etc.) of 14 crops (soybean, cucumber, tomato, etc.). The performance test of knowledge distillation on four different network structures of VGG16, AlexNet, GoogleNet, and ResNet showed that when VGG16 was used as a teacher model, the accuracy of the model was improved to 97.54%. Using single disease recognition rate, average accuracy rate, model memory and average recognition time as 4 indicators to evaluate the accuracy of the trained Distilled-MobileNet model in a real environment, the results showed that, the average accuracy of the model reached 97.62%, and the average recognition time was shortened to 0.218 s, only accounts for 13.20% of the VGG16 model, and the model size was reduced to only 19.83 MB, which was 93.60% smaller than VGG16. Compared with traditional neural networks, distilled-mobile model has a significant improvement in reducing size and shorting recognition time, and can provide a new idea for disease recognition on devices with limited memory and computing resources.

Key words: disease identification; deep learning; model compression; knowledge distillation; convolutional neural network