

面向边缘计算的轻量级植物病害识别模型

王 冠¹, 王建新¹, 孙 钰²

(1. 北京林业大学信息学院, 北京 100083; 2. 北京航空航天大学网络空间安全学院, 北京 100191)

摘要: 【目的】传统深度学习模型因参数和计算量过大不适用于边缘部署, 在网络边缘的植物病害自动识别是实现长时间大范围低成本作物监测的迫切需求。【方法】联合使用多种模型压缩方法, 得到可部署于算力有限的嵌入式系统的轻量级深度卷积神经网络, 在边缘节点实现植物病害智能识别。模型压缩分 2 个阶段: 第 1 阶段利用基于 L_1 范数的通道剪枝方法, 压缩 MobileNet 模型; 第 2 阶段将模拟学习与量化相结合, 在模型量化的同时恢复识别精度, 得到高精度轻量级的端模型。【结果】在 PlantVillage 数据集 58 类植物病害的实验结果表明: 通道剪枝将 MobileNet 压缩了 3.6~14.3 倍, 量化又将模型的参数精度由 32 bit 降低至 8 bit。整体压缩率达到了 14.4~57.2 倍, 识别准确率仅降低 0.24%~1.65%。与通道剪枝后无模拟学习训练、通道剪枝结合量化后无模拟学习训练这 2 种压缩方法相比, 具有更高的模型压缩率和识别准确率。【结论】联合使用多种模型压缩方法可以少量的精度损失深度压缩人工智能模型, 可为农林业提供面向边缘计算的植物病害识别模型。图 3 表 2 参 23

关键词: 深度学习; 植物病害识别; 边缘计算; 模型压缩; MobileNet

中图分类号: S126 文献标志码: A 文章编号: 2095-0756(2020)05-0978-08

Lightweight plant disease recognition model for edge computing

WANG Guan¹, WANG Jianxin¹, SUN Yu²

(1. School of Information Science and Technology, Beijing Forestry University, Beijing 100083, China; 2. School of Cyber Science and Technology, Beihang University, Beijing 100191, China)

Abstract: [Objective] The traditional deep learning model is not suitable for edge deployment because of too many parameters and too much calculation. Automatic identification of plant diseases on the edge of the network is urgently needed to realize long-term and large-scale low-cost crop monitoring. [Method] By using multiple model compression methods, a light weight deep convolution neural network was obtained, which could be deployed in the embedded system with limited computing power to realize intelligent identification of plant diseases at edge nodes. The model compression was divided into two stages. The first stage used the channel pruning method based on L_1 norm to compress the MobileNet model. In the second stage, simulation learning and quantization were combined to restore the recognition accuracy while the model was quantized, and a high-precision lightweight end model was obtained. [Result] Experimental results of 58 kinds of plant diseases in PlantVillage dataset showed that channel pruning compressed MobileNet by 3.6–14.3 times, and quantization reduced the parameter accuracy of the model from 32 bit to 8 bit. The overall compress rate reached 14.4–57.2 times, and the recognition accuracy was only reduced by 0.24% to 1.65%. Compared with the pruning method trained by common learning, and pruning with quantization trained by common learning,

收稿日期: 2019-10-13; 修回日期: 2020-04-14

基金项目: “十三五”国家重点研发计划项目(2018YFD0600200); 贵州省科研机构能力建设专项(黔科合服企〔2019〕4007); 贵州省科研机构服务企业行动计划项目(黔科合服企〔2018〕4002)

作者简介: 王冠, 博士研究生, 从事人工智能技术及其在林业中的应用研究。E-mail: wangguan@bjfu.edu.cn。通信作者: 王建新, 教授, 博士, 博士生导师, 从事计算机应用技术研究。E-mail: wangjx@bjfu.edu.cn

this method achieved higher compression rate and recognition accuracy. [Conclusion] The combination of multiple model compression methods can compress the artificial intelligence models in depth with only tiny loss of accuracy, and provide plant disease recognition models for agriculture and forestry based on edge computing. [Ch, 3 fig. 2 tab. 23 ref.]

Key words: deep learning; plant disease identification; edge computing; model compression; MobileNet

植物病害是导致现代农林业减产的主要原因, 及时准确的植物病害识别技术是实施有效防治的关键。在实际生产中, 植物病害识别主要依靠人工肉眼观察及经验判断, 需要人们在实地进行持续监测^[1-2]。这种人工评估方法耗时费力且具有一定的主观性, 阻碍了现代农林业的快速发展, 因此, 快速准确的植物病害自动识别成为了精准农业、高通量植物表型和智能温室等领域的研究热点^[3-4]。基于图像处理的植物病害识别方法得到了广泛的研究和应用。早期的识别过程需要从图片中分割病斑, 人工提取病斑特征, 再利用机器学习算法对特征进行分类。HIARY 等^[5] 提取病斑的纹理特征, 采用 k -means 聚类算法和人工神经网络 (artificial neural network, ANN) 对 5 种植物病害进行识别, 准确率达 94%。TIAN 等^[6] 提出用基于支持向量机 (support vector machine, SVM) 的多分类器识别小麦 *Triticum aestivum* 叶部病害。秦丰等^[7] 对 4 种苜蓿 *Medicago* 叶部病害进行识别研究, 分析比较了多种分割方法、特征选择和分类方法。虽然以上方法在特定场景取得了较好效果, 但仍无法实现病害的现场实时诊断。这些方法极大程度上基于阈值的病斑分割算法, 对亮度、物体形态和遮挡程度都非常敏感^[8-9], 都只适合背景单一且对比度高的扫描式图像。此外, 特征提取和选择复杂耗时, 仅局限于有限几种病害, 难以处理复杂背景的大数据。近年来, 深度学习在计算机视觉领域取得重大突破。深度卷积网络神经网络 (convolutional neural network, CNN) 可在大数据中自动端到端提取特征, 避免了人工图像分割和特征工程^[10]。MOHANTY 等^[11] 针对 PlantVillage 数据集^[11] 54 306 张植物病害图像, 使用 AlexNet^[12] 和 GoogLeNet^[13] 识别 38 种植物病害。孙俊等^[14] 在同样的数据集上, 将 AlexNet 进行改进, 提出一种批归一化与全局池化相结合的识别模型。龙满生等^[15] 采用参数精调的迁移学习方式训练 AlexNet, 用于油茶 *Camellia oleifera* 病害图像识别。张建华等^[16] 基于改进的 VGG16 模型, 通过迁移学习实现自然条件下棉花 *Anemone vitifolia* 病害图像分类。DECHANT 等^[17] 提出了集成多个 CNN 的方法, 实现玉米 *Zea mays* 大斑病图像的高精度识别。PICON 等^[18] 利用深度残差网络 ResNet 对 3 种早期小麦病害进行识别, 改善了复杂背景下的病害识别率。通常深度学习模型部署在云平台, 需要将拍摄图像上传至云平台进行识别。但这种方法严重依赖高速的 4G/5G 无线网络和强大的云平台, 不仅无法覆盖广大偏远农田林地, 长时间大范围的上传与识别还导致能耗、流量及云服务成本大幅上涨, 限制了物联网的建设规模。然而, 目前的监控设备借助低成本低功耗加速芯片, 即可支持边缘计算, 仅在发现病害时通过低功耗广覆盖的 NB-IoT 网络^[19] 上报, 可显著降低网络及云服务成本, 促进大规模的农林业物联网普及。但现有的 CNN 模型计算量和参数量过大, 不适用于边缘部署。轻量级模型 MobileNet^[20] 在速度和精度两者间达到了一个较好的均衡, 但其目标平台是手机等高端嵌入式平台, 参数量及运算量仍超过 PaddlePi 等廉价边缘设备的承受能力。近年来, 学术界也提出了多种模型压缩方法。模型通道剪枝^[21] 剪裁掉模型一部分冗余或低权重的卷积核, 减少模型的参数量。量化^[22] 将模型由 32 bit 浮点数转化为定点整数, 减少模型参数占用的空间。然而上述压缩方法仅应用于 ResNet 等重量级模型, 尚未对 MobileNet 等轻量级模型压缩进行优化, 而且这些压缩方法彼此相互独立, 未能联合使用实现模型的深度压缩。为解决上述问题, 本研究提出了面向边缘计算的植物病害识别模型构建方法, 主要贡献为: ①首次针对轻量级模型 MobileNet 实现深度压缩。②通过联合通道剪枝、量化等多种模型压缩方法, 得到了深度压缩的轻量级边缘端模型, 可在廉价边缘节点运行。③将模拟学习方法^[23] 与量化相结合, 实现模型压缩的同时, 提升识别效果, 最后得到的边缘端模型可达到与原模型相近的识别准确率。

1 材料与方法

1.1 材料

本研究使用 PlantVillage 植物病害数据集。PlantVillage 既包含单一背景下的植物叶片扫描式图像,

也收录自然背景下的植物叶片图像，包括叶片重叠、阴影和土壤干扰等情形。截至目前已收集了87 280张图像，包括25种植物和29种病害组成的58类植物-病害组合（图1）。

数据集按图1所示的编号将各种叶片归类并制作标签。随机抽取数据集中60%图像作为训练集，剩余的40%作为测试集。单一背景图像与自然背景图像使用相同的分割比例。由于PlantVillage数据集包含从不同角度对同一叶片拍摄的多张图像，因此相同叶片的图像仅存在于训练集或测试集中。



1. 苹果 *Malus domestica* 健康叶；2. 苹果疮痂病；3. 苹果黑腐病；4. 苹果锈病；5. 香蕉 *Musa paradisiaca* 健康叶；6. 香蕉黑条叶斑病；7. 香蕉叶斑病；8. 蓝莓 *Vaccinium* spp. 健康叶；9. 卷心菜 *Brassica oleracea* 健康叶；10. 卷心菜黑腐病；11. 哈密瓜 *Cucumis melo* 健康叶；12. 木薯 *Manihot esculenta* 褐斑病；13. 木薯绿螨病；14. 芹菜 *Apium graveolens* 早疫病；15. 樱桃 *Prunus* spp. 健康叶；16. 樱桃白粉病；17. 玉米 *Zea mays* 健康叶；18. 玉米灰斑病；19. 玉米锈病；20. 玉米大斑病；21. 黄瓜 *Cucumis sativus* 健康叶；22. 黄瓜霜霉病；23. 茄 *Solanum melongena* 健康叶；24. 葫芦 *Cucurbita* spp. 霜霉病；25. 葡萄 *Vitis vinifera* 健康叶；26. 葡萄黑腐病；27. 葡萄黑痘病；28. 葡萄枯病；29. 洋葱 *Allium cepa* 健康叶；30. 柑橘 *Citrus sinensis* 黄龙病；31. 桃 *Prunus persica* 健康叶；32. 桃细菌性斑点病；33. 辣椒 *Capsicum annuum* 健康叶；34. 辣椒细菌性斑点病；35. 马铃薯 *Solanum tuberosum* 健康叶；36. 马铃薯晚疫病；37. 马铃薯早疫病；38. 圆南瓜 *Cucurbita* spp. 花叶病；39. 树莓 *Rubus* spp. 健康叶；40. 大豆 *Glycine max* 健康叶；41. 大豆霜霉病；42. 大豆灰斑病；43. 大豆壳针孢叶枯病；44. 长南瓜 *Cucurbita* spp. 健康叶；45. 长南瓜白粉病；46. 草莓 *Fragaria* spp. 健康叶；47. 草莓叶焦病；48. 番茄 *Lycopersicum esculentum* 健康叶；49. 番茄细菌性叶斑病；50. 番茄早疫病；51. 番茄晚疫病；52. 番茄壳针孢叶斑病；53. 番茄二斑叶螨病；54. 番茄花叶病；55. 番茄叶霉病；56. 番茄靶斑病；57. 番茄黄化曲叶病；58. 西瓜 *Citrullus lanatus* 健康叶。同一编号的图像中，“-1”为单一背景图像，“-2”为自然背景图像

图1 PlantVillage 数据集植物病害示例图

Figure 1 Example of plant disease images from the PlantVillage dataset

1.2 方法

选择轻量级卷积神经网络 MobileNet^[20]作为本研究的基准模型。MobileNet 模型将传统的卷积分解为一个深度卷积 (depthwise convolution, DC) 和一个卷积核为 1×1 的逐点卷积 (pointwise convolution, PC)，计算速度比传统卷积快8~9倍，主要面向智能手机等高端嵌入式系统。为深度压缩 MobileNet，本研究提出了如图2所示的面向边缘计算的植物病害识别模型二阶段构建方法。

第1阶段使用通道剪枝压缩迁移学习训练的MobileNet模型。与从头训练方法相比，迁移学习可以有效提升模型的识别准确率。迁移学习是用ImageNet数据集上预训练好的参数初始化模型，然后在PlantVillage数据集上通过标准多分类损失函数优化模型参数。最后使用基于 L_1 范数的通道剪枝^[21]精简低权值的卷积核，同时将该卷积核所有的输入输出连接从网络中删除，降低了模型计算量和存储空间。

第2阶段对剪枝后的模型通过量化方法进一步压缩, 得到轻量级的边缘端模型。量化是将模型的权值和激活值由32 bit降低至8 bit, 分为训练时量化和训练后量化。虽然训练时量化方法更适用于轻量级模型, 但直接使用该方法压缩剪枝后的MobileNet模型仍会导致识别精度显著下降, 因此第2阶段将模拟学习与训练时量化相结合, 利用迁移学习的MobileNet监督剪枝后模型的量化训练过程, 实现量化模拟学习, 在模型压缩的同时提升识别准确率。

1.2.1 通道剪枝 模型通道剪枝采用均匀剪枝方法对MobileNet模型的分离卷积层进行通道剪枝, 即每层都减掉同样比例的卷积核。依据需要减少模型的浮点运算数量(floating point operations, FLOPs)来确定每层的剪枝比例。计算各层中每个卷积核权值的绝对值和(即 L_1 范数), L_1 范数越大, 代表该卷积核对模型的贡献越大, 反之越小。每层按 L_1 范数由高到低的顺序排序卷积核, 优先剪枝 L_1 范数低的卷积核。为实现模型的深度压缩, 需进行较高比例的通道剪枝, 分别对模型减掉70%、80%和90%的FLOPs。

1.2.2 量化模拟学习 通常训练时, 量化的损失函数是标准的多分类损失函数。量化模拟学习是用模拟学习损失函数作为训练时量化的损失函数。模拟学习方法使剪枝量化后模型的输出特征尽量接近迁移学习训练的MobileNet输出特征。利用2个输出特征之间的 L_2 范数作为模拟损失函数, 即:

$$L_{L_2}(W_s, W_t) = \|F(x; W_t) - F(x; W_s)\|_2^2. \quad (1)$$

式(1)中: W_t 和 W_s 分别是迁移学习训练的MobileNet和剪枝量化后模型的权值矩阵, $F(x; W_t)$ 和 $F(x; W_s)$ 分别表示这2个模型的输出特征值。

剪枝量化后模型的输出特征再经Softmax归一化得到预测类别概率, 与分类标签比较后得到交叉熵, 作为标准多分类损失函数 $L_{\text{class}}(W_s)$ 。模拟学习的完整损失函数就是分类损失函数与模拟损失函数的权重和:

$$L(W) = L_{\text{class}}(W_s) + \alpha L_{L_2}(W_s, W_t). \quad (2)$$

式(2)中: α 为平衡损失权重的超参数。相较于普通多分类问题的损失函数, 模拟学习方法可提供额外的监督信息。

将训练时量化方法与模拟学习相结合, 实现量化模拟学习具体的训练步骤为: ①在训练的前向传播中, 将模型的权值 w_f 和激活值 a_f 进行量化得到定点值 w_q 和 a_q , 对于浮点数 x 具体的量化过程为:

$$x_{\text{int}} = \text{round}\left(\frac{x}{A}\right); x_Q = \text{clamp}\left[-\left(\frac{N}{2} - 1\right), \frac{N}{2} - 1, x_{\text{int}}\right]. \quad (3)$$

x_Q 即为得到的量化值。其中: clamp函数对于输入的变量 a , b , c 输出为:

$$\begin{aligned} \text{clamp}(a, b, c) &= a & x \leq a \\ &= x & a < x \leq b \\ &= b & x > b. \end{aligned} \quad (4)$$

也就是将浮点数除以缩放因子 A , 再最近邻取整, 最后把范围限制到1个区间内。 N 与量化后整数类型占用的比特数有关。本研究采用有符号8 bit整数类型, $N=256$ 。对于权值, 每层权值的最大绝对值作为缩放因子。对于激活值, 计算各训练批次激活的最大绝对值的滑动平均值作为缩放因子。②计算剪

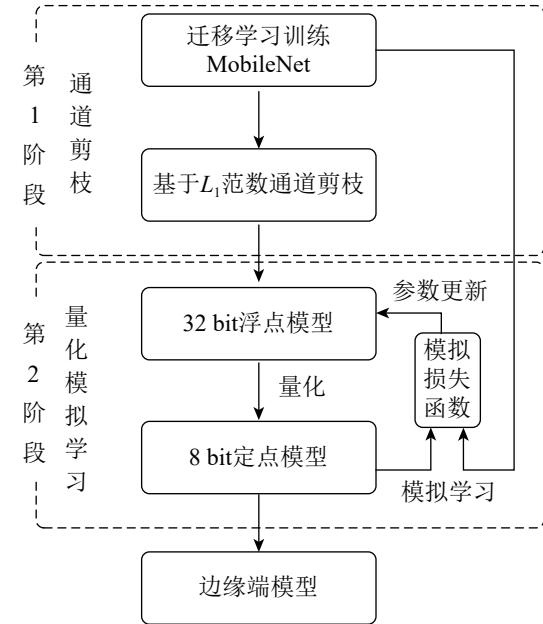


图2 面向边缘计算的植物病害识别模型构建方法: 通道剪枝和量化模拟学习

Figure 2 Plant disease recognition model for edge computing building pipeline: channel pruning and quantized mimic learning

枝量化后模型对迁移学习训练的 MobileNet 进行模拟学习的损失函数, 即计算公式(2), 得到损失值 $L(w_q)$ 。③后向传播过程, 利用步骤②得到的损失函数值对量化之后的权值求梯度, 公式为 $\frac{\partial L(w_q)}{\partial w_q}$ 。④用步骤③计算梯度去更新量化前的浮点值, 也就是将模型的权值反量化回有误差的浮点类型。公式为 $w_f = w_f - v \frac{\partial L(w_q)}{\partial w_q}$, 其中: v 为学习率。因此, 模型的后向传播过程仍然是浮点数计算。⑤重复步骤①至步骤④, 直至完成训练。最后再对模型按照步骤①量化, 得到最终的边缘端模型。

2 结果与分析

模型实现和训练采用的软件环境为 Ubuntu 16.04 操作系统和 PaddlePaddle 深度学习框架, 硬件环境为 GPU 工作站, 使用 NVIDIA Titan X 显卡 (12 GB 显存) 和 AMD Ryzen 7 1700X 处理器 (32 GB 内存)。采用模型的平均识别准确率 (accuracy) 作为衡量模型精度的标准。同时为了更好地评价模型的鲁棒性, 将每类病害样本分别进行测试, 计算每个类别的查准率 (precision)、查全率 (recall) 以及查全率与查准率的加权平均分数, 并在所有类别上求平均。

2.1 MobileNet 迁移学习训练试验

训练 CNN 模型需要对输入图片进行预处理。首先, 利用数据增广技术对原图像进行变换, 将训练图像变换为 256×256 大小, 然后再随机剪枝成 224×224 , 再进行随机水平翻转和随机垂直翻转。该过程极大扩充了训练数据集的多样性, 可提升 CNN 模型的准确率, 降低网络过拟合的风险。之后, 计算训练集的红 (R)、绿 (G)、蓝 (B) 3 个颜色通道的均值和方差, 所有图像都减去该均值, 除以方差, 得到归一化后的数据作为 CNN 的输入, 可加速训练过程收敛。对于测试集中的每一张图片, 需要变换至 224×224 大小, 减去训练集各通道均值, 除以其方差进行归一化后就可以输入 CNN 模型进行识别。

利用迁移学习训练 MobileNet, 使用 ImageNet 数据集预训练的参数初始化模型, 采用批量训练的方法将训练集分为多个批次 (batch), 使用随机梯度下降算法来实现模型优化, 批次大小为 32, 遍历 1 次训练集中的所有图片作为 1 个周期 (epoch), 共迭代 50 个周期, 初始学习率为 0.005, 动量值为 0.9, 之后每迭代 20 个周期就将学习率减小为原来的 0.1 倍。训练好的模型参数量为 3.3 M, 识别准确率为 96.23%, 查准率、查全率和加权平均分数分别为 96.62%、95.46% 和 95.75%。

2.2 边缘端植物病害识别模型训练试验

研究不同压缩率下本研究方法的有效性, 使用不同的剪枝率, 分别对模型减掉不同比例的 FLOPs。结果表明: 当剪枝率低于 60% 时, 即使使用无模拟训练方法重新训练模型, 得到的识别准确率与原 MobileNet 模型差别很小, 说明原模型在该数据集上具有较高的冗余性, 只有当剪枝率高于 70% 时, 才能体现不同压缩方法表现的差距。因此, 设置剪枝率为 70%、80% 和 90%, 对应的模型参数量大小为 0.91、0.58 和 0.23 M, 模型的参数量压缩了 3.6、5.7、14.3 倍, 量化又将精度由 32 bit 降低至 8 bit, 压缩率为 4 倍, 得到的边缘端模型的整体压缩率分别为 14.4、22.8 和 57.2 倍。为快速恢复剪枝后模型精度, 首先利用模拟学习损失函数进行 30 个周期的 32 bit 浮点模型训练, 使用随机梯度下降算法优化模型, 批次大小为 32, 初始学习率为 0.005。之后, 每迭代 15 个周期就将学习率减小为原来的 0.1 倍。公式(2)的 α 值设置为 1。之后再进行 20 个周期的量化模拟学习, 学习率为 0.005, 公式(2)的 α 值为 0.1, 其余超参数值不变。训练结果如表 1 所示。

表 1 表明: 整体压缩率分别为 14.4、22.8 和 57.2 倍的边缘端模型, 识别准确率分别为 95.99%、

表 1 边缘端模型植物病害识别结果

Table 1 Plant disease recognition results of models on the edge

剪枝率/%	参数量/M	剪枝压缩率/倍	量化压缩率/倍	整体压缩率/倍	准确率/%	查准率/%	查全率/%	加权平均分数/%
70	0.91	3.6	4	14.4	95.99	96.18	94.41	94.92
80	0.58	5.7	4	22.8	95.55	95.51	93.52	93.99
90	0.23	14.3	4	57.2	94.58	94.87	92.41	93.15

95.55% 和 94.58%，与迁移学习训练的 MobileNet 模型相比仅下降了 0.24%、0.68% 和 1.65%。同时查准率、查全率和加权平均分数值也表明边缘端模型具有较高的鲁棒性。

不同压缩率的边缘端模型在测试集的混淆矩阵如图 3 所示。图 3 列出了 58 个类中的每类被正确分类的比例(对角线上的值)和被误识别为其他类的比例(非对角线上的值)。每类的编号与图 1 一致。可以看出: 边缘端模型对不同植物的不同病害均具有较强的识别能力, 但不同病害识别结果之间存在着较大的差异。58 类病害中, 这 3 个边缘端模型的识别准确率均超过 90% 的有 43 类, 均超过 80% 的有 51 类, 均超过 70% 的有 55 类。其中有 11 号哈密瓜健康叶、24 号葫芦霜霉病、25 号葡萄健康叶、39 号树莓健康叶、46 号草莓健康叶这 5 类的识别准确率在 3 个模型均达到了 100%。识别效果最差, 在 3 个模型上识别准确率几乎均低于 70% 的病害是 12 号木薯褐斑病 (3 个模型识别率分别为 48.15%、58.33%、48.15%), 35 号马铃薯健康叶 (3 个模型识别率分别为 48.33%、33.33%、48.33%), 21 号黄瓜健康叶 (3 个模型识别率分别为 67.92%、66.04%、77.36%)。这些病害大多都被误识别为外形相似的其他病害, 例如 12 号木薯褐斑病被误识别为 13 号木薯绿螨病, 35 号马铃薯健康叶被误识别为病斑较小的 36 号马铃薯晚疫病, 21 号黄瓜健康叶被误识别为 11 号哈密瓜健康叶。

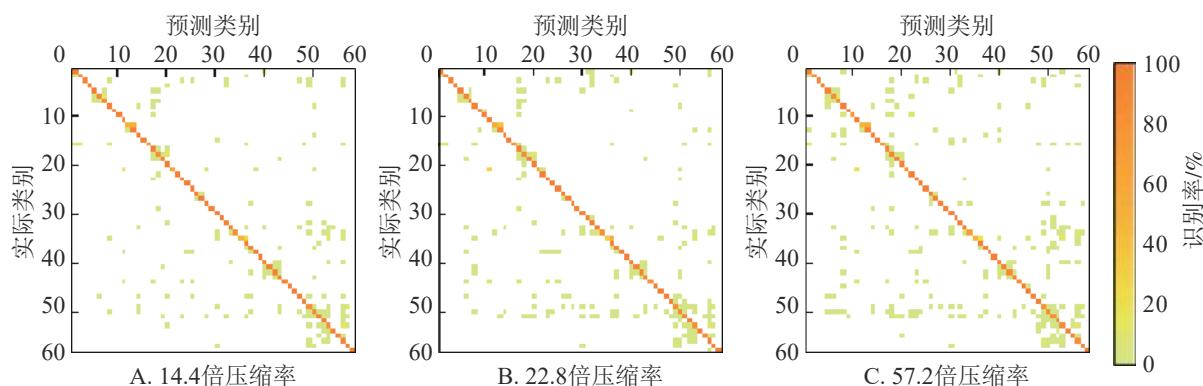


图 3 边缘端模型在测试集的混淆矩阵

Figure 3 Confusion matrix on the test set of models on the edge

2.3 边缘端植物病害识别模型对比试验

为进一步测试边缘端模型性能, 分别在剪枝率 70%、80% 和 90% 的条件下, 利用无模拟学习方法, 即标准的多分类损失函数分别训练通道剪枝后模型和通道剪枝并量化模型, 训练的超参数与本研究的训练超参数一致, 与本研究模型进行对比实验。从表 2 可见: 在不同的剪枝率的情况下, 本研究模型与其他模型压缩方法相比均具有更高的模型压缩率和识别准确率, 而且压缩率越高, 识别准确率相比其他方法提升越明显, 能更好识别植物病害类别并部署于边缘设备。

表 2 不同压缩方法边缘端模型植物病害识别结果

Table 2 Plant disease recognition results of models on the edge compressed by different methods

剪枝率/%	参数量/M	边缘端模型	精度/bit	压缩率/倍	准确率/%
70	0.91	剪枝+无模拟学习	32	3.6	95.48
		剪枝+量化+无模拟学习	8	14.4	95.45
		本研究模型	8	14.4	95.99
80	0.58	剪枝+无模拟学习	32	5.7	94.95
		剪枝+量化+无模拟学习	8	22.8	94.92
		本研究模型	8	22.8	95.55
90	0.23	剪枝+无模拟学习	32	14.3	93.40
		剪枝+量化+无模拟学习	8	57.2	93.53
		本研究模型	8	57.2	94.58

3 结论

本研究针对边缘环境下计算资源的限制，在迁移学习训练的 MobileNet 模型基础上，联合使用 2 种压缩算法降低模型参数量和运算量，并结合模拟学习恢复识别精度，得到深度压缩的边缘端模型。在 PlantVillage 的实验结果表明：利用本研究方法对 MobileNet 进行不同程度的深度压缩，均能够大大减少网络计算量并保留原始识别能力。其中减少 70%~90% FLOPs 的模型，参数量压缩了 3.6~14.3 倍，再经过量化模拟学习后整体压缩率为 14.4~57.2 倍，准确率达到了 95.99%~94.58%，较迁移学习训练的 MobileNet 模型仅降低 0.24%~1.65%，同时还具有较高的鲁棒性，对不同植物的不同病害均具有较强的识别能力。实验结果证明了该压缩方法的可行性和有效性。

随着 PlantVillage 数据集的不断扩展，深度学习模型能更多更准地识别植物病害。本研究提出的模型构建方法可平衡识别的速度和精度，满足植物病害识别边缘部署的需求。

4 参考文献

- [1] SHAMSHIRBAND S, ANUAR N B, KIAH M L M, et al. Survey an appraisal and design of a multi-agent system based cooperative wireless intrusion detection computational intelligence technique [J]. *Eng Appl Artif Intell*, 2013, **26**(9): 2105–2127.
- [2] HILLNHUTTER C, MAHLEIN A. Early detection and localisation of sugar beet diseases: new approaches [J]. *Gesunde Pflanzen*, 2008, **60**(4): 143–149.
- [3] CAMARGO A, SMITH J S. An image-processing based algorithm to automatically identify plant disease visual symptoms [J]. *Biosyst Eng*, 2009, **102**(1): 9–21.
- [4] MUTKA A M, BART R S. Image-based phenotyping of plant disease symptoms [J]. *Front Plant Sci*, 2015, **5**: 1–8.
- [5] HIARY H A, AHMAD S B, REYALAT M, et al. Fast and accurate setection and classification of plant diseases [J]. *Int J Comput Appl*, 2011, **17**(1): 31–38.
- [6] TIAN Yuan, ZHAO Chunjiang, LU Shenglian, et al. Multiple classifier combination for recognition of wheat leaf diseases [J]. *Intell Automation Soft Comput*, 2011, **17**(5): 519–529.
- [7] 秦丰, 刘东霞, 孙炳达, 等. 基于图像处理技术的 4 种苜蓿叶部病害的识别 [J]. 中国农业大学学报, 2016, **21**(10): 65–75.
- QIN Feng, LIU Dongxia, SUN Bingda, et al. Recognition of four different alfalfa leaf diseases based on image processing technology [J]. *J China Agric Univ*, 2016, **21**(10): 65–75.
- [8] BARBEDO J G A. An automatic method to detect and measure leaf disease symptoms using digital Image processing [J]. *Plant Dis*, 2014, **98**(12): 1709–1716.
- [9] BARBEDO J G A. A new automatic method for disease symptom segmentation in digital photographs of plant leaves [J]. *Eur J Plant Pathol*, 2016, **147**(2): 349–364.
- [10] HE Kaiming, ZHANG Xiangyu, REN Shaoqing, et al. Deep residual learning for image recognition [C]// BAJCSY R, LI Feifei, TUYTELAARS T. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas: IEEE Press, 2015: 770–778.
- [11] MOHANTY S P, HUGHES D P, SALATHÉ M. Using deep learning for image-based plant disease detection [J]. *Front Plant Sci*, 2016, **7**: 1419.
- [12] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks [J]. *Commun ACM*, 2017, **60**(6): 84–90.
- [13] SZEGEDY C, LIU Wei, JIA Yangqing, et al. Going deeper with convolutions [C]// BISCHOF H, FORSYTH D, SCHMID C, et al. 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Boston: IEEE Press, 2015: 1–9.
- [14] 孙俊, 谭文军, 毛罕平, 等. 基于改进卷积神经网络的多种植物叶片病害识别 [J]. *农业工程学报*, 2017, **33**(19): 209–215.
- SUN Jun, TAN Wenjun, MAO Hanping, et al. Recognition of multiple plant leaf diseases based on improved convolutional neural network [J]. *Trans Chin Soc Agric Eng*, 2017, **33**(19): 209–215.
- [15] 龙满生, 欧阳春娟, 刘欢, 等. 基于卷积神经网络与迁移学习的油茶病害图像识别 [J]. *农业工程学报*, 2018, **34**(18):

- 194 – 201.
- LONG Mansheng, OUYANG Chunjuan, LIU Huan, *et al.* Image recognition of *Camellia oleifera* diseases based on convolutional neural network & transfer learning [J]. *Trans Chin Soc Agric Eng*, 2018, 34(18): 194 – 201.
- [16] 张建华, 孔繁涛, 吴建寨, 等. 基于改进 VGG 卷积神经网络的棉花病害识别模型[J]. 中国农业大学学报, 2018, 23(11): 161 – 171.
- ZHANG Jianhua, KONG Fantao, WU Jianzhai, *et al.* Cotton disease identification model based on improved VGG convolution neural network [J]. *J China Agric Univ*, 2018, 23(11): 161 – 171.
- [17] DECHANT C, WIESNER-HANKS T, CHEN Siyuan, *et al.* Automated identification of northern leaf blight-infected maize plants from field imagery using deep learning [J]. *Phytopathology*, 2017, 107(11): 1426 – 1432.
- [18] PICON A, ALVAREZ-GILA A, SEITZ M, *et al.* Deep convolutional neural networks for mobile capture device-based crop disease classification in the wild [J]. *Comput Electron Agric*, 2019, 161: 280 – 290.
- [19] WANG Y P E, LIN Xingqin, ADHIKARY A, *et al.* A primer on 3GPP narrowband internet of things [J]. *IEEE Commun Mag*, 2017, 55(3): 117 – 123.
- [20] 陈方. MobileNet 压缩模型的研究与优化[D]. 武汉: 华中师范大学, 2018.
- CHEN Fang. *Research and Optimization of MobileNet Compression Model* [D]. Wuhan: Central China Normal University, 2018.
- [21] HAN S, POOL J, TRAN J, *et al.* Learning both weights and connections for efficient neural networks [C]// CORTES C, LAWRENCE N D, LEE D D, *et al.* *The 28th International Conference on Neural Information Processing Systems*. Montreal: MIT Press, 2015: 1135 – 1143.
- [22] JACOB B, KLIGYS S, CHEN B, *et al.* Quantization and training of neural networks for efficient integer-arithmetic-only inference [C]// BROWN M, MORSE B, PELEG S, *et al.* *2018 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Salt Lake City: IEEE Press, 2018: 2704 – 2713.
- [23] BA J L, CARUANA R. Do deep nets really need to be deep [C]// GHAHRAMANI Z, WELLING M, CORTES C, *et al.* *The 27th International Conference on Neural Information Processing Systems* [C]. Montreal: MIT Press, 2014: 2654 – 2662.