

DOI:10.19651/j.cnki.emt.2212253

基于改进 MobileViT 网络的番茄叶片病害识别

陈晓^{1,2} 夏颖¹(1.南京信息工程大学电子与信息工程学院 南京 210044; 2.南京信息工程大学江苏省
大气环境与装备技术协同创新中心 南京 210044)

摘要: 针对卷积神经网络对番茄叶片型病害分类效果不佳的问题,提出了一种基于改进 MobileViT 轻量级网络的番茄病害识别方法。首先,删除输入和全局表征层的特征融合,将局部和全局表征层进行特征融合,使局部表征与全局表征更加密切相关;其次,为了避免模型在缩放时参数和 FLOPS 的大幅增加,在融合块中使用 1×1 卷积层替换 3×3 卷积层;然后,还添加了输入与融合块的残差结构,优化了网络模型中的更深层次;最后,将 ReLU6 激活函数替换成 H-Swish 激活函数,进一步提高了模型准确率。实验结果表明,改进后的 MobileViT 模型可以很好地实现番茄病害的识别,平均识别准确率达到 99.16%。相较于其它的卷积神经网络模型,具有更高的识别精度。

关键词: 番茄;病害识别;MobileViT;卷积神经网络;Transformer;H-Swish 激活函数

中图分类号: TP391.4 **文献标识码:** A **国家标准学科分类代码:** 520.20

Improved MobileViT network for tomato leaf disease identification

Chen Xiao^{1,2} Xia Ying¹

(1. School of Electronic and Information Engineering, Nanjing University of Information Science and Technology, Nanjing 210044, China; 2. Jiangsu Collaborative Innovation Center of Atmospheric Environment and Equipment Technology, Nanjing University of Information Science and Technology, Nanjing 210044, China)

Abstract: To solve the problem of poor classification effect of convolution neural network on tomato leaf type diseases, a tomato disease identification method based on MobileViT lightweight network was proposed in this paper. Firstly, feature fusion of input and global representation is deleted and local and global representation are fused to make local representation more closely related to global representation. Secondly, in order to avoid the large increase of parameters and FLOPS when the model is scaled, the 3×3 convolution layer is replaced by 1×1 convolution layer in the fusion block. Then, the residual structure of input and fusion blocks is added to optimize the deeper level in the network model. Finally, the model accuracy is further improved by replacing the ReLU6 activation function with the H-Swish activation function. The experimental results showed that the improved MobileViT model can well recognize tomato diseases, with an average recognition accuracy of 99.16%. Compared with other convolution neural network models, it has higher recognition accuracy.

Keywords: tomato; disease identification; MobileViT; convolutional neural networks; transformer; H-Swish activation function

0 引言

2020年3月,粮农组织、世界卫生组织和世界贸易组织总干事联合发表声明称,各国必须采取措施阻止新型冠状病毒肺炎疫情加速蔓延和确保粮食安全^[1]。植物病害一直是粮食生产中的棘手问题^[2]。番茄作为当今世界上最普遍存在的高营养价值农作物之一,受到了许多病害的影响,这些病害的精确识别对农学家来说是一项具有挑战性的任

务^[3]。在农业生产实践中,农民利用经验目测识别番茄病害。但这也带来了效率和可靠性问题^[4]。因为病害种类繁多,症状相似,即使是有经验的农民或者专家也可能无法正确识别番茄病害^[5]。对番茄叶片病害进行准确的自动识别,有助于提高番茄生产管理水平,为番茄提供良好的生长环境。

随着计算机技术的飞速发展,计算机视觉、机器学习、深度学习等技术^[6-16]在农业中得到了广泛的应用。传统机

收稿日期:2022-12-01

器学习方法通过颜色、纹理或形状特征对作物病害的 RGB 图像进行分割^[17]。但不同疾病特征相似,难以判断疾病类型,疾病识别准确率较差^[18]。卷积神经网络(CNN)采用端到端的结构,相比其机器学习简化了识别过程^[19]。使用深度学习技术通过早期疾病检测和后续疾病管理来提高蔬菜、水果以及大田作物的成活率,目前已经进行了多项研究^[20]。Memon 等^[21]将 VGG16, InceptionV3 和 Resnet50 进行并行训练,使用集成方法从不同的模型中学习,在自制的七类病害的棉花数据集上准确率达到 98.53%。Fuentes 等^[22]提出了一种基于 CNN 的番茄叶部病害识别算法,能够快速有效地识别 9 种不同的病虫害,通过几组不同模型比较,最终的最高识别准确率为 83.1%。以上研究证明了 CNN 网络在这一领域的可行性和有效性,但也还存在着一些缺陷和不足,比如深层网络的识别准确率高,但计算量和参数量都较大,训练成本也较高,而浅层网络计算量和参数量都较小,但是不能充分学习到特征,因此识别准确率较低^[23]。另外,卷积神经网络在通过卷积核之后一般只关注到图像的局部特征,而忽略了图像的全局特征^[24]。

近年来,自然语言处理(natural language processing, NLP)领域的 Transformer 模型引起了研究者的极大关注^[25]。Dosovitskiy 等^[26]将 Transformer 用于图像分类任务,在超大规模数据集上的性能超越了 CNN,但是它的训练成本很高;对于中小尺度的数据集,因 Vision Transformer 是对全局信息的直接学习,加上模型参数量较多,一般都是处于欠拟合的,所以表现性能不佳^[27]。苹果公司在 2021 年提出了一种 CNN 与 Vision Transformer

相结合的混合架构模型 MobileViT^[28], 希望能在移动设备或嵌入式终端^[29-35]上实现模型预测,但由于融合块模块产生了缩放问题,并具有复杂的学习任务,因此性能一般。

针对上述问题,本文提出了一种改进 MobileViT 的番茄叶片病害识别方法,首先利用浅层 CNN 提取图像的局部特征信息,再利用 Vision Transformer 网络提取图像全局特征信息,最后将局部特征信息和全局特征信息进行特征融合后与输入特征残差连接,来实现对番茄病害的快速准确识别。实验表明,在自制的番茄数据集上,该方法相比原网络与卷积神经网络能够更准确地实现番茄病害的识别。

1 MobileViT 网络

传统的卷积神经网络是对局部特征信息进行提取,而对全局特征信息的提取能力较差;而 Vision Transformer 模型善于提取全局特征信息,但是其模型参数太大,对计算机的算力要求太高,导致基本无法进行移动端的部署。而 MobileViT 网络是将卷积神经网络和 Vision Transformer 相结合,并提取自身网络的优点建立起来的通用且轻量化的网络模型,在保证精度的同时也具有较好的实时性。

MobileViT 网络主要是由普通卷积、MobileNetV2 模块、MobileViT 模块、全局池化和全连接层共同组成。网络结构如图 1 所示,其中 MV2 是 MobileNetV2 模块的缩写,2×是执行该模块两次,↓2 是执行 stride=2 的下采样操作标志,MViT 是 MobileViT 模块的缩写,h=w=2 表示将划分成的每个 batch 大小设置为 2×2,L=2 表示将 MobileViT 模块中的 Transformer 模块连续执行两次。

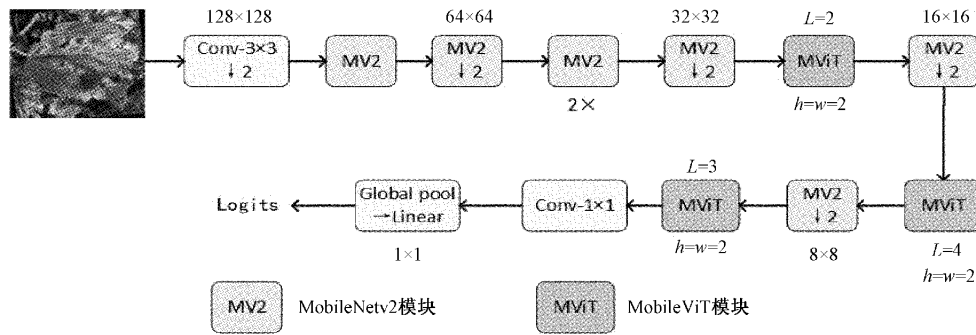


图 1 MobileViT 网络结构图

MobileViT 模块主要由 3 个子模块构成,分别为局部特征模块、全局特征模块和融合特征模块,它可以在参数量较少的前提下,全面地提取出图像中的特征信息。该模块结构如图 2 所示,为了能够学习具有空间归纳偏置的全局表征,首先要将 X_L 展开成 N 个不重叠的 patch $X_U = R^{P \times N \times d}$ 。其中, $P = wh$, $N = \frac{HW}{P}$ 为 patch 的数量, $h \leq n$, $w \leq n$ 为 patch 的高和宽,patch 中的每个像素 $p \in \{1, 2, \dots, P\}$ 通过 Transformer 来进行全局特征建模,得到:

$$X_{G(p)} = Transformer(X_U(p)), 1 \leq p \leq P \quad (1)$$

与丢失像素空间顺序的 Vision Transformer 网络不同,MobileViT 既不会丢失 patch 顺序,也不会丢失每个 patch 内像素的空间顺序。为了减少参数量大小,将卷积得到的局部特征信息通过 Unfold→Transformer→fold 结构,如图 3 所示,为特征图的 Unfold 和 fold 操作示例图,首先对输入的特征图划分成一堆的 patch,以 2×2 大小为例,划分完后再做 Attention 时,是将每个 patch 中对应相同位置的做 Self-Attention,即图中颜色相同的 token 才做 Self-Attention,颜色不同的 token 不做信息交互,这样可减少做 Attention 的参数量。通过 Unfold 方法后得到

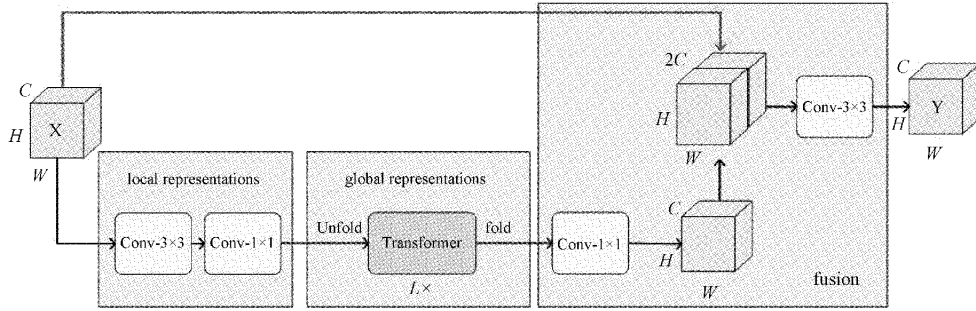


图 2 MobileViT 模块

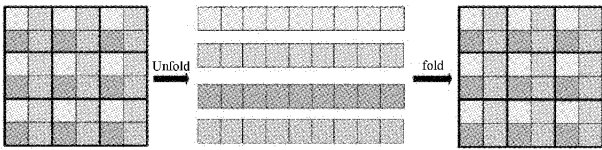


图 3 Unfold 和 fold 操作示例图

4 个序列,再将每个序列输入到 Transformer 模块中进行全局建模,最后再通过 fold 方法将这些特征折叠成原特征图的形状大小。

MobileNetV2 模块是 MobileNetV2^[36] 网络中的倒残差模块,其结构如图 4 所示。首先对输入特征图进行通道数的先增加后减少,再与残差模块相连。特征图内通道数越少卷积层的复杂度就越低,只通过通道数量较小的特征图也很难提取出足够多的特征。MobileNetV2 模块正是权衡了这二者之间的关系,采用了先增加通道数,后减少通道数的方法。所以 MobileNetV2 模块可以用很少的计算量得到很高识别精度。

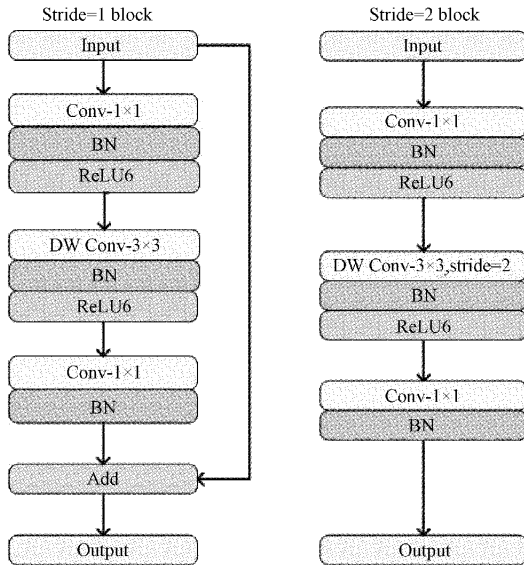


图 4 MobileNetV2 模块

2 改进的 MobileViT 网络

MobileViT 模块是整个 MobileViT 网络的核心。它

首先将特征图通过一个卷积核大小为 3×3 的卷积层进行局部的特征建模,然后通过卷积核大小为 1×1 的卷积层来调整通道的数量,接着通过 Unfold \rightarrow Transformer \rightarrow fold 结构进行全局的特征建模,然后再通过卷积核大小为 1×1 的卷积层将通道数调整回原始大小,接着通过 shortcut 捷径分支与原始输入特征图进行沿通道 channel 方向的 Concat 拼接,最后再通过一个卷积核大小为 3×3 的卷积层做特征融合得到输出。

2.1 局部和全局特征融合

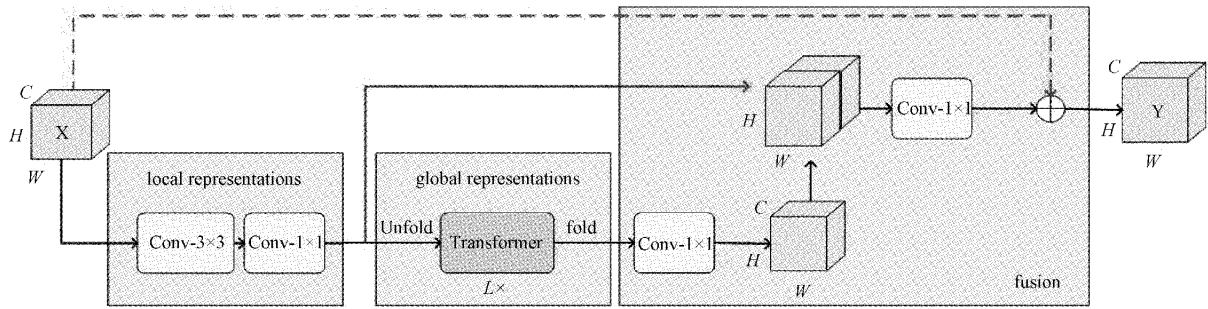
尽管 MobileViT 网络通过输入与全局进行特征融合,来提高图像分类的准确性,但由于番茄叶片病害种类很多,相似度很高和背景复杂等问题,直接将 MobileViT 网络应用到番茄叶片病害图像识别中仍存在一些不足。

在 MobileViT 模块中,直接将输入与全局进行特征融合,容易忽略特征图中的局部特征信息,不利于对番茄叶片病害图像进行准确的识别。由于局部特征信息和全局特征信息的关系要比输入特征信息和全局特征信息的关系更为密切,因此我们删除了原 MobileViT 模块原始输入特征与全局表征的 Concat 拼接,直接在全局特征信息提取之前,添加一条 shortcut 捷径分支,将得到的全局特征信息与局部特征信息沿通道方向进行 Concat 拼接,使网络能够提取更多关键的番茄叶片病害图像特征信息。如图 5 所示,为改进后的 MobileViT 模块,其中红色实线为添加的局部特征信息和全局特征信息特征融合。

考虑到局部特征层的输出通道略高于输入特征中的通道,因此在全局特征信息与局部特征信息拼接后,总通道数会大于改进之前的通道数,由于的 3×3 卷积层融合了输入特征、全局特征以及其他位置的输入和全局特征,这是一项复杂的任务,若直接进行卷积核大小为 3×3 的卷积层特征融合,会使网络在缩放时参数和 FLOPS 的大幅增加,因此我们使用 1×1 的卷积层,替换原网络的 3×3 卷积层,使融合模块的目标简化。

2.2 输入特征的残差融合

2016 年,He 等^[37]以卷积神经网络为基础,设计了残差神经网络(residual neural network, ResNet),它既能有效解决网络较深情况下引起的梯度消失及梯度爆炸等现象,又能将输入特征以捷径分支的方式整合起来,使得网



注:红色实线 → 代表concat拼接, 红色虚线 - → 代表残差连接

图 5 改进的 MobileViT 模块

络变得更深,从而能够实现网络深度较浅时,对网络加深,提高网络特征表达能力。在残差神经网络中引入残差学习概念,利用多个首尾连接的残差块来学习残差特征,残差块结构如图 6 所示。

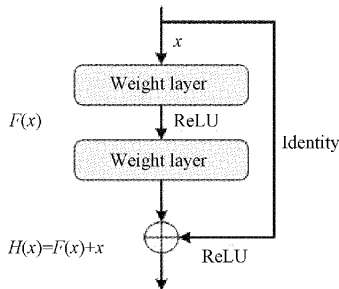


图 6 残差块结构

其中, x 为残差神经网络的输入, $H(x)$ 为输出, $F(x)$ 为残差映射函数, $H(x) = F(x, \{W_i\}) + x$ 为恒等映射函数; Weight layer 为卷积层。高分辨率图像一般深度较浅,大量信息保存在空间维度,能够很好地从空间维度上进行信息提取。所以,通过引入残差连接的方法,将其用于尺寸大、深度较浅的输入特征图和 MobileViT 模块最后一层融合块的输出特征中,如图 5 所示,红色虚线代表残差连接。这样有助于优化模型中的深层结构,提高模型精度。

2.3 替换 ReLU6 激活函数

MobileNetV2 模块中使用的是 ReLU6 激活函数,ReLU 激活函数的激活范围没有约束,它的输出范围是 0 到无穷大,而 ReLU6 函数作为一种特殊的 ReLU 函数,它的最大输出范围被约束为 6。当移动设备 Float16 精度较低时,它有很好的数值分辨率。但在激活函数输出值较高时,Float16 很难准确描述较大范围内的值,这也将造成精度下降。ReLU6 函数公式如式(2)所示。

$$ReLU6(x) = \min(\max(x, 0), 6) \quad (2)$$

如图 7 所示,为 ReLU6 与 H-Swish 函数对比图。Hard-Swish 函数是由 MobileNetV3^[38] 中提出的一种近似的激活函数,它的函数图像与原始的 Swish 函数类似,与 Swish 相比,移动部署设备中带 Hard-Swish 激活函数的模

型更具优越性。网络层数越多,图像的分辨率就越低,每激活一层内存将减少一半,非线性函数应用成本降低,在实践中具有重要意义,函数表现性能优越、求导方便且推理速度变化不大,所以为了进一步提升该模型性能,便采用 H-Swish 激活函数代替原 ReLU6 激活函数。H-Swish 公式如式(3)所示。

$$H-Swish(x) = x \cdot \frac{ReLU6(x+3)}{6} = \begin{cases} 0, & x \leq -3 \\ x, & x \geq +3 \\ x \cdot (x+3)/6, & \text{其他} \end{cases} \quad (3)$$

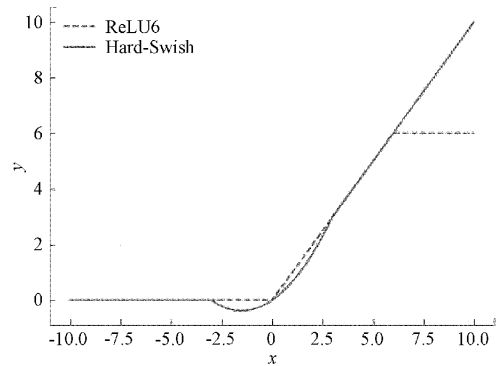


图 7 ReLU6 与 H-Swish 函数对比图

3 实验结果与分析

3.1 数据收集

番茄叶片病害图像分为单一背景和复杂背景,单一背景的番茄叶片病害图像在实际环境中识别较差,复杂背景番茄叶片病害图像背景复杂多样,对实验环境下的识别效果也不理想。因此本文将 Plant Village, Plant Doc 和 Dataset of Tomato Leaves 三个公开植物病害数据集组合成一个复合番茄叶片病害数据集,使得该数据集既含有简单背景环境,又含有复杂背景环境下的番茄叶片病害图像。

如图 8 所示,复合番茄叶片病害数据集包括 9 类番茄病害叶片图像(细菌斑点病,早疫病,晚疫病,叶霉病,斑枯

病,黄曲叶病毒病,花叶病毒病,灰叶斑病,双斑蜘蛛螨病)及健康叶片图像,共 10 类番茄叶片图像,每类病害分别有其对应的单一背景和复杂背景叶片图像。

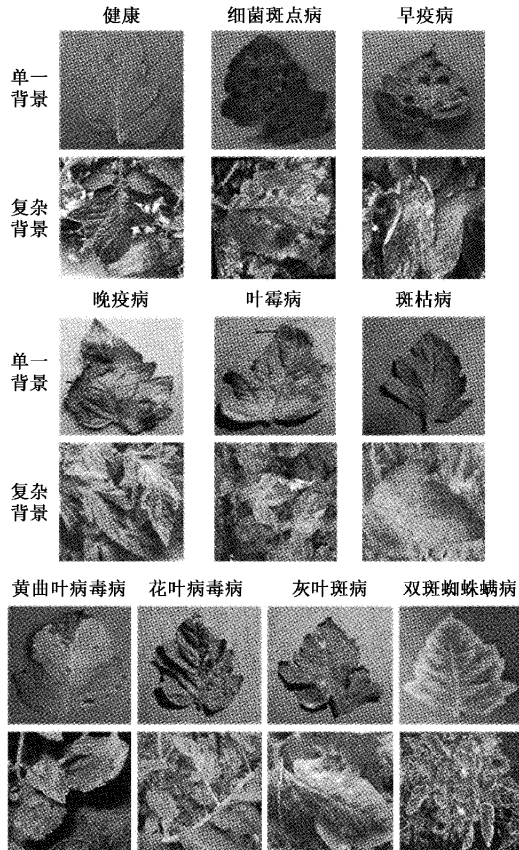


图 8 番茄病害样本图像

番茄叶片病害样本的详细情况如表 1,其中在图像增强之前数据集样本总量不多,不同番茄病害类别间样本数量相差悬殊,这种样本分布不平衡的状况将造成模型过拟合和难以特征提取等消极作用。

表 1 数据集详细信息

病害类型	单一背景/张	复杂背景/张
健康	1 607	689
细菌斑点病	1 824	1 216
早疫病	798	534
晚疫病	1 591	1 061
叶霉病	1 064	710
斑枯病	1 512	1 007
黄曲叶病毒病	1 775	762
花叶病毒病	2 188	547
灰叶斑病	1 827	457
双斑蜘蛛螨病	1 745	437
总计	15 931	7 420

在确保模型学习性能的前提下,构造了更多自然情况下病害识别的场景,使得模型能更适应于复杂情况下的工作环境,提高了其模型的鲁棒性。所以采用数据增强技术,通过图像平移、旋转、翻转、亮度或对比度的调节和添加噪声这 5 种数据增强技术分别完成了训练集与验证集病害图像的组组合增强。如图 9 所示,为增强效果示例图,对数据增强进行了适当的扩充调整后,数据可进行网络的训练。

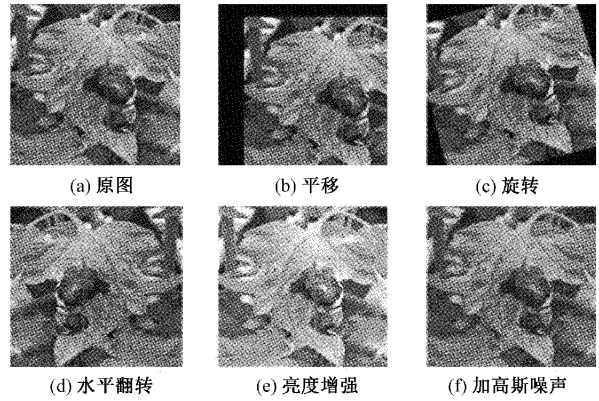


图 9 增强效果示例图

3.2 实验配置

本文实验是在 Ubuntu 21.04 64 位系统环境下完成的,CPU 为 Intel(R) Core(TM) i7-11200H CPU@2.40 GHz,GPU 为 NVIDIA RTX2080Ti,11 G 运行内存。模型框架采用的是 Pytorch1.10.1,编程语言的版本采用的是 Python3.8.10。

3.3 参数设置

实验过程中将数据集按照训练集:验证集:测试集 7:2:1 的比例进行随机划分,分别用作对模型进行训练、验证和测试。将模型全连接层改为 10 层,分别代表了番茄病害叶片的 9 种类型和番茄健康叶片的 1 种类型。番茄病害叶片的图像尺寸大小均设置为 224×224,批处理大小设置为 64,模型迭代 100 个 epoch。为了识别准确率和损失值能进行较好的收敛,因此实验采用了学习率衰减策略,将初始学习率设为 0.002,每迭代 10 个 epoch,则将学习率衰减 20%。优化器模块使用 Adam,参数默认。为了避免该模型出现过拟合的现象,将 Dropout 值设置为 0.2。

3.4 评价指标

为了客观评价改进模型的性能,以分类准确率(Accuracy)与模型占用内存大小为评价指标,分类平均准确率越高,并且模型占用内存大小越小,一定程度上说明该模型的性能越好。其中,分类准确率计算公式如下

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \quad (4)$$

式中:TP 表示数据集中的正样本被正确地划分为正样本的个数,TN 表示数据集中的负样本被正确地划分为负样本的个数,FP 表示数据集中的负样本被错误地划分为正

样本的个数, FN 表示数据集中的正样本被错误地划分为负样本的个数。

3.5 结果与分析

本文实验使用的是 MobileViT 模型中的 MobileViT-xs 模型,如图 10 所示,通过对原始 MobileViT-xs 网络模型和本文所提出的改进 MobileViT-xs 网络模型在实验数据集上 Accuracy 和 Loss 的结果对比表明,改进后模型收敛速度比原模型更快,收敛效果更加稳定,收敛时的准确率较高且损失值较小,证明了改进后的方法可靠性与有效性。

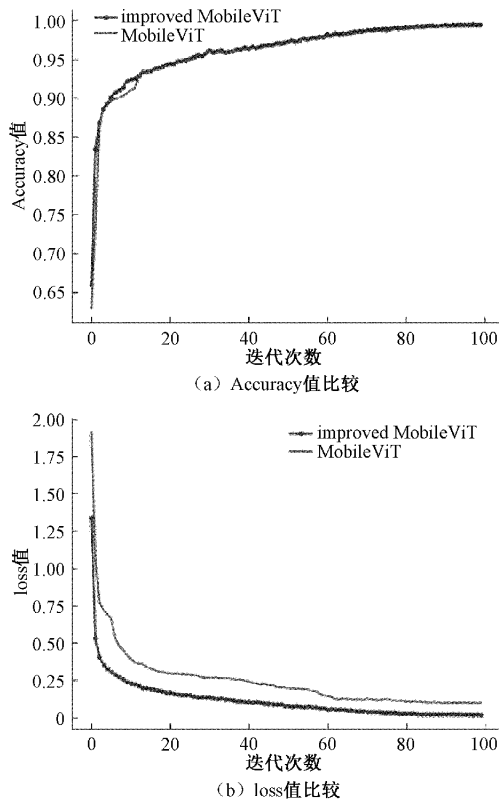


图 10 原始模型与所提改进模型比较

针对番茄叶片病害图像中存在的单一背景与复杂背景图像,使用原始 MobileViT 网络模型和改进后的网络模型分别进行了实验。在单一背景中,每张图像只有一个目标叶片;在复杂背景中,每张图像中有多个叶片,还包含复杂的背景干扰。实验结果如表 2 所示。

表 2 单一背景和复杂背景测试集准确率

实验	背景类型	平均准确率/%
MobileViT	单一背景	99.26
	复杂背景	95.61
本文方法	单一背景	99.52
	复杂背景	98.39

表 2 中的实验结果表明:在单一背景下,改进后的模

型相对于原始 MobileViT 模型方法的平均准确率只提高了 0.26%,提升效果不明显;在复杂背景下,改进后的模型相比原始模型方法的平均准确率提高了 2.78%,模型识别准确率大大提高。这表明改进后的模型具有更强的特征提取能力,在一定程度上可以去除复杂背景等干扰信息的影响,增强模型的鲁棒性。

为了进一步验证改进模型的效果,我们进行了以下实验来验证不同改进模块对该模型的效果。改进模型 1 是仅替换了 ReLU6 激活函数;改进模型 2 是在原模型上删除了输入和全局表示块的特征融合,将局部和全局表示块进行特征融合,并在融合块中使用 1x1 卷积层替换 3x3 卷积层;改进模型 3 是在改进模型 2 的基础上替换了 ReLU6 激活函数;最后是本文改进模型,在改进模型 3 的基础上添加了输入与融合块的残差结构。改进模型的消融实验结果如表 3 所示。

表 3 改进模型之间的性能对比

实验	平均准确率/%	FLOPs	Params/个	模型大小/M
原模型	98.10	约 1 G	约 2.3×10^6	7.53
改进模型 1	98.22	约 1 G	约 2.3×10^6	7.53
改进模型 2	98.75	约 1 G	约 2.3×10^6	7.65
改进模型 3	98.83	约 0.93 G	约 2.5×10^6	7.65
本文方法	99.16	约 0.93 G	约 2.5×10^6	7.65

从表 3 中可以看出,在替换了 ReLU6 激活函数之后,与原始模型相比模型大小不变,准确率提高了 0.12%;之后将局部和全局表示块进行特征融合,并使用 1x1 卷积层替换 3x3 卷积层,与原始模型相比,模型大小仅增加了 0.12 MB,但准确率提高了 0.65%;接着在此基础上替换了 ReLU6 激活函数,准确率又提高了 0.73%;当同时加入所有改进部分时,此时的模型比原模型准确率提高了 1.06%。

我们采用 Grad-CAM 热力图的分析法验证不同改进方法的有效性。不同改进方法可视化生成的热力图如图 11 所示,可以看出,每种模型对于某个类别的关注区域。图中原模型热力图中的热力区域主要集中在番茄叶片病害的细微特征上;改进模型 1 热力图中的热力区域主要集中在番茄叶片病害的一部分主要特征上;改进模型 2 热力图中的热力区域主要集中在番茄叶片病害的大部分主要特征和一些无关特征上;改进模型 3 热力图中的热力区域主要集中在番茄叶片病害的主要特征上,但是热力部分较为分散;本文方法热力图中的热力区域覆盖了番茄叶片病害的主要特征,并且比较集中。虽然采用不同改进方法都可以不同程度地增强模型对番茄叶片病害的几何特征提取能力,但是最后的本文方法可以更好地表征番茄叶片病害图像。

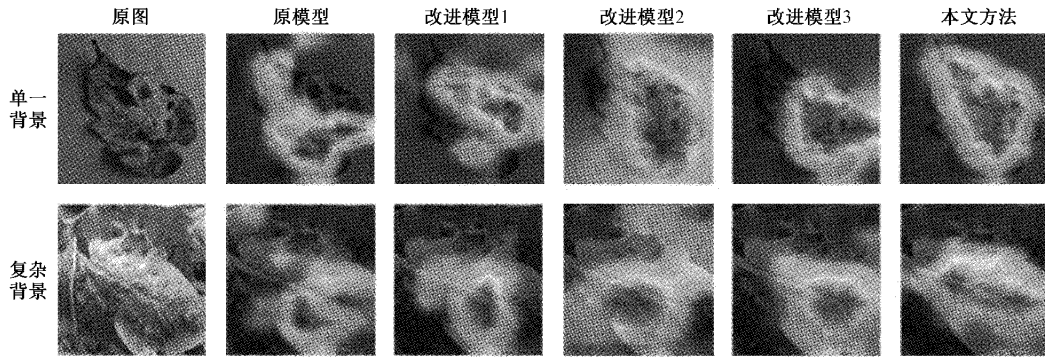


图 11 不同改进方法可视化生成的热力图

为了进一步验证改进后的 MobileViT 算法在分类番茄叶片病害的优势,将改进后的网络模型结构与其他算法进行对比,实验结果如表 4 所示。

表 4 不同模型之间的性能对比

实验	准确率/%	模型大小/MB
ResNet50	98.34	90.06
DenseNet	98.12	27.14
MobileNetV2	88.72	8.67
MobileNetV3	91.56	5.95
ShuffleNet	96.56	4.99
MobileViT	98.10	7.53
本文方法	99.16	7.65

可以发现,改进后的 MobileViT 识别准确率可以达到 99.16%,与 DenseNet^[39] 和 ResNet50 相比较,改进后的 MobileViT 在识别准确率上有小范围提高,但模型大小却远低于 DenseNet 和 ResNet50,这说明改进的 MobileViT 网络对番茄病害图像识别有一定的增强作用,同时具有较好的泛化能力,也极大地减少了该模型占用的内存。与 MobileNetV2, MobileNetV3 和 ShuffleNet 轻量级网络相比,在模型大小相差不多的情况下,改进后的 MobileViT 网络识别准确率有很大提升。因此改进后的 MobileViT 网络能够很好的完成番茄病害识别的任务。

为了更直观地展示分类结果,绘制混淆矩阵,如图 12 所示,0 表示细菌斑点病,1 为早疫病,2 为健康,3 为晚疫病,4 为叶霉病,5 为斑枯病,6 为黄曲叶病毒病,7 为花叶

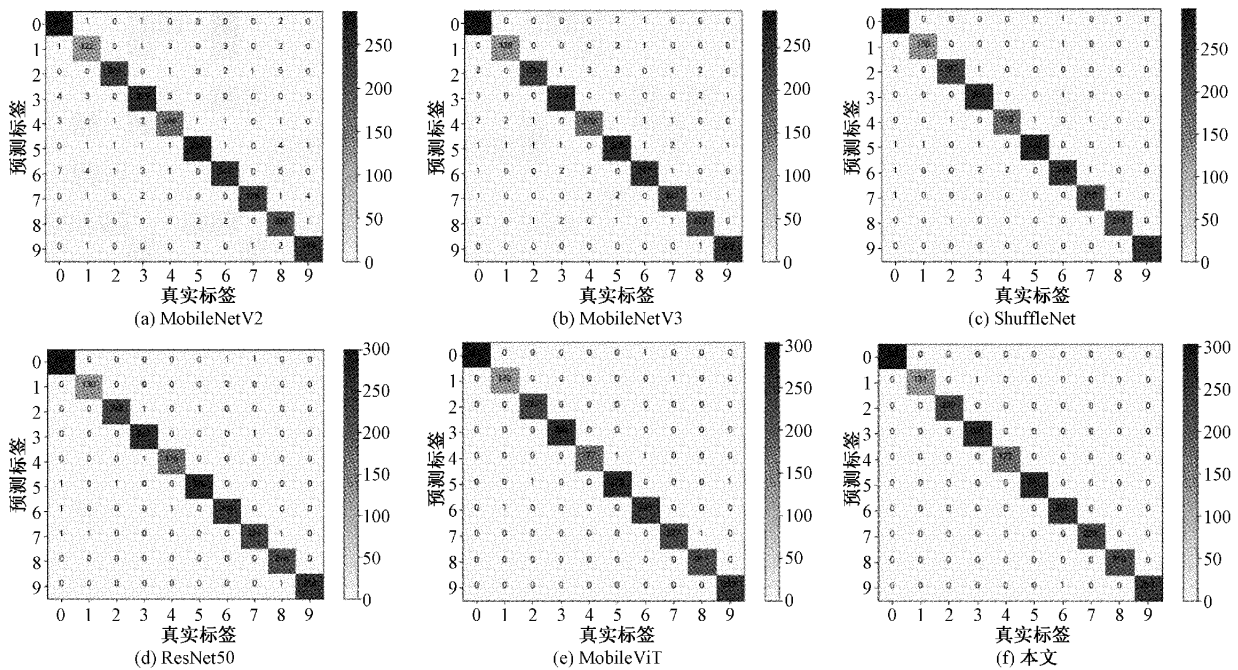


图 12 网络混淆矩阵图

病毒病,8 为灰叶斑病,9 为双斑蜘蛛螨病。对角线上的元素表示正确分类样本的数量。

由 MobileNetV2 和 MobileNetV3 网络混淆矩阵可知,网络对每一类病害的分类都没有完全正确;由 ShuffleNet

和 ResNet50 网络混淆矩阵可知,只有对于第 9 类分类正确,其他类别虽然相比与 MobileNetV2 和 MobileNetV3 网络有很大的提升,但是可以看出,还是有个别的分类错误;在 MobileViT 网络混淆矩阵上也得到了类似的结果;由本文方法的网络混淆矩阵中可以看出,仅第 3 和第 6 类样本中有分类错误,其他类别均分类正确。可见改进模型中对番茄叶片中病害类别划分有误的样本数量明显减少。这也表明了此改进方法是可行的,有助于番茄叶片病害的分类识别。

4 结 论

针对卷积神经网络对番茄叶片病害的分类效果不佳,且计算资源消耗多、占用内存大等问题,本文提出一种改进 MobileViT 算法,通过对其网络的 MobileViT 模块和损失函数进行改进,并在自制的番茄数据集上进行实验对比。结果表明,该模型在参数量和准确率之间都表现了较高水平的性能,且病害识别平均准确率达到 99.16%,相较原始 MobileViT 算法和其他卷积神经网络算法都有显著提高,可以更精确更快速的识别番茄叶片的病害类型,满足了在低性能设备上的轻量化网络快速推理识别病害的实际部署要求。本文模型识别准确率较高,整体性能较好,可为后续更智能化地识别其他植物叶片病害数据集研究提供参考。且在自然背景下的应用优势更为突出,后续工作将针对自然背景继续改进网络模型。

参考文献

- [1] DONGYU Q, ADHANOM G T, AZEVEDO R. Mitigating impacts of COVID-19 on food trade and markets[J]. Food and Agriculture organization of the United Nations, Accessed, 2020, 9.
- [2] TIAN H, WANG T, LIU Y, et al. Computer vision technology in agricultural automation—A review[J]. Information Processing in Agriculture, 2020, 7(1): 1-19.
- [3] VALENZUELA M E M, RESTOVIĆ F. Valorization of tomato waste for energy production[M]. London: Tomato Chemistry, Industrial Processing and Product Development, 2019: 245-258.
- [4] KAUR S, PANDEY S, GOEL S. Plants disease identification and classification through leaf images: A survey[J]. Archives of Computational Methods in Engineering, 2019, 26(2): 507-530.
- [5] LIU J, WANG X. Tomato diseases and pests detection based on improved Yolo V3 convolutional neural network[J]. Frontiers in Plant Science, 2020, 11: 898.
- [6] 陈晓, 荆茹韵. 单图像超分辨率方法综述[J]. 电子测量技术, 2022, 45(9):104-112.
- [7] 陈晓, 杨瑶. 基于长期递归卷积网络的无创血压测量[J]. 电子测量技术, 2022, 45(4): 139-146.
- [8] 孙超文, 陈晓. 基于多尺度特征融合反投影网络的图像超分辨率重建[J]. 自动化学报, 2021, 47(7): 1689-1700.
- [9] 杨瑶, 陈晓. 基于神经网络的说话人识别实验设计[J]. 实验室研究与探索, 2020, 39(9): 38-41.
- [10] 朱娟, 陈晓. 改进胶囊网络的有序重叠手写数字识别方法[J]. 激光杂志, 2019, 40(7): 43-46.
- [11] 郭妍, 陈晓, 任晓晔. 一种优化随机森林模型的室内定位方法[J]. 激光杂志, 2018, 39(10): 70-74.
- [12] 徐畅, 陈晓, 季仟亿. 基于稀疏编码的 Shearlet 域图像去噪[J]. 激光杂志, 2017, 38(10): 96-100.
- [13] 邹胜男, 陈晓, 陈霞. 基于改进的 UKF 算法的室内测距定位[J]. 激光杂志, 2017, 38(4): 61-65.
- [14] 陈霞, 陈晓, 邹胜男. 基于正态检验的室内定位算法[J]. 激光杂志, 2017, 38(3): 41-45.
- [15] 沈千里, 陈晓, 支亚京, 等. 一种新的人脸图像去噪算法[J]. 传感器与微系统, 2015, 34(11): 133-136.
- [16] 韩鹏飞, 陈晓. 基于 MFCC-IMFCC 和 GA-SVM 的鸟声识别研究[J]. 计算机系统应用, 2022, 31(11): 393-399.
- [17] NAZKI H, YOON S, FUENTES A, et al. Unsupervised image translation using adversarial networks for improved plant disease recognition[J]. Computers and Electronics in Agriculture, 2020, 168: 105117.
- [18] FERENTINOS K P. Deep learning models for plant disease detection and diagnosis[J]. Computers and Electronics in Agriculture, 2018, 145: 311-318.
- [19] AGARWAL M, GUPTA S K, BISWAS K K. Development of efficient CNN model for tomato crop disease identification [J]. Sustainable Computing: Informatics and Systems, 2020, 28: 100407.
- [20] CHEN X, ZHOU G, CHEN A, et al. Identification of tomato leaf diseases based on combination of ABCK-BWTR and B-ARNet [J]. Computers and Electronics in Agriculture, 2020, 178: 105730.
- [21] MEMON M S, KUMAR P, IQBAL R. Meta deep learn leaf disease identification model for cotton crop [J]. Computers, 2022, 11(7): 102.
- [22] FUENTES A, YOON S, KIM S C, et al. A robust deep-learning-based detector for real-time tomato plant diseases and pests recognition [J]. Sensors, 2017, 17(9): 2022.
- [23] 潘仁勇, 张欣, 陈孝玉龙, 等. 基于 DTS-ResNet 的苹果叶片病害识别方法[J]. 国外电子测量技术, 2022, 41(9):142-148.

- [24] 朱相华,智敏,殷雁君.基于 2D CNN 和 Transformer 的人体动作识别[J].电子测量技术,2022,45(15):123-129.
- [25] WANG Z, DUAN Z, ZHANG H, et al. Friendly topic assistant for transformer based abstractive summarization [C]. Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing(EMNLP), 2020: 485-497.
- [26] DOSOVITSKIY A, BEYER L, KOLESNIKOV A, et al. An image is worth 16x16 words; Transformers for image recognition at scale[J]. ArXiv Preprint, 2020, ArXiv:2010.11929.
- [27] QI X, CHEN G, LI Y, et al. Applying neural-network-based machine learning to additive manufacturing: Current applications, challenges, and future perspectives [J]. Engineering, 2019, 5(4): 721-729.
- [28] MEHTA S, RASTEGARI M. Mobilevit; Light-weight, general-purpose, and mobile-friendly vision transformer [J]. ArXiv Preprint, 2020, ArXiv: 2110.02178.
- [29] 赵苏徽,陈晓.基于树莓派和云平台的智能灌溉系统[J].计算机系统应用,2022,31(4):123-129.
- [30] 陈晓,张凯.嵌入式自动气象站故障检测系统研究[J].电子测量技术,2021,44(23):158-164.
- [31] 荣百川,陈晓.基于蓝牙的智能防摔监测系统设计[J].激光杂志,2019,40(7):32-34.
- [32] 邢玉品,陈晓.基于 TDC-GP22 的超声波测风仪设计[J].现代电子技术,2018,41(21):128-131.
- [33] 杨应鹏,陈晓.基于蓝牙的可穿戴式脉搏波形采集系统设计[J].激光杂志,2018,39(6):74-77.
- [34] 卢明乔,陈晓.基于 zigbee 组网的智能照明系统的研究与设计[J].信息技术,2013,12:56-59.
- [35] 卢明乔,陈晓.基于 ZigBee 传感网的农业智能照明系统[J].湖北农业科学,2014,53(23):5858-5860.
- [36] SANDLER M, HOWARD A, ZHU M, et al. Mobilenetv2; Inverted residuals and linear bottlenecks[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 4510-4520.
- [37] 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, 2016: 770-778.
- [38] HOWARD A, SANDLER M, CHU G, et al. Searching for mobilenetv3 [C]. Proceedings of the IEEE/CVF International Conference on Computer Vision, 2019: 1314-1324.
- [39] JÉGOU S, DROZDZAL M, VAZQUEZ D, et al. The one hundred layers tiramisu; Fully convolutional densenets for semantic segmentation[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2017: 11-19.

作者简介

陈晓(通信作者),博士,教授,主要研究方向为现代电子系统统计、信号与信息处理、图像处理、成像等。

E-mail: chenxiao@nuist.edu.cn

夏颖,硕士研究生,主要研究方向为计算机视觉,深度学习,图像处理等。

E-mail: 1269904282@qq.com