

史宝明, 贺元香, 赵霞. 一种轻量级 CNN 玉米病害图像识别方法[J]. 江苏农业科学, 2024, 52(5): 201–207.
doi:10.15889/j.issn.1002-1302.2024.05.030

一种轻量级 CNN 玉米病害图像识别方法

史宝明¹, 贺元香¹, 赵霞²

(1. 兰州文理学院数字媒体学院, 甘肃兰州 730010; 2. 甘肃农业大学信息科学技术学院, 甘肃兰州 730070)

摘要:针对传统卷积神经网络模型参数和运算量超大,难以部署在资源受限的移动终端或嵌入式设备上的问题,以 VGG16 作为基础框架,结合 MobileNet v3 模型思想,提出了一种轻量级卷积神经网络玉米病害图像识别方法,通过逐级渐进的方式建立网络模型,用线性瓶颈的倒残差深度可分离卷积代替标准卷积,用卷积层来代替全连接层,大幅度降低了模型的参数量和运算量。在深度卷积和点卷积之间加入改进的 squeeze and excitation 通道注意力模块,来增强模型精度。注意力模块的第 2 个全连接层的激活函数使用 hard-swish 代替 sigmoid,可以大幅度提高运算速度。试验样本数据为 PlantVillage 数据集的玉米病害子数据集,由于样本数据集偏小,通过随机旋转、随机缩放大小、随机宽度高度偏移、水平翻转、垂直翻转、随机错切变换、随机亮度变化、样本零均值化等方式对样本数据进行了增强和扩充,扩充后的数据集在改进模型上进行试验。试验结果表明,和 VGG16 对比,改进模型的准确率提高了 1.48 百分点,参数量是原模型的 1/5,运算量是原模型的 1/15;在不降低准确率的前提下,模型的参数量和运算量大幅度降低,实现了模型的轻量化。改进模型可以部署在移动终端等手持设备上,为农业病害识别提供指导和参考。

关键词:玉米病害;轻量级卷积神经网络;倒残差结构;深度可分离卷积;VGG;病害图像识别

中图分类号:TP391.41;S435.131 **文献标志码:**A **文章编号:**1002-1302(2024)05-0201-07

随着深度学习技术的快速发展,相继诞生了一大批优秀的卷积神经网络,如 VGG^[1]、ReNet^[2]、InceptionNet^[3]、DenseNet^[4]等,模型的准确率不断提升,但同时网络模型的深度越来越深,参数量和运算量越来越大,对硬件设备的算力要求也越来越高。这些网络模型适合在 PC 端进行部署应用,但在移动设备和嵌入式设备中却不适合部署应用。国内学者通过将上述经典卷积神经网络引入农业病害识别中^[5-11],取得了不错的效果,但实用性不高。对于农业病害识别来说,开发能够部署在诸如手机移动设备或嵌入式设备的轻量级网络模型^[12]

具有更好的市场前景。

轻量化卷积神经网络的设计主要从卷积结构轻量化、卷积模块轻量化和卷积运算轻量化 3 个方面来进行^[13]。在轻量化卷积网络研究和应用方面,2016 年最早公开的轻量化网络 SqueezeNet^[14],在 ImageNet 数据集上的表现和 AlexNet 相当,但其参数量仅为 AlexNet 的 2%。后来随着 Xception^[15]、ShuffleNet^[16-18]等网络的提出,轻量级神经网络得到了快速发展,并在农业病害防治方面进行了应用和改进^[19-21]。2017 年 Google 研究团队首次提出 MobileNet^[22],使用深度可分离卷积技术,大大降低了模型的参数和运算量,在 MobileNet v1 的基础上,对网络模型的主体模块引入了线性瓶颈的倒残差结构,提出了 MobileNet v2^[23]。相比之前的模型,它的准确率更高,模型更小。MobileNet v3^[24]更新了层块的设计,加入了通道注意力机制,使用 H-

收稿日期:2023-04-20

基金项目:甘肃省教育厅创新基金(编号:2023A-181)。

作者简介:史宝明(1981—),男,甘肃平凉人,硕士,副教授,主要研究方向为深度学习、计算机视觉、虚拟现实技术。E-mail:zlsbm@qq.com。

[28] Tan M X, Pang R M, Le Q V. EfficientDet: scalable and efficient object detection [C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). June 13–19, 2020, Seattle, WA, USA. IEEE, 2020:10778–10787.

[29] Li H L, Li J, Wei H B, et al. Slim-neck by GSConv: a better design paradigm of detector architectures for autonomous vehicles [EB/OL]. 2022; arXiv: 2206.02424. [https://arxiv.org/abs/](https://arxiv.org/abs/2206.02424)

2206.02424. pdf.

[30] Fenu G, Mallocci F M. DiaMOS plant: a dataset for diagnosis and monitoring plant disease[J]. Agronomy, 2021, 11(11):2107.

[31] Peng S D, Jiang W, Pi H J, et al. Deep snake for real-time instance segmentation [C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). June 13–19, 2020, Seattle, WA, USA. IEEE, 2020:8530–8539.

swish 激活函数代替 ReLU6 激活函数,使用神经网络结构搜索(neural architecture search, NAS)^[25] 技术搜索网络参数,并重新设计优化了耗时层,模型的准确率和速度有了进一步的提升。国内学者们针对 MobileNet 网络进行了一系列的改进。刘洋等对 MobileNet 和 Inception v3 网络进行迁移学习,实现了 PlantVillage 数据集和自建葡萄叶片病害数据集的识别,并对 2 种模型的识别效果进行了对比分析研究^[26]。孙俊等基于 MobileNet v2,引入轻量级的坐标注意力机制和多尺度特征融合技术对 MobileNet v2 模型进行改进,在 11 类病害叶片和 4 类健康叶片数据集上进行了识别验证,模型精度提升了 2.91%^[27]。贾鹤鸣等基于 VGG 网络,引入深度可分离卷积和全局平均池化技术,构建了一个轻量级网络模型并在 PlantVillage 数据集上进行验证,也取得了较好效果^[28]。徐振南等提出了一种基于轻量级卷积神经网络 MobileNet v3 的网络模型,通过迁移学习方式对马铃薯叶部病害进行了识别,识别准确率达到 98.00%^[29]。上述研究表明,轻量级卷积神经网络模型在减少网络参数量的同时,可以保证较高的识别准确率,能够应用在移动端等终端设备上,可以满足农业病害识别的实际需求。这些研究为农作物病害识别在资源受限的移动设备和

嵌入式设备上部署提供了参考。

本研究以 VGG16 作为基础框架,使用了深度可分离卷积思想提出了一种 CNN 玉米病害图像识别方法,用线性瓶颈的倒残差深度可分离卷积代替标准卷积实现模型的轻量化,引入通道注意力模块来增强模型精度,同时通过改进注意力模块中的激活函数来提升运算速度。

1 研究理论与方法

1.1 VGG 网络

VGG 网络^[1]获得了 2014 年 ImageNet 图像分类比赛的亚军。VGG 网络主要由 5 层卷积层、3 层全连接层和 softmax 输出层组成(图 1)。由于堆叠 2 个 3×3 卷积核的感受野和 1 个 5×5 卷积核的感受野相同,堆叠 3 个 3×3 卷积核的感受野和 1 个 7×7 卷积核的感受野相同,因此 VGG 网络使用多个 3×3 卷积核来代替更大的卷积核,能够有效降低网络的参数量,所有隐藏层之间的激活单元使用 ReLU 激活函数,在进行下采样时,统一使用了 2×2 小池化核。相比之前的网络模型,它的网络深度更深,表示的特征图更宽。VGG 网络结构简单,易于扩展改进。

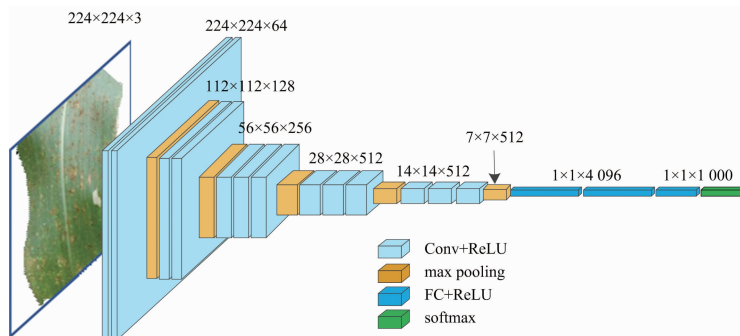


图1 VGG 网络模型

1.2 深度可分离卷积

经研究发现,使用深度可分离卷积代替标准卷积可以大大减少模型的参数量^[22]。深度可分离卷积(depthwise separable convolution, DSC)由深度卷积(depthwise convolution, DWConv)和逐点卷积(pointwise convolution, PWConv)构成。标准卷积和深度可分离卷积的对比如图 2 所示,其中,BN 表示批量归一化层,ReLU 表示激活函数层,DWConv 表示深度卷积层,PWConv 表示逐点卷积层。

在标准卷积中,卷积核的深度和输入特征图的

通道数相同,输出特征图的通道数和卷积核的个数相同。在 DWConv 中,卷积核的个数和输入特征图的通道数相同,卷积核的深度均为 1,每个卷积核只负责与输入特征图的单个通道进行卷积运算,因此输出特征图的通道数与输入特征图的通道数相同。而在 PWConv 中,使用 DWConv 卷积的输出特征图在深度方向进行加权混合,生成新的特征图,卷积核的大小为 1,其深度和 DWConv 的输出通道数相同。卷积核的个数决定了输出特征图的通道数,PWConv 卷积本质上就是进行卷积核为 1×1 的标准卷积。

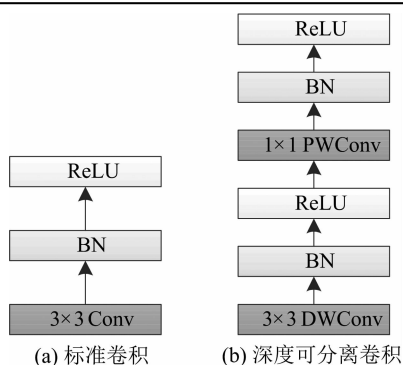


图2 普通卷积和深度可分离卷积

假设输入特征图的大小为 $D_F \times D_F$, 其通道数为 M , 卷积核的大小为 $D_k \times D_k$, 输出特征图的通道数为 N , 则标准卷积的参数量 P_{Conv} 的大小为:

$$P_{\text{Conv}} = D_F \times D_F \times M \times N \times D_k \times D_k$$

深度可分离卷积的参数量 P_{DSC} 的大小为:

$$P_{\text{DSC}} = D_F \times D_F \times M \times D_k \times D_k + D_F \times D_F \times M \times N$$

两者之比为:

$$\frac{P_{\text{DSC}}}{P_{\text{Conv}}} = \frac{D_F \times D_F \times M \times D_k \times D_k + D_F \times D_F \times M \times N}{D_F \times D_F \times M \times N \times D_k \times D_k} = \frac{1}{N} + \frac{1}{D_k^2}$$

当选择 $D_k = 3$ 时, $P_{\text{DSC}}/P_{\text{Conv}} = 1/N + 1/9$, 可以看到深度可分离卷积的参数量约是标准卷积的 $1/9$ 。因此, 使用深度可分离卷积代替标准卷积, 可以设计轻量化的卷积神经网络。

1.3 深度可分离倒残差结构

深度可分离残差结构^[23]是先使用 1×1 的卷积降维, 接着进行 3×3 的卷积, 再用 1×1 的卷积升维, 是一个两头大、中间小的细腰结构, 而深度可分离倒残差结构是先进行 1×1 卷积扩维, 接着进行 3×3 卷积, 再用 1×1 的卷积降维, 是一个两头小、中间大的瓶颈结构, 前 2 个卷积使用非线性激活函数, 最后 1 个卷积使用线性激活函数, 因此也称为线性瓶颈倒残差结构, 以下称其为 bneck 块。本研究提出 2 种倒残差结构, 区别是一个深度卷积的步长值为 1, 需要进行残差跳连; 另一个深度卷积的步长值为 2, 无需进行残差跳连。两者的结构如图 3 所示。其中, Stride 为步长值, ReLU 为非线性激活函数, Linear 为线性激活函数。从图 3 中可知, 2 种倒残差结构分别对应了 Stride 为 1 和 2 的 2 种不同形态, 2 种形态主要都由 1×1 升维卷积层、DWConv 卷积层、 1×1 降维卷积层构成。前 2 层使用了 ReLU 激活函数, 最后 1 层使用 Linear 线性激活函数, 即无需做任何操作。当 Stride = 1 且输入特征通道数等于输出特征通道数时, 需要残差跳连, 其余情况无需

残差跳连。要注意的是 2 种形态中的 1×1 卷积步长均为 1, 只有 DWConv 卷积步长值不同。

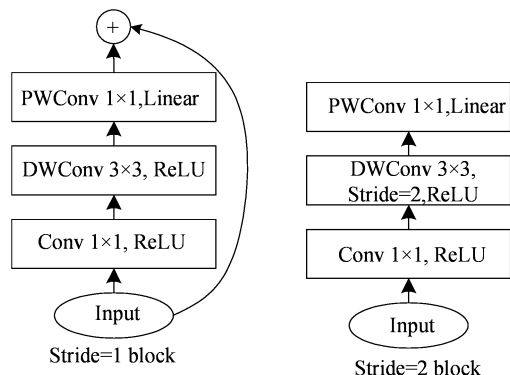


图3 Bneck 块

1.4 激活函数

目前最常用的激活函数是 ReLU 激活函数, 其定义如下:

$$\text{ReLU}(x) = \max(0, x)$$

在文献[24]中给出了一种新的激活函数, 即 swish 激活函数, 可有效代替 sigmoid 函数, 其定义如下:

$$\text{swish}(x) = x \cdot \sigma(x);$$

$$\sigma(x) = 1/(1 + e^x)$$

式中: $\sigma(x)$ 为 sigmoid 函数, 由于 $\sigma(x)$ 需要进行指数运算, 效率较低, 而 $\sigma(x)$ 和 ReLU 的图像比较相近, 因此借助 ReLU 定义了 h-sigmoid(x) 函数来近似模拟 $\sigma(x)$, 即有:

$$\text{h-sigmoid}(x) = \text{ReLU6}(x+3)/6$$

式中: $\text{ReLU6} = \min[\text{ReLU}(x), 6]$ 。在 swish 函数中, 用 h-sigmoid(x) 代替 $\sigma(x)$, 可得到 h-swish 函数, 即有:

$$\text{h-swish}(x) = x \cdot \text{ReLU6}(x+3)/6$$

相关激活函数的对比如图 4 所示, 可以看到, sigmoid 函数与 h-sigmoid 函数、swish 函数与 h-swish 函数这 2 组函数的图像非常近似, 但由于 ReLU 函数的运算规则简单, 没有复杂的指数运算, 所以基于 ReLU 的 h-sigmoid、h-swish 函数的运算效率要高的多。

1.5 通道注意力机制

通道注意力机制(channel attention mechanism, CAM)^[30]能够有效提升卷积神经网络的性能。在上述 bneck 块的 DWConv 层和 PWConv 层之间, 引入了 squeeze and excitation(SE)通道注意力模块, 即先对每个 DWConv 输出的特征图进行全局平均池化, 再进行全连接、ReLU6 激活, 接着再进行全连接、

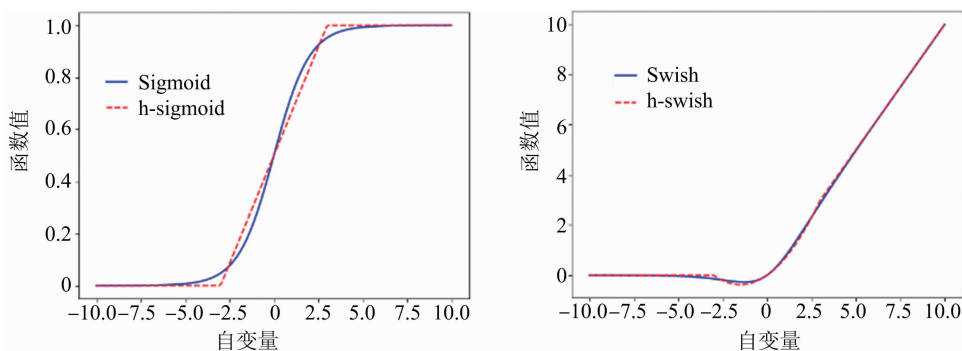


图4 激活函数对比

sigmoid 或 h-swish 激活,第 1 个全连接层的节点个数设为 DWConv 输出特征图的 1/4,第 2 个全连接层的节点个数和 DWConv 输出特征图的个数相同,经过上述运算就得到了 1 组特征图权重系数,根据权重系数对特征图进行融合作为 PWConv 的输入,

这样可以增加有用特征的权重,从而将注意力聚焦在更加有用的特征上,在深度可分离卷积中引入通道注意力机制后的模块命名为 SE-bneck 块,其结构见图 5。

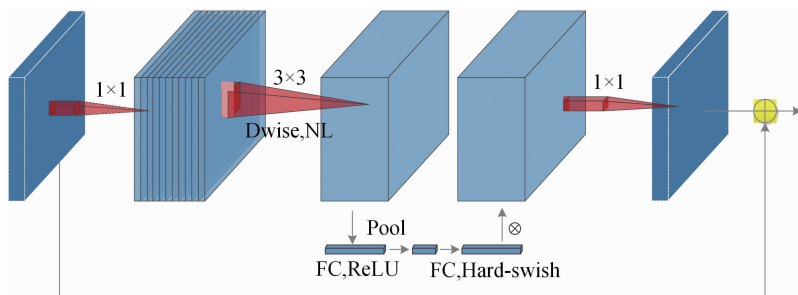


图5 SE-bneck 块

在通道注意力模块中,第 1 个全连接层后采用了 ReLU6 激活函数。当 ReLU6 的输入过大时,其输出被限制为 6,这样可有效避免因权重梯度相差过大引起的量化误差,提高模型的表达能力。而第 2 个全连接层的后面采用了 h-swish 激活函数。

2 网络结构设计

本研究以 VGG16 作为基础模型来进行改进,VGG16 主要由 13 层卷积层和 3 层全连接层构成,在特定的卷积层间加入 MaxPooling 层进行下采样。由于 VGG 模型是在 1 000 个分类的 ImageNet 网络上进行训练的,而本研究中的玉米病害类型只有 4 类,因此本研究将 VGG 模型全连接层前 2 层的神经元个数由 4 096 替换为 512,最后 1 层的分类数由 1 000 变为 4,将该模型记为模型 0。以下通过逐级递进的方式设计完成了一个轻量级 CNN 玉米病害图像识别模型。

在模型 0 的基础上,在每个卷积层和激活函数之间加入批量归一化(batch normalization, BN)层。

BN 层是将输入数据按照批量样本特征的均值 u 和方差 σ 对样本特征进行标准化处理,即

$$x_i' = \frac{x_i - u}{\sqrt{\sigma^2 + \varepsilon}}$$

其中, x_i' 为特征图像中第 i 个像素 x_i 归一化后的值, ε 为微小常数,保证分母大于 0。BN 操作可以将卷积后原本偏离的数据重新拉回到均值为 0、标准差为 1 的正态分布区域,从而将进入激活函数的数据分布在激活函数的线型区,使得输入数据的微小变化能够更明显地体现到激活函数的输出,进而增强模型的表达能力。经过对改进的 VGG16 网络分析发现,其主要参数量来源于后面的全连接层,即使全连接层的神经元降为原来的 1/4,全连接层的参数量仍占整个网络参数量的 47% 左右,因此考虑用卷积层来代替全连接层,舍弃 flatten 层,用 1 个卷积核为 $1 \times 1 \times 512$ 的卷积层代替第 1 个全连接层,后接全局平均池化层,再用 1 个卷积核为 $1 \times 1 \times 512$ 的卷积层代替第 2 个全连接层,用卷积核为 $1 \times 1 \times \text{classes}$ 的卷积层代替第 3 个全连接层,后接

flatten 层和 softmax 层,改进模型记为模型 1。

在模型 1 的基础上,引入深度可分离卷积和线性瓶颈的倒残差结构。前面的 13 个卷积层除了第 1 个之外,其余的 12 个卷积层用 bneck 块来代替,得到了模型 2。注意只有在下采样时,bneck 块的步长为 2,其余均为 1。模型 2 相比模型 1,在保持模型精度基本不变的前提下,进一步降低了模型的参数量,提高了运算速度。在模型 2 的基础上,通过在部分 bneck 块中引入 SE 模块,使得在进行 PW 卷积时能够关注更加有用的特征。具体做法是在特征提取层使用 SE 模块。而在下采样过程中不使用 SE 模块,另外相比 ReLU 激活函数,h-swish 激活函数能够更加有效地进行非线性变换,因此将网络第 1 个卷积层和部分深层的激活函数替换为 h-swish,同时模型的第 1 个卷积层的步长由 1 调整为 2,后接的第 1 个 bneck 层的步长由 2 调整为 1,加速了模型的轻量化,改进后的模型记为模型 3。各模型的结构对比见表 1,模型 3 的详细配置情况见表 2。

表 1 模型结构对比

模型 0	模型 1	模型 2	模型 3
输入图像(224×224×3)			
conv3-64	conv-bn3-64	conv-bn3-16	conv-bn3-16s2
conv3-64	conv-bn3-64	bneck64-32s2	bneck64-32
最大池化	最大池化		
conv3-128	conv-bn3-128	bneck128-32s1	se-bneck128-32s1
conv3-128	conv-bn3-128	bneck128-64s2	bneck128-64s2
最大池化	最大池化		
conv3-256	conv-bn3-256	bneck256-64s1	se-bneck256-64s1
conv3-256	conv-bn3-256	bneck256-64s1	se-bneck256-64s1
conv3-256	conv-bn3-256	bneck256-128s2	bneck256-128s2
最大池化	最大池化		
conv3-512	conv-bn3-512	bneck512-128s1	se-bneck512-128s1
conv3-512	conv-bn3-512	bneck512-128s1	se-bneck512-128s1
conv3-512	conv-bn3-512	bneck512-256s2	bneck512-256s2
最大池化	最大池化		
conv3-512	conv-bn3-512	bneck512-256s1	se-bneck512-256s1
conv3-512	conv-bn3-512	bneck512-256s1	se-bneck512-256s1
conv3-512	conv-bn3-512	bneck512-512s2	bneck512-512s2
最大池化	最大池化		
flatten	conv-bn1-512	conv-bn1-512	conv-bn1-512
dense-512	global_avg_pool	global_avg_pool	global_avg_pool
dense-512	conv-bn1-512	conv-bn1-512	conv-bn1-512
dense-classes	conv1-classes	conv1-classes	conv1-classes
	flatten	flatten	flatten
softmax	softmax	softmax	softmax

注:conv 后的第 1 个数值表示卷积核的大小,第 2 个数值表示输出通道数,bn 表示使用了 BN 层。bneck 块后面的第 1 个数值表示扩张通道数,第 2 个数值表示输出通道数,s 后的数值表示深度卷积的步长,se 表示使用了 SE 通道注意力模块。

表 2 模型 3 网络模型

输入	运算	扩维 通道数	输出 通道数	SE	激活 函数	步长
224 ² ×3	conv2d,3×3		16	No	HS	2
112 ² ×16	bneck,3×3	64	32	No	RE	1
112 ² ×32	bneck,3×3	128	32	Yes	RE	1
112 ² ×32	bneck,3×3	128	64	No	RE	2
56 ² ×64	bneck,3×3	256	64	Yes	RE	1
56 ² ×64	bneck,3×3	256	64	Yes	RE	1
56 ² ×64	bneck,3×3	256	128	No	RE	2
28 ² ×128	bneck,3×3	512	128	Yes	RE	1
28 ² ×128	bneck,3×3	512	128	Yes	HS	1
28 ² ×128	bneck,3×3	512	256	No	HS	2
14 ² ×256	bneck,3×3	512	256	Yes	HS	1
14 ² ×256	bneck,3×3	512	256	Yes	HS	1
14 ² ×256	bneck,3×3	512	512	No	HS	2
7 ² ×512	conv2d,3×3		512	Yes	RE	1
7 ² ×512	global_avg_pool		512	No		1
1 ² ×512	conv2d,1×1		512	No	RE	1
1 ² ×512	conv2d,1×1		classes	No		1

注:SE 表示是否采用注意力机制;激活函数 RE 表示 ReLU 激活函数,HS 表示 h-swish 激活函数;classes 表示模型分类数。

3 数据集预处理

本研究使用的数据集图像来自公开数据集 PlantVillage 中的玉米叶片病害图像,共包含了 4 类图像共计 4 354 张图片,其中健康玉米叶片图像 1 162 张,锈病叶片图像 1 192 张,大斑病叶片图像 1 000 张,灰斑病叶片图像 1 000 张。对应的 4 类数据标签分别为:corn_healthy、corn_commonrust、corn_northernblight、corn_cercospora灰斑病。将上述图像先按 8:2 的比例划分为训练集和测试集。由于数据集图片量偏少,模型在训练的过程中容易过拟合,为了提升模型的泛化能力,降低模型对图像的敏感度,可以通过数据增强的方式对数据集进行扩充,增加样本数量和多样性。创建 Python 数据增强脚本,使用 keras 的 ImageDataGenerator 对象对样本数据进行随机旋转 90°、随机缩放大小 0.2、随机宽度高度偏移 0.3、水平翻转、垂直翻转、随机错切变换 0.2、亮度变化范围[0.1,0.9]、样本零均值化等处理对训练集数据进行增强,增强效果见图 6。

增强后的数据集包含健康玉米叶片图像 2 982 张、锈病叶片图像 3 022 张、大斑病叶片图像和灰斑病叶片图像各 3 040 张。将增强后的数据集按 8:2

的比例再次划分为训练集和验证集。划分后的数据分布见表 3。

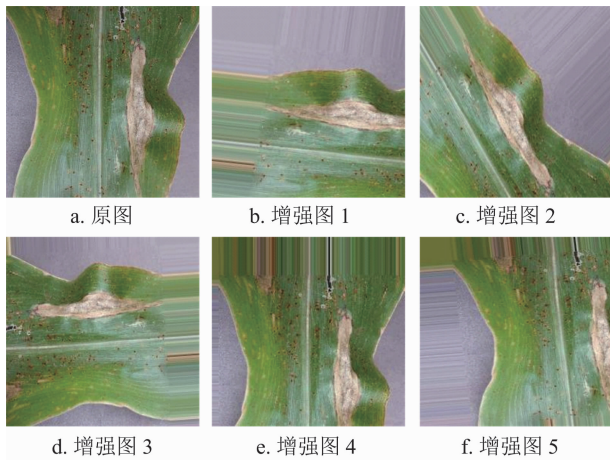


图6 数据增强效果

表 3 增强玉米叶片病害数据分布

病害名称	训练集数量 (张)	验证集数量 (张)	测试集数量 (张)	图片 格式
健康叶片	2 386	596	232	JPEG
锈病	2 418	604	238	JPEG
大斑病	2 432	608	200	JPEG
灰斑病	2 432	608	200	JPEG

4 试验设计

4.1 试验环境

试验在 Windows 10 系统环境下完成,具体软硬件配置如下: Intel® Core™ i5 – 6500 CPU@3.20 GHz 处理器, NVIDIA GeForce RTX3060 显卡, 16 G 内存; 软件采用 Anaconda + PyCharm 集成开发环境, Python 版本为 3.8; TensorFlow 深度学习框架, 版本为 2.5; CUDA 版本为 CUDA 11.1。

4.2 试验细节

试验中输入图像的大小统一裁剪为 224 × 224, 并进行了归一化处理。batchsize 设为 32, 模型选用 Adam 优化器和交叉熵损失函数, 训练迭代轮次设为 30 个 epoch, 学习率采用衰减学习率, 初始学习率为 0.001, 每个迭代 10 个 epoch。学习率降为原来的 1/10。

4.3 试验结果分析

对模型 0 至模型 3 在增强数据集上依次进行训练, 记录各模型在训练过程中的识别准确率和损失函数, 准确率和损失值变化曲线结果分别见图 7、图 8。

对模型 0 至模型 3 的准确率、参数量和运算量 (FLOPs) 这几个指标进行统计, 结果见表 4。

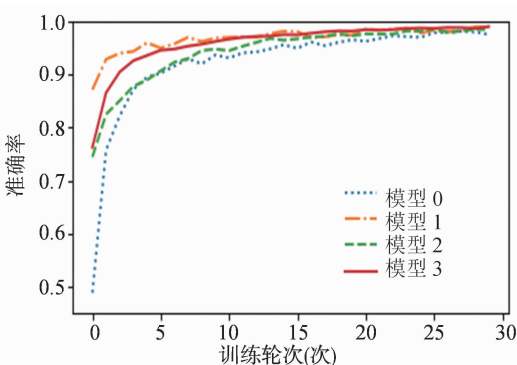


图7 准确率变化曲线

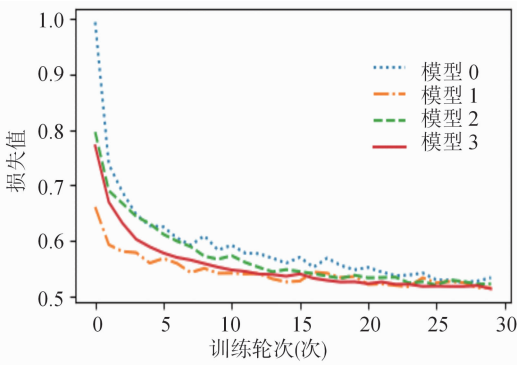


图8 损失函数变化曲线

表 4 4 个模型评价指标对比

模型	准确率 (%)	参数量 (M)	FLOPs (G)
模型 0	97.56	27.82	30.74
模型 1	99.01	15.26	30.75
模型 2	98.51	4.22	2.19
模型 3	99.04	4.82	2.17

通过对图 7、表 4 以及试验过程中记录的其他相关数据分析可知, 经过 30 个 epoch 的迭代, 4 个模型都达到了较高的准确率。模型 1 在模型 0 基础上, 准确率提高了 1.45 个百分点, 参数量减少了将近 50%, 但运算量并未减少。深入分析后发现模型 1 相比模型 0 在全连接层的运算量有所减少, 但由于在各卷积层后面增加了 BN 层, 这些 BN 层的运算量累加在一起也是非常可观的。模型 2 使用了深度可分离卷积层代替标准卷积层, 在准确率略微降低的情况下, 使得模型的参数量降为模型 1 的 1/4 左右, 运算量约为模型 1 的 1/15, 实现了模型的轻量化。在模型 2 的基础上, 进一步引入通道注意力机制 SE 模块, 并且调整了第 1 个卷积层和 bneck 层的步长后, 在参数量略微增加的情况下, 进一步降低了运算量, 且提高了模型的准确率。改进后的模型 3 相比模型 0, 准确率提高了 1.48 个百分点, 而模型的参

数量减少为原来的 1/5 左右,浮点运算量降为原来的 1/15 左右,在没有降低准确率的前提下,实现了模型的轻量化。

5 结语

本研究提出了一种轻量级 CNN 玉米病害图像识别方法,在 VGG16 网络架构的基础上,引入深度可分离卷积代替标准卷积,以减少模型参数和运算量;在深度卷积层和点卷积之间,引入 SE 通道注意力模块,使得在确保准确率不降低的情况下,模型参数和运算量大幅度降低,实现了模型的轻量化。该模型可部署在资源受限的移动端设备和嵌入式设备上,方便在实际农业生产中进行推广和应用。

参考文献:

- [1] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition [J]. ArXiv e-Prints, 2014; arXiv: 1409.1556.
- [2] He K M, Zhang X Y, Ren S Q, et al. Deep residual learning for image recognition [C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, NV, USA. IEEE, 2016: 770-778.
- [3] Szegedy C, Vanhoucke V, Ioffe S, et al. Rethinking the inception architecture for computer vision [C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, NV, USA. IEEE, 2016: 2818-2826.
- [4] Huang G, Liu Z, van der Maaten L, et al. Densely connected convolutional networks [C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, HI, USA. IEEE, 2017: 2261-2269.
- [5] 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.
- [6] 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.
- [7] 方晨晨, 石繁槐. 基于改进深度残差网络的番茄病害图像识别[J]. 计算机应用, 2020, 40(增刊 1): 203-208.
- [8] 刘翱宇, 吴云志, 朱小宁, 等. 基于深度残差网络的玉米病害识别[J]. 江苏农业学报, 2021, 37(1): 67-74.
- [9] 鲍文霞, 吴刚, 胡根生, 等. 基于改进卷积神经网络的苹果叶部病害识别[J]. 安徽大学学报(自然科学版), 2021, 45(1): 53-59.
- [10] 樊湘鹏, 周建平, 许燕, 等. 基于改进卷积神经网络的复杂背景下玉米病害识别[J]. 农业机械学报, 2021, 52(3): 210-217.
- [11] 何欣, 李书琴, 刘斌. 基于多尺度残差神经网络的葡萄叶片病害识别[J]. 计算机工程, 2021, 47(5): 285-291, 300.
- [12] 葛道辉, 李洪升, 张亮, 等. 轻量级神经网络架构综述[J]. 软件学报, 2020, 31(9): 2627-2653.
- [13] 马金林, 张裕, 马自萍, 等. 轻量化神经网络卷积设计研究进展[J]. 计算机科学与探索, 2022, 16(3): 512-528.
- [14] Iandola F N, Han S, Moskewicz M W, et al. SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5 MB model size [EB/OL]. (2016-06-01) [2023-04-12]. <http://ariv.org/abs/1602.07360>.
- [15] Chollet F. Xception: deep learning with depthwise separable convolutions [C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, HI, USA. IEEE, 2017: 1800-1807.
- [16] Zhang X Y, Zhou X Y, Lin M X, et al. ShuffleNet: an extremely efficient convolutional neural network for mobile devices [C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, UT, USA. IEEE, 2018: 6848-6856.
- [17] Ma N N, Zhang X Y, Zheng H T, et al. ShuffleNet v2: practical guidelines for efficient CNN architecture design [C]//Computer Vision - ECCV 2018: 15th European Conference, Munich, Germany, September 8-14, 2018, Proceedings, Part XIV. ACM, 2018: 122-138.
- [18] 杨祥, 段军明, 董明刚. 面向移动端的植物病害图像识别方法及其应用[J]. 江苏农业科学, 2023, 51(4): 191-197.
- [19] 项小东, 翟蔚, 黄言志, 等. 基于 Xception-CEMs 神经网络的植物病害识别[J]. 中国农机化学报, 2021, 42(8): 177-186.
- [20] 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.
- [21] 李书琴, 陈聪, 朱彤, 等. 基于轻量级残差网络的植物叶片病害识别[J]. 农业机械学报, 2022, 53(3): 243-250.
- [22] Howard A G, Zhu M, Chen B, et al. MobileNets: efficient convolutional neural networks for mobile vision applications [EB/OL]. (2017-06-06) [2023-04-12]. <https://ariv.org/abs/1704.04861>.
- [23] Sandler M, Howard A, Zhu M L, et al. MobileNet v2: inverted residuals and linear bottlenecks [C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, UT, USA. IEEE, 2018: 4510-4520.
- [24] Howard A, Sandler M, Chen B, et al. Searching for MobileNet v3 [C]//2019 IEEE/CVF International Conference on Computer Vision (ICCV). Seoul, Korea (South). IEEE, 2019: 1314-1324.
- [25] Lv Z Q, Qian C, Yen G G, et al. Analysis of expected hitting time for designing evolutionary neural architecture search algorithms [EB/OL]. (2022-10-11) [2023-04-12]. <https://arxiv.org/abs/2210.05397>.
- [26] 刘洋, 冯全, 王书志. 基于轻量级 CNN 的植物病害识别方法及移动端应用[J]. 农业工程学报, 2019, 35(17): 194-204.
- [27] 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.
- [28] 贾鹤鸣, 郎春博, 姜子超. 基于轻量级卷积神经网络的植物叶片病害识别方法[J]. 计算机应用, 2021, 41(6): 1812-1819.
- [29] 徐振南, 王建坤, 胡益嘉, 等. 基于 MobileNetV3 的马铃薯病害识别[J]. 江苏农业科学, 2022, 50(10): 176-182.
- [30] Zhang Z X, Wang M W. Convolutional neural network with convolutional block attention module for finger vein recognition [EB/OL]. [2023-03-10]. <http://arxiv.org/abs/2202.06673>. pdf.