

李 帅,薄敬东,龚瑞昆,等. 基于多尺度特征增强的轻量化黄瓜病害识别模型[J]. 江苏农业科学,2024,52(20):267-276.
doi:10.15889/j.issn.1002-1302.2024.20.032

基于多尺度特征增强的轻量化黄瓜病害识别模型

李 帅,薄敬东,龚瑞昆,崔传金
(华北理工大学电气工程学院,河北唐山 063210)

摘要:在复杂的背景环境下对农作物病害进行准确识别与分类,为农作物病害的诊断及防治提供可靠依据,具有重要经济意义。提出了一种新的网络模型——MeNet(multiscale enhance on me),用于对大田中黄瓜的 8 种形态(其中包含 6 种病害和鲜黄瓜、鲜叶)进行精准识别。该模型的设计包括适用于网络前端的特征增强模块,对原始图像进行像素级多尺度特征增强,从而提升模型的特征表达效率;运用特征挑选的思想进行后续的特征提取和增强,再加入基于空域抑制的 SimAM 注意力,进一步突出了显著特征,提高特征效用;运用逐点卷积对特征图进行通道间信息交互,再以全局平均池化总结特征图。结果表明,相较于其他模型,本研究的 MeNet 性能更为优越,在复杂背景病害数据集上,平均准确率达到 92.38%,最高准确率达到了 92.92%,而模型的参数量仅为 0.33 M,浮点运算量仅为 0.30 G,证明 MeNet 模型在图像识别领域具有实际应用的潜力和继续研究的价值。

关键词:黄瓜病害;图像识别;卷积神经网络;轻量化;多尺度特征增强;空域抑制

中图分类号:S126;S436.421.1;TP391.41 **文献标志码:**A **文章编号:**1002-1302(2024)20-0267-10

生物防治是农业可持续发展模式的重要内容,生物防治学科为农业可持续发展提供了重要的科学支撑。近年来,我国已进入加快推进农业可持续发展的历史阶段,这为生物防治学科的发展提供了新的机遇,同时也对生物防治学科的发展提出了更高的要求^[1]。

在深度学习兴起之前,人们主要使用传统的机

器学习方法来解决图像分类问题。这些方法通常涉及人工设计的特征提取器和分类器,如支持向量机(SVM)、随机森林等。这些方法在小规模问题上效果较好,但在复杂图像数据上的分类任务中表现不佳。但是,关于深度学习的探索一直都在进行,一直到 2012 年 Alex 等的 AlexNet 在 ImageNet 图像分类比赛中获胜,标志着卷积神经网络(CNN)正式崭露头角。随着卷积提取特征的特性被认可,深度学习开始向着更深的网络进发,此后的 VGG 和 ResNet 都将网络推向了更深的方向,而 GoogleNet 则提出了多尺度卷积的概念,又让网络变得更宽,通过大量卷积的叠加,追求网络的识别性能,因为人们普遍认为卷积神经网络中特征的复杂性是随网络深度加深而提高的^[2]。在卷积不能满足需求后,又发展出了叠加注意力机制的网络模型,进一步提升识别准确度。

收稿日期:2023-11-09

基金项目:河北省自然科学基金(编号:F2015209308-PT);唐山市科技计划项目(编号:20150212C);河北省高等学校科学技术研究项目(编号:ZD2016070);河北省省级研究生示范课程建设项目(编号:KCJSX2021061)。

作者简介:李 帅(1996—),男,安徽阜阳人,硕士,主要从事检测技术及智能装置研究。E-mail:ncstlishuai@163.com。

通信作者:龚瑞昆,博士,教授,硕士生导师,主要从事检测技术及智能装置研究。E-mail:ncstgongruikun@163.com。

[22]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.

[23]Touvron H, Cord M, Douze M, et al. Training data-efficient image transformers & distillation through attention [EB/OL]. (2020-12-23) [2023-11-01]. <https://arxiv.org/abs/2012.12877>.

[24]Howard A G, Zhu M L, Chen B, et al. MobileNets: efficient convolutional neural networks for mobile vision applications [EB/

OL]. (2017-04-17) [2023-11-01]. <https://arxiv.org/abs/1704.04861>.

[26]Otsu N. A threshold selection method from gray-level histograms [J]. IEEE Transactions on Systems, Man, and Cybernetics, 1979, 9(1): 62-66.

[27]Rother C, Kolmogorov V, Blake A. GrabCut: interactive foreground extraction using iterated graph cuts [M]//Seminal Graphics Papers: Pushing the Boundaries, Volume 2. New York, NY, USA: ACM, 2023: 593-598.

刘阳等以 PlantVillage 数据集为对象,对经典轻量级卷积神经网络 SqueezeNet 进行改进,通过修剪模型,更改通道参数,以牺牲小部分性能换取了大幅度参数量减少^[3];孙俊等以 Kaggle 网站收集到的 11 类病害万张数据为对象,将 CA 坐标注意力和多尺度特征金字塔融入 MobileNet v2,使网络准确率提高了 2.91%^[4];彭红星等以自 IP102 和 PlantVillage 收集到的葡萄病虫害数据为对象,用深度可分离卷积构建双分支特征提取,辅以 CA 坐标注意力改进 MobileNet v2,取得了 1.83 百分点的准确率提升^[5];贾璐等以自农管家平台收集到的 3 种病害为对象,用 Inception 结构和加入 SE 注意力的倒残差结构自建模型,在其自建数据集上取得了平均识别准确率 87.93% 的成绩^[6];王焕鑫等以 PlantVillage 中的 4 种作物为对象,用 ECA 注意力和 AFF 特征融合改进 MobileNet v2,取得了参数量减少 15.37%,识别准确率提升 0.9 百分点的成绩^[7];朱学岩等以自采集的油橄榄自建数据集为对象,以 EfficientNet - B0 为主干,辅以 CBAM 注意力和双线性池化层,搭建双线性网络,提升了识别准确率^[8];杜海顺等以 AI challenger 2018 数据集为对象,用大量的平均池化组成的特征信息融合改进 ConvNeXt 网络,使 ConvNeXt 识别准确率提升了 2.01 百分点^[9];李建威等以国际皮肤影像合作组织的 2018 挑战赛数据集为对象,用 SE 通道注意力和 SimAM 无参注意力改进 ConvNeXt 网络,也取得了较好的成绩^[10];Chen 等以自采集的橡胶树病害数据集为对象,用倒残差结构、SE 注意力、EIoU 损失函数改进了 YOLO v5 模型,在其自建数据集上平均精度达到了 70%,比 YOLO v5 提高了 5.4 百分点^[11]。

经过对参考文献及其他诸多文献的细致分析,总结了以往深度学习模型的一些关键问题:

(1) 在图像分类的研究中出现的数据集大致分为 2 类,一类是单一背景的实验室条件图片,另一类是复杂背景的大田条件图片。通常情况下,后者的图像识别准确率普遍较低,这是一个重要的挑战。

(2) 尽管已有的网络模型已经在植物病害识别中取得了较好的效果,但由于其本身的参数量和存储空间的需求量很大,使其很难应用在移动或嵌入式设备中大量推广^[12]。传统的网络模型通常具有大量参数,且单个参数的利用率较低,也正因为如此让深度学习变得难以解释,宛如一个黑匣子。因此,精简模型参数不仅有利于推广应用,还有利于

进一步解释深度学习,这是一项很有意义的工作。

本研究基于以上 2 个问题,着眼于精简模型参数和参数的可解释性,设计了一种轻量化深度学习网络用以解决黄瓜病害的分类问题。

1 黄瓜病虫害数据集构建及增广处理

本研究所使用的数据集是来源于 Mendeley Data 的一个黄瓜病害数据集,其中包含 8 种类型的黄瓜,即炭疽病、细菌性枯萎病、肚腐病、霜霉病、苜蓿果腐病、蔓枯病、鲜叶和鲜黄瓜,每类 160 张,共 1 280 张图片。其黄瓜病害分类是在农业专家的合作下完成的,该数据集原材料取自真实田间,对比实验室条件下采集的数据集背景噪声更复杂,也具有更强的真实性。

为降低病害诊断模型在训练过程中的过拟合风险,并提高其泛化效果,在现有图像数据基础上,采用数据增广方法对样本进行扩充。采用的主要方法包括翻转、亮度、对比度、饱和度和色调技术生成增强图像,用以模拟不同的拍摄角度和不同的拍摄距离,以尽可能地还原真实的拍摄条件并增加数据数量,这些方法均采用 torchvision 库中的 transforms 方法实现。从数据集中分别随机抽取 80% 作为训练集,在训练集上使用上述数据增广方法;剩余 20% 作为测试集,不使用数据增广方法,用于后续试验。样本质量如图 1 所示。

为了普适性地反映网络模型的性能,评估模型的泛化性能,本研究还给出了在 PlantVillage 数据集上的测试结果,这是一个公开的农作物病害数据集,其中的所有图像均是由专业人士出于科学研究目的而拍摄并标注的,具有很强的专业性,其中共包含 13 种农作物的 38 类叶片图像信息,这 38 类图像细分为 26 类病害叶片和 12 类健康叶片,总计有 54 305 张图片。该数据集的特点是数据量大,标注清晰,但其背景是单一的实验室背景,背景便于区分。在试验中,扰乱数据集的次序,并按照比例 8:2 将其分成训练集和测试集。同时将数据集中的原始图像统一裁剪为 $3 \times 224 \times 224$,使其适应模型的输入。样本质量如图 2 所示。

2 相关方法

2.1 MeNet 模型设计思想和结构

本研究提出了一种轻量化深度学习模型,能够在黄瓜病害的分类任务中实现较高的识别准确率。



图1 黄瓜数据集的图像样本



图2 PlantVillage 数据集的图像样本

与主流网络相比,该模型具有更精简的参数,更容易解释,同时满足小型模型的要求,有利于推广应用。

在图像分类任务中,每张图片是以 1 个三通道的数值矩阵的形式呈现的,每个数值对应 1 个像素点,通过卷积操作后得到特征图,在反向传播的时候不断修正,直到特征图适应我们的预期。也就是说可以把原图片看作一个原始特征图集合;把原图片的每一个通道看作一个原始特征图;把组成原图片的数值矩阵中的每个数值看作是一个原始特征;将源自同一张图片的原始特征图集合、原始特征图、原始特征看作原始特征集合。已有网络只在第一层卷积中利用了上述的原始特征集合,随后便以最浅层特征集合代替原始特征集合向前传播,之后的每一层都是如此更替的模式。笔者提出一种新的主张,不再用卷积得到的特征集合替代上一层的

特征集合,而是将卷积得到的特征集合看作一种关注,叠加在原始特征集合上,以原始特征集合为“骨”,贯穿整个模型,以每一层卷积得到的特征集合为“肉”,不断地附着在“骨”上。在前向传播的过程中,单位“骨”上“肉”越多,卷积时所获得的关注越多,作用越大。用这样的方式可以规范特征图的更新路径,进而精简模型参数,并且能最大化地将原始图片中的有效信息保留在每一层的特征图上传递出来,以提高识别的准确率。

本研究将其命名为 MeNet (multiscale enhance on me)。MeNet 的整体模型如图 3 所示。

结合图 3 可知,MeNet 模型主要由多尺度特征增强模块 (MFE)、特征挑选和空域抑制模块 (CC)、最大池化层、可分离卷积层、平均池化层和全连接层构成。将上述思想用于黄瓜图像分类的任务。

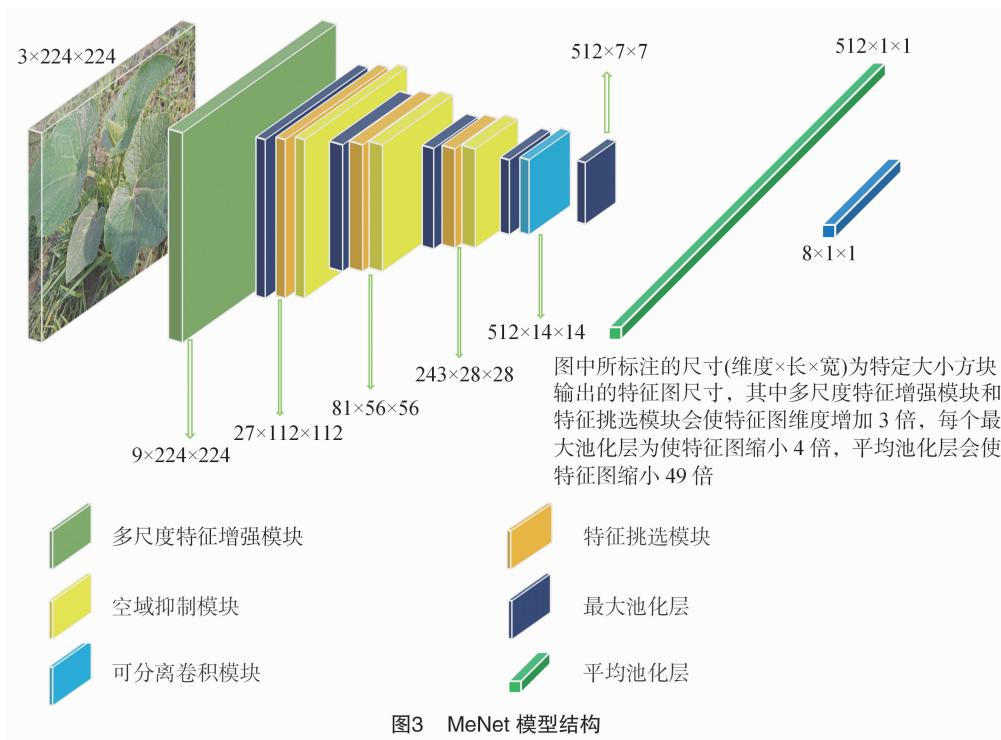


图3 MeNet 模型结构

2.2 多尺度特征增强模块 (multi-scale feature enhancement module, MFE)

在许多模型中,第一次下采样操作通常通过控制卷积操作的步长来实现,用以缩小特征图,后续提取图像的高层特征以支持识别任务^[12]。然而,由于这种下采样方式导致原始图像中不同像素点的计算次数不同,不同的计算次数先天性地给予了不同像素点不同的重要性,这可能会使模型丢失一些原始图像的重要特征^[13]。尤其是对于真实田间背景这种复杂情况,这种损失很可能会导致模型忽略图像边缘的特征,进而影响模型的性能。为了解决这个问题,本研究提出了一种新的方法:在第一次下采样前,让所有像素点的计算次数相同,提取底层特征后对原始图像进行像素级特征增强,使网络在进行第一次下采样前能够保留更多的有效特征,从而提升网络的表达上限,搭配最大池化下采样,起到突出特征、过滤噪声的效果。

多尺度结构的核心思想在于利用不同感受野的卷积核提取不同尺度的信息,其实质是让网络学习一个不同尺度局部区域内像素之间的上下文关系,这有助于更好地捕捉不同尺寸的特征信息^[14-15]。

本模块(MFE)学习输入图像不同尺度的特征信息,并且给予不同尺度特征图不同的权重后,再叠加相应尺度的特征增强;其中的权重由特征图整体平均池化后经过线性变换和 sigmoid 激活后得到,

特征增强所用的增强矩阵由相应尺度的最大池化和 1×1 卷积得到;最后将原始特征集合增维后与权重矩阵和增强矩阵进行像素级融合。

本模块包含2个感受野,分别是 3×3 和 5×5 ,2个分支分别提取不同感受野的特征,由于输入图像都是3通道的RGB图像,为了每个通道的特征图都能均匀得到增强,本模块所得到的特征图通道须为3的倍数。

本模块的结构图如图4所示。

2.3 特征挑选和空域抑制模块 (feature selection and spatial suppression module, SS)

本模块(SS)的主要目的是提取图像特征,并用不同尺寸的最大池化对不同感受野内的特征进行挑选,之后将挑选出来的突出特征与上一层特征图进行像素级融合,让网络在提取特征的同时挑选更为突出的特征来加强自身。

在近些年,注意力机制一直是一个重要的研究方向^[16]。Yang 等认为应由神经科学中的一些统一原则来指导注意力机制的计算,因此基于一些成熟的神经科学理论提出了一种新的方法^[17],可以有效地产生真正的三维权重。要想成功地实现注意力,需要根据特征图来估算单个神经元的重要性。在神经科学中,信息量最大的神经元通常是那些在周围神经元中显示出独特放电模式的神经元。而且,一个活跃的神经元通常会抑制周围神经元的活动,

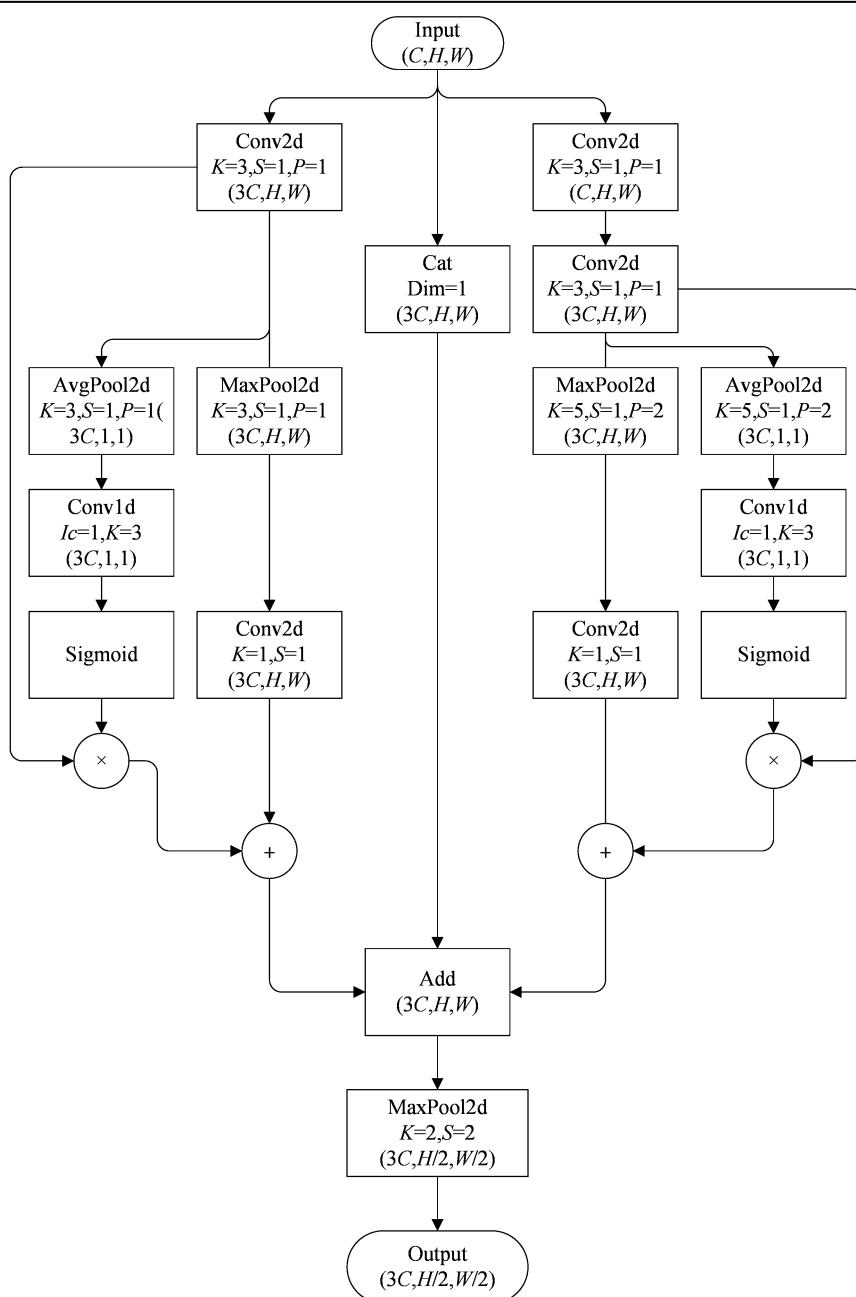


图4 多尺度特征增强模块

这种现象被称为空域抑制。由此得到启发,在视觉加工过程中,表现出明显空域抑制效应的神经元应被赋予更高的优先权(即重要性)。在寻找重要神经元时,最简单的方法就是度量神经元之间的线性可分性。基于这一思想,为每个神经元定义以下能量函数。

$$\begin{cases} e_t(w_t, b_t, y, x_i) = (y - \hat{t})^2 + \frac{1}{M-1} \sum_{i=1}^{M-1} (y - \hat{x}_i)^2 \\ \hat{t} = w_t t + b_t \\ \hat{x}_i = w_t x_i + b_t \end{cases} \quad (1)$$

式中: t 和 x_i 为输入特征 $X \in R^{C \times H \times W}$ 的单个通道中的目标神经元和其他神经元; i 是空间维度上的索引; $M = H \times W$ 是该通道上的神经元数量; w_t 还有 b_t 是计算中的权重和偏置。经过研究人员一步步的计算后,最终给出了一个可以简单计算的公式:

$$e_t^* = \frac{4(\sigma^2 + \varphi)}{(t - \mu)^2 + 2\sigma^2 + 2\varphi} \quad (2)$$

式中: μ 和 σ^2 是除 t 以外的所有神经元的平均值和方差,系数 φ 在 0.000 1 时取得较好的性能。其能量越低,意味着神经元 t 与周围神经元的差异越大,

其重要性也越高。因此,神经元的重要性可以通过 $\frac{1}{e^*}$ 得到。最后,根据各神经元的重要性,对特征图进行优化,并将其总结在一个通道上,则 SimAM 的计算流程可概况为

$$\bar{X} = \text{sigmoid}\left(\frac{1}{E}\right) \cdot X. \quad (3)$$

式中:输出结果 \bar{X} 为增强后的特征, X 为输入特征,

• 为点积运算, E 为每个通道上所有神经元能量函数的值列表, sigmoid 则用来限制 E 值的范围。

通过这种方式,可以更准确地在神经元群体中捕捉到在特定任务中扮演关键角色的神经元以重现人脑中的注意力协同机制,并将其体现在 SimAM 注意力模块中。

整体的模块结构如图 5 所示。

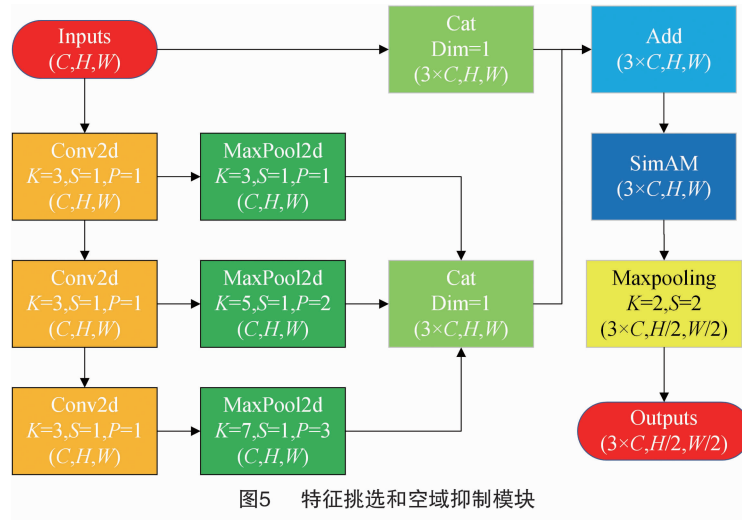


图5 特征挑选和空域抑制模块

2.4 通道混合和压缩模块 (channels and compression module, CC)

深度可分离卷积 (DP) 由 1 个卷积核大小为 3 的深度卷积 (DW) 和 1 个卷积核大小为 1 的逐点卷积 (PW) 组成。相比于普通卷积, DW 能节省非常多的参数量和运算量, 而特征提取效果与普通卷积相比相差不大, 且 PW 可以让各通道间的信息进行融合, 在降低计算成本的同时增加非线性容量, 这有助于模型拟合数据分布, 故而深度可分离卷积模块对于轻量型模型来说是最好的选择之一。

上述几个模块不断将模型的关注点聚焦到病害特征所在位置, 这就使得此时的特征图中包含 2 类特征点: 其一为病灶位置的特征点, 特征效率高, 越靠近病灶位置的特征值就越高; 其二为非病灶位置的特征点, 离病灶位置越远特征值就越低。非病灶位置的特征点在分类器中是有害的, 一方面增加运算量, 一方面干扰分类结果。要解决这一问题需要在最后分类前对特征图进行优化, 采用全局平均池化将特征图压缩为一个值, 由一个值的输出来总结一个特征图的关键信息, 这将会使模型忽略非病灶位置的特征点, 从而提高分类性能, 同时能进一步减少参数量。

本模块 (CC) 的计算过程公式化表达为

$$\begin{cases} X_2 = \text{PW}[\text{DW}(X)] \\ X_3 = \text{MP}\{\text{ReLU}[\text{BN}(X_2)]\} \\ X_4 = \text{AVG}(X_3) \end{cases} \quad (4)$$

式中: X 为模块输入; X_4 为模块输出; DW 为分组卷积; PW 为点卷积; ReLU 为激活函数; BN 为归一化层; MP 为最大池化操作; AVG 为平均池化操作。

3 试验结果

3.1 试验平台

本研究采用的试验平台硬件环境: 处理器为 12 vCPU Intel® Xeon® Platinum 8255C CPU @ 2.50 GHz, 内存为 40 GB, 显卡为 NVIDIA GeForce RTX 2080 Ti, 显存为 11 GB。编程语言为 Python, 深度学习框架: Pytorch 1.11.0, CUDA 版本为 11.3, 进行模型的搭建和训练。

3.2 训练参数

为试验的有效性考量, 所有试验均采用相同的环境参数: 模型在训练和测试时均采用相同的批处理的方式, 每批数据包含 64 幅图像, 即批训练数据量的大小 (batch) 设置为 64。模型训练的迭代总次数 (Epoch) 设置为 300, 采用 Adam 优化器, 初始学

习率设置为 0.001, betas 设置为 (0.9, 0.999), 损失函数使用交叉熵损失函数。

3.3 模型评价指标

本研究是轻量化分类模型在黄瓜病害图像上的应用, 所以将采用 5 个指标^[18]来判断模型的整体性能, 分别是最高准确率、平均准确率、参数量、浮点运算数、混淆矩阵。

其中的平均准确率为 10 次试验在测试集上准确率的平均值, 且平均准确率更高意味着模型的稳定性更好; 最高准确率为 10 次试验在测试集上的最高值, 最高准确率更高意味着模型的上限更好。

参数量、浮点运算量由 thop 库中的 profile 方法导出, 其中浮点运算量为 1 张 224 像素 × 224 像素分辨率的 RGB 图像输入网络后前向传播时共执行的浮点运算次数。

混淆矩阵是由 sklearn 库中的 metrics 方法导出, 经过 pandas 库对混淆矩阵进行美化, 方便分析。本研究混淆矩阵的纵坐标为预测值, 横坐标为真实值。混淆矩阵中自左上至右下的对角线数值为模型对于每种图像预测正确的数量, 哪种图像对应的数值越大则说明模型在此种图像上的分类效果越好。

3.4 消融试验

由于本研究的目的是完成一种轻量化模型的设计, 大体上有 13 层的卷积结构, 故而以 VGG16 网络进行网络剪枝, 以只保留 13 层卷积的 CNN 网络为基础模型, 以下称为 BaseCNN, 以此为基础在试验中验证所包含的模块性能。

按照基础模型 (BaseCNN) 中的特征图尺寸变化规律, 将 BaseCNN 中的普通卷积全部替换为深度可分离卷积模块 (DP), 该模型以 DPNet 代指; 将 DPNet 中的第 1 次下采样 (第 1 个 DP) 替换为本研究所采用的多尺度特征增强模块 (MFE), 称为 MFE_DP; 将模型的中间部分每一个 DP 都 (中间 3 个 DP) 替换为本研究所采用的特征挑选和空域抑制模块 (SS), 称为 MFE_SS_DP; 将 MFE_SS_DP 中的 SimAM 空域抑制删去, 称为 MFE_S_DP; 最后将模型的尾部 (最后一个 DP) 替换为本研究所采用的通道混合和压缩模块 (CC), 称为 MFE_SS_CC (MeNet)。

由表 1 可知, DPNet 的最高准确率和平均准确率均比 BaseCNN 要高, 而且参数量和运算量均超大幅度减少。这是因为深度可分离卷积 (DP) 本身分

表 1 用本研究采用的模块不断替换基础模型中对应模块的性能对比

模型	参数量 (M)	浮点 运算量 (G)	平均 准确率 (%)	最高 准确率 (%)
BaseCNN	10.38	18.15	80.93	82.42
DPNet	0.65	0.41	86.28	87.11
MFE_DP	0.64	0.34	87.73	88.67
MFE_S_DP	0.53	0.30	88.06	89.06
MFE_SS_DP	0.53	0.30	89.45	89.84
MFE_SS_CC (MeNet)	0.33	0.30	92.38	92.92

组卷积加点卷积的组合可以大幅降低参数量和运算量; DP 中的点卷积能有效进行跨通道的特征整合, 且点卷积的本质也可以对原特征图的每一个像素点进行 1 次像素级的加权, 这使得其准确率也有大幅度提高, 这个优点将保留在 CC 模块中。

MFE_DP 的最高准确率和平均准确率均比 DPNet 要高, MFE 模块给 DPNet 带来了 1.45 百分点的平均准确率提升和 1.56 百分点的最高准确率提升, 运算量减少 17.07%, 参数量略微减少。因为 MFE 模块为模型保留了不同感受野的完整图像信息, 对图像边缘的特征保留得更全面, 且能够针对不同尺寸的病害信息给予特征图相应的增强信号。

MFE_SS_DP 和 MFE_DP 的比较说明 SS 模块给 MFE_DP 带来了 1.72 百分点的平均准确率提升和 1.17 百分点的最高准确率提升, 参数量减少 17.19%, 运算量减少了 11.76%。其中, MFE_S_DP 中只包含了特征挑选, 可以使 MFE_DP 的平均准确率提升 0.33 百分点, 最高准确率提升 0.39 百分点。从 MFE_SS_DP 和 MFE_S_DP 的比较中看出, 在特征挑选的基础上加入 SimAM (神经元挑选) 后平均准确率提升了 1.39 百分点, 最高准确率提升了 0.78 百分点, 说明空域抑制和特征挑选是可以配合使用的, 并不会冲突, 且空域抑制模块挑选出的重要神经元在分类中发挥了更强大的作用。通过反复 3 次的特征挑选和神经元挑选, 特征有效性大大提高, 这是轻量化模型可以用少量参数实现较好分类效果的重要条件。

MFE_SS_CC (MeNet) 和 MFE_SS_DP 的比较说明 CC 模块给 MFE_SS_DP 带来了 2.93 百分点的平均准确率提升和 3.08 百分点的最高准确率提升, 参数量减少了 37.74%, 运算量几乎不变。CC 模块与被替换掉的 DP 模块相比, 虽然只多了一个特征图

压缩的步骤,但是这个压缩的过程正是贯彻设计思想的关键一招,大量非病灶位置的特征点的取值比较小,对分类任务起不利影响,用压缩特征图的方式将取值较小的特征点忽略,只有这样才能真正地将前 2 个模块处理好的高效特征的作用发挥出来,筛选有效的特征进入分类器。这种做法直接将下一步全连接分类时的输入特征数减少为原来的 1/49,减少了大量的冗余连接,而保留了重要连接,让分类更加准确,也大大减少了参数量。

从整体上来看,试验验证了“2.1”节中的设计思想,可以实现黄瓜病害的分类,且大大提高了特征的有效性,让模型用更少的参数完成了分类的任务。接下来,将用 MeNet 和其他主流模型对比以证明其优越性。

3.5 对比试验

将本研究的网络模型与主流网络模型在相同数据集和训练环境下作对比试验,以评估 MeNet 的网络性能,结果如表 2 所示。

表 2 各主流模型与 MeNet 在黄瓜数据集上的性能对比

模型	参数量 (M)	浮点运算量 (G)	平均准确率 (%)	最高准确率 (%)
VGG16	138.37	15.52	81.33	82.03
MobileNet v2 ^[18]	2.41	0.5	86.76	87.89
MobileNet v3	4.32	0.31	89.80	91.02
EfficientNetB0	14.05	1.74	90.55	91.79
EfficientNetB1 ^[19]	27.23	2.98	90.55	91.79
Inceptionv2	7.34	1.73	89.94	91.41
ShuffleNet v1	2.11	0.27	87.96	88.67
ShuffleNet v2	4.57	0.55	85.01	86.32
GhostNet ^[20]	3.13	0.22	87.30	88.28
ConvNeXt ^[21]	49.42	8.68	86.52	87.11
MeNet	0.33	0.30	92.38	92.92

由表 2 可知,本研究提出的模型在复杂背景的黄瓜数据集上有着出色的性能表现。与老牌网络 VGG16 相比,MeNet 无论是参数量、浮点运算量还是准确率都更优秀,充分说明对于深度学习来说,特征数量不是越多越好;与模型体量最接近的 GhostNet、ShuffleNet v1、MobileNet v3 网络相比,MeNet 在参数量和准确率上更具有优势,参数量仅为 ShuffleNet v1 的 15.64%,但是浮点运算数却比 GhostNet 高出了 0.08 G,平均准确率则是最高的,比 GhostNet 高出 5.08 百分点,比 ShuffleNet v1 高出 4.42 百分点,比 MobileNet v3 高出 2.58 百分点;与准确率表现最接近的 EfficientNetB0、Inception v2 网络相比,MeNet 在网络体量上有着巨大的优势,参数量不足它们的 1/20,浮点运算数也只有它们的 1/6,MeNet 准确率也高出二者至少 1.13 百分点;这说明 MeNet 提取的特征更加简洁、高效,具有更强的可解释性,继续研究的价值也更高。

3.6 可视化分析

使用 Grad-CAM++ 方法对模型的最后一层卷积层进行可视化分析,用叠加热力图的方式查看网

络经过特征增强和挑选后所关注的特征。

计算网络模型的混淆矩阵,通过数据集内容和混淆矩阵的对照,可以观察出数据集的难分点和网络模型的缺点,这对指明以后的研究方向很有帮助。可视化结果如图 6 和图 7 所示。

由图 6 可知,MeNet 模型可以很好地将关注集中在病灶区域,这是模型能够正确分类的保障,同时也验证了模型切实提高了特征表达效率,使得模型在参数减少的同时保证准确率。

由图 7 可知,MeNet 模型整体分类效果较好,最影响结果的是炭疽病 (anthracnose) 和蔓枯病 (gummy stem blight),这 2 种病害的识别最为困难,因 2 种病害病灶区域极其相似,故而容易相互混淆。2 种病害的图像如图 8 所示。

3.7 泛化性试验

将本研究的模型在大型公开数据集 PlantVillage 上运行,训练集与测试集比例均为 8:2,用以测试模型的泛化能力。在对比试验中挑选参数量最接近的和准确率最接近的模型共同参与。试验结果如表 3 所示。

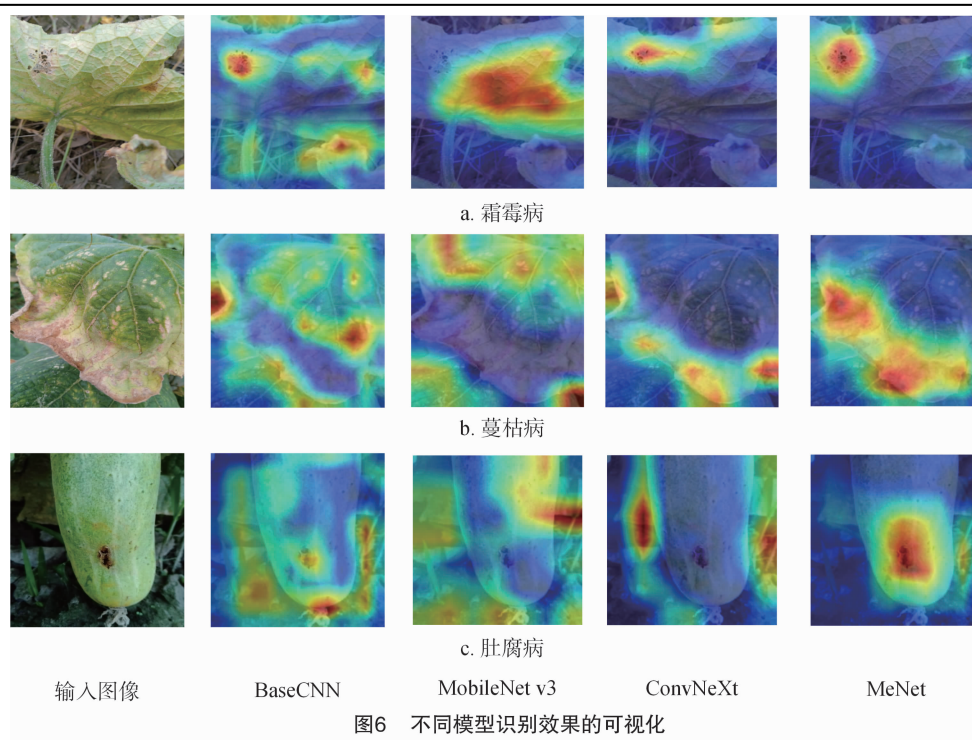


图6 不同模型识别效果的可视化

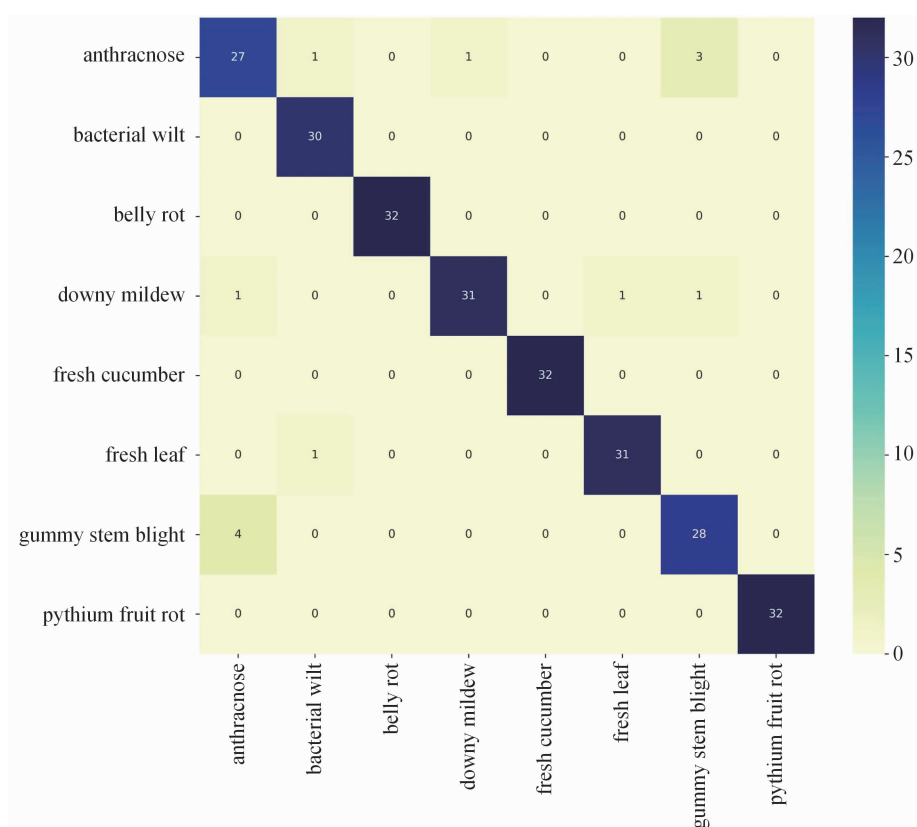


图7 MeNet 的混淆矩阵

从表 3 可以明显看出,与 GhostNet 和 EfficientNetB0 模型相比,本研究提出的 MeNet 模型不仅在准确率方面表现更出色,而且模型的体积也

更小。就 MeNet 模型本身而言,其准确率完全满足需求。且依据 MeNet 在大型数据集上的表现可以推断,MeNet 模型同样可以很好地拟合大型数据集,



炭疽病

蔓枯病

图8 炭疽病和蔓枯病图像对比

表3 3个模型在 PlantVillage 数据集上的性能表现

模型	参数量 (M)	浮点运算量 (G)	平均准确率 (%)	最高准确率 (%)
EfficientNetB0	14.08	1.74	99.04	99.20
GhostNet	3.16	0.22	98.73	98.98
MeNet	0.35	0.30	99.45	99.53

有着很大的移植应用空间。

4 结论

以往神经网络模型存在复杂背景下病害识别困难和参数量大 2 个问题,导致网络难解释、难推广、难应用,归根结底是由于以往模型提取的特征有效性不高所致。本研究据此展开研究并得出以下结论:采用“2.1”节提出的以输入图像原始特征集合为“骨”,以特征矩阵为“肉”构建图像特征的方法是可行的;本研究的模型以 0.33 M 的参数在复杂背景下达到了 92.92% 的准确率,优于现有主流模型,证实本研究的方法是有效的;模型可视化结果证实了本研究的方法可以规范特征图的更新路径,让特征更易解释;参数量的大幅降低验证了本方法能用少量高质量特征替换大量低质量特征,提高了特征效率,显著降低了模型参数。

本研究模型参数量与运算量极小,可以轻松移植到许多性能较低的设备上使用。这不仅降低了设备成本,增强了模型的可移植性;还有助于生产商减小设备的体积,增强设备便携性,对于广泛推广电子设备用于农作物病害的识别具有巨大的推动作用。

参考文献:

- [1] 陈学新,杜永均,黄健华,等. 我国作物病虫害生物防治研究与应用最新进展[J]. 植物保护,2023,49(5):340-370.
- [2] Li Z W, Liu F, Yang W J, et al. A survey of convolutional neural networks: analysis, applications, and prospects [J]. IEEE Transactions on Neural Networks and Learning Systems, 2022, 33(12):6999-7019.
- [3] 刘阳,高国琴. 采用改进的 SqueezeNet 模型识别多类叶片病害[J]. 农业工程学报,2021,37(2):187-195.
- [4] 孙俊,朱伟栋,罗元秋,等. 基于改进 MobileNet v2 的田间农作物叶片病害识别[J]. 农业工程学报,2021,37(22):161-169.
- [5] 彭红星,徐慧明,刘华鼎. 融合双分支特征和注意力机制的葡萄病虫害识别模型[J]. 农业工程学报,2022,38(10):156-165.
- [6] 贾璐,叶中华. 基于注意力机制和特征融合的葡萄病害识别模型[J]. 农业机械学报,2023,54(7):223-233.
- [7] 王焕鑫,沈志豪,刘泉,等. 基于改进 MobileNet v2 模型的农作物叶片病害识别研究[J]. 河南农业科学,2023,52(4):143-151.
- [8] 朱学岩,陈锋军,郑一力,等. 融合双线性网络和注意力机制的油橄榄品种识别[J]. 农业工程学报,2023,39(10):183-192.
- [9] 杜海顺,张春海,安文昊,等. 基于多层信息融合和显著性特征增强的农作物病害识别[J]. 农业机械学报,2023,54(7):214-222.
- [10] 李建威,吕晓琪,谷宇. 基于改进 ConvNeXt 的皮肤镜图像分类方法[J]. 计算机工程,2023,49(10):239-246,254.
- [11] Chen Z Y, Wu R H, Lin Y Y, et al. Plant disease recognition model based on improved YOLO v5[J]. Agronomy, 2022, 12(2):365.
- [12] 王军,冯孙铖,程勇. 深度学习的轻量化神经网络结构研究综述[J]. 计算机工程,2021,47(8):1-13.
- [13] 吴建成,郭荣佐,成嘉伟,等. 注意力特征融合的快速遥感图像目标检测算法[J]. 计算机工程与应用,2024,60(1):207-216.
- [14] Ioffe S, Szegedy C. Batch normalization: Accelerating deep network training by reducing internal covariate shift[C]. 32nd International Conference on Machine Learning, ICML 2015, 2015, 1:448-456.
- [15] 齐国红,许新华,师晓丽. 基于多尺度注意力 U-Net 的结球甘蓝青虫检测方法[J]. 江苏农业学报,2023,39(6):1349-1357.
- [16] Wang Q L, Wu B G, Zhu P F, et al. ECA-net: efficient channel attention for deep convolutional neural networks[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle, WA, USA. IEEE, 2020:11531-11539.
- [17] Yang L, Zhang R Y, Li L, et al. Simam: A simple, parameter-free attention module for convolutional neural networks[C]//International conference on machine learning. Virtual: IEEE, 2021: 11863-11874.
- [18] 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. IEEE, 2018:4510-4520.
- [19] Tan M X, Le Q V. EfficientNet: rethinking model scaling for convolutional neural networks[EB/OL]. (2019-05-28)[2023-11-09]. <http://arxiv.org/abs/1905.11946>.
- [20] Han K, Wang Y H, Tian Q, et al. GhostNet: more features from cheap operations[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle, WA, USA. IEEE, 2020:1577-1586.
- [21] Liu Z, Mao H Z, Wu C Y, et al. A ConvNet for the 2020s[C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). New Orleans, LA, USA. IEEE, 2022: 11966-11976.