

基于改进 ShuffleNet V2 的荔枝叶片病虫害图像识别

谢家兴^{1,2,3}, 陈斌瀚¹, 彭家骏¹, 何培华¹, 景庭威¹,
孙道宗^{1,3}, 高 鹏¹, 王卫星^{1,3}, 郑代德¹, 李 君^{2,4*}

(¹华南农业大学电子工程学院(人工智能学院), 广州 510642; ²岭南现代农业科学与技术广东省实验室茂名分中心, 广东茂名 525000; ³广东省农情信息监测工程技术研究中心, 广州 510642; ⁴华南农业大学工程学院, 广州 510642)

摘要:【目的】探索及时、准确识别危害荔枝叶片病虫害的方法, 做好防护和治理工作。【方法】以常见荔枝叶片病虫害图像为研究对象, 针对荔枝病虫害图像的病斑分布区域广、面积大小不一等特点, 对 ShuffleNet V2 模型进行相应改进。首先, 采集荔枝 5 类叶片病虫害制作数据集, 采用数据增强操作构建更加丰富的数据集; 其次, 在网络特征提取模块采用混合空洞卷积, 避免传统卷积在下采样过程中造成图像分辨率降低及信息丢失; 然后, 通过嵌入轻量型通道注意力模块 ECA (efficient channel attention), 增进特征图之间相互依赖关系。此外, 删减模型中不必要的层数和通道数, 降低模型的参数量及运算量。【结果】研究所改进模型在 5 类荔枝病虫害图像(毛毡病、炭疽病、煤烟病、叶瘿蚊、藻斑病)中达到了 99.04% 的识别准确率, 比原网络 ShuffleNet V2 高出 2.55%。相较于经典网络 AlexNet、ResNet-18、DenseNet 和 MobileNet V2 等, 改进模型不仅有着更高的识别准确率, 并且改进后模型参数量仅为 0.059×10^6 , 为原模型的 4.92%, 浮点运算量仅为 0.183×10^9 。【结论】研究结果适合部署在移动终端等嵌入式资源受限设备上, 有助于实现对作物病虫害实时、准确地识别。

关键词:荔枝; 叶斑症状; 图像识别; ShuffleNet V2 模型; 模型参数

中图分类号: S667.1

文献标志码: A

文章编号: 1009-9980(2023)05-1024-12

Image recognition of leaf pests and diseases based on improved ShuffleNet V2 in litchi

XIE Jiaying^{1, 2, 3}, CHEN Binhan¹, PENG Jiajun¹, HE Peihua¹, JING Tingwei¹, SUN Daozong^{1, 3}, GAO Peng¹, WANG Weixing^{1, 3}, ZHENG Daide¹, LI Jun^{2, 4*}

(¹College of Electronic Engineering (College of Artificial Intelligence) of South China Agricultural University, Guangzhou 510642, Guangdong, China; ²Maoming Branch, Guangdong Laboratory for Lingnan Modern Agriculture, Maoming 525000, Guangdong, China; ³Engineering Research Center for Monitoring Agricultural Information of Guangdong Province, Guangzhou 510642, Guangdong, China; ⁴College of Engineering of South China Agricultural University, Guangzhou 510642, Guangdong, China)

Abstract: 【Objectives】Litchi suffers from many kinds of pests and diseases. Therefore, it is necessary to invest enough energy and funds to control them to ensure the normal growth of litchi. At present, the leaf pest and disease identification of litchi is a problem urgently to be solved. In order to explore ways to identify litchi leaf pests and diseases in a timely and accurate manner, the experiment was undertaken, so as to take preventive and control measures in a timely manner. In this study, common leaf pest and disease images were taken as the research objects, and the ShuffleNet V2 model was improved accordingly for the difficulty to accurately identify if the types of pests and diseases on litchi leaves had the characteristics of large distribution area and different size of lesions. 【Methods】First, five types of

收稿日期: 2022-11-10 接受日期: 2022-12-08

基金项目: 华南农业大学新农村发展研究院农业科技合作共建项目(2021XNYNYKJHZGJ032); 茂名实验室自主科研项目(2021ZZ002); 广东省现代农业产业技术体系创新团队建设专项(2022KJ108); 岭南现代农业实验室科研项目(NT2021009); 广东省乡村振兴战略专项(农业科技能力提升)(TS-1-4); 财政部和农业农村部: 国家现代农业产业技术体系(CARS-32-14); 广东省科技创新战略专项资金(“攀登计划”专项资金)项目(pdjh2021b0077, pdjh2021b0077); 大学生创新创业训练计划项目

作者简介: 谢家兴, 讲师, 博士, 研究方向为农情信息获取与智慧农业。Tel: 13602464331, E-mail: xjx1998@scau.edu.cn

*通信作者 Author for correspondence. Tel: 13533517092, E-mail: autojunli@scau.edu.cn

litchi leaf pests and diseases (Algal leaf spot, *Aceria litchi*, Sooty mold, Anthraconse and *Dasineura* sp.) were collected as the data set for the model test. In order to improve the robustness of the model, the original data were augmented by methods such as flipping, cropping, adding noise and changing contrast to obtain a more abundant data set. Second, hybrid dilated convolution was used in the network feature extraction module to obtain a larger receptive field to avoid the loss of information during the downsampling process, and to eliminate the local information loss caused by the use of ordinary dilated convolution stacking. The bigger the receptive field, the larger the range of the corresponding original image, which means that it contains more global and higher semantic level features. In terms of image classification, the region of interest is often distributed in multiple areas of the image, and more global information and higher-level feature information are needed to better identify the target. Third, the attention mechanism can better aggregate the feature information of the target to be recognized by the network model and reduce the influence of irrelevant background. By embedding a lightweight channel attention mechanism ECA (Efficient Channel Attention) in the model, the interdependence between feature maps is improved. In addition, by using hybrid atrous convolution and embedding ECA attention mechanism to effectively improve the information extraction, the image data do not need deeper structure and more channels to obtain the final classification structure. Therefore, in order to achieve a lightweight model purpose, it is required to delete the unnecessary number of layers and channels in the model, adjust the number of modules in stage2, stage3, stage4 to 3, 6, 3, and change the number of channels to 16, 24, 48, 96, 192, reducing the model parameters and calculations. **【Results】** The improved model in this study achieved a recognition accuracy of 99.04% in terms to five types of litchi pests and disease images (Algal leaf spot, *Aceria litchi*, Sooty mold, Anthraconse and *Dasineura* sp.), which was 2.55% higher than the original network ShuffleNet V2. The traditional convolution reduces the size of the feature map in downsampling, which leads to the loss of information along with the decrease in resolution; when the feature map is the same, the dilated convolution can obtain a larger receptive field, thereby obtaining denser feature information. It has a better recognition rate for small targets, but the ordinary dilated convolution has a grid effect, which causes part of the information to be lost when the same dilated rate is superimposed; this study uses a hybrid dilated convolution, which avoids the use of traditional convolution or ordinary dilated convolution. Part of the information is lost during downsampling by convolution, and the accuracy is higher than 95.53% by using ordinary dilated convolution and 96.61% by using the original network; Compared with the SE (Squeeze-and-Excitation Attention Module) attention module and the CBAM attention module (Convolutional Block Attention Module), the ECA attention module adopts a local cross-channel interaction strategy without dimensionality reduction and adaptive selection of the method of 2D convolution kernel size, and the method of summarizing cross-channel information through 1D convolution layer to obtain more accurate attention information has better accuracy and lower parameter amount. Compared with the classic networks AlexNet, ResNet-18, DenseNet and MobileNet V2, the improved model has great advantages in accuracy, parameter amount and computation amount, and can maintain a low parameter amount and computation amount, which has the highest accuracy. The number of parameters of the improved model is only 0.059×10^8 , which is 4.92% of the original model. The amount of floating point operations is only 0.183×10^8 . **【Conclusion】** The results of this study show that atrous convolution with atrous rate intersected by [1, 2, 3] in feature extraction maintains high resolution and prevents information loss during downsampling, while embedding ECA attention module in the model improves key information. Therefore, more useful information can be obtained, and a better classification effect can be obtained; and the model, the

number of network layers and the number of channels can be further deleted on the premise of achieving the recognition effect, which better balances the amount of parameters of network. The three indicators of accuracy and calculation amount greatly reduce the amount of model parameters and calculation amount of the model, and at the same time improve the performance of the model to a higher level, which is conducive to deployment in embedded resource-constrained devices like mobile terminals. It helps to realize real-time and accurate identification of crop pests and diseases.

Key words: Litchi; Leaf spot symptom; Image recognition; ShuffleNet V2 model; Model parameter

中国荔枝种植面积、产量和产值均居世界首位, 种植面积达 47.57 万 hm^2 ^[1]。荔枝作为目前广东最重要的经济作物之一, 种植面积以 27.3 万 hm^2 居全国第一, 品种最为丰富, 种质资源在全国乃至国际上独具优势^[2]。但是荔枝在种植过程易受到多种病虫害的侵袭, 因此, 快速、准确地判断荔枝病虫害类型是植保精准施药、降低损失的前提和基础^[3]。传统病虫害识别和检测主要依靠人工根据经验现场观察判定的方法, 存在判定效率低、成本高、受主观性影响等问题, 无法满足现代农业大规模、实时检测的需求^[4-6]。

随着计算机技术的发展, 研究人员开始将机器学习和图像处理的方法作为对作物病虫害进行识别的一种技术手段。在作物病虫害识别研究方面, Hlaing 等^[7]通过 Johnson SB 分布模型对番茄图像的尺度不变特征转换(scale invariant feature transform, SIFT)特征进行建模, 获得图像的纹理统计信息, 结合颜色特征信息, 经过支持向量机训练, 平均精度可达 85.1%。党满意等^[8]通过将马铃薯晚疫病叶片提取的颜色、纹理及形状特征参数结合起来进行建模, 识别率达 90%, 识别时间为 9 s。卜俊怡等^[9]通过提取诱虫板图像 RGB 颜色模型中的 B 分量和 HSV 颜色模型中 V 分量, 利用最大类间方差法分割提取颜色、形状及纹理特征, 识别率达到 90.40%。上述基于传统图像处理和机器学习方法的试验, 关键在于提取到丰富、准确的特征。然而, 由于作物病害图像纹理、颜色等特征非常复杂且不规则, 传统方法中提取的低层次特征无法应用于大量、多类目标的识别^[10]。

近年来, 深度学习技术在不断发展, 卷积神经网络(convolutional neural network, CNN)作为深度学习的典型代表^[11-12], 相比于传统图像处理和机器学习的图像分类算法, 卷积神经网络通过卷积操作对输入图像进行特征提取, 不仅能提取到更高层、表达能力更好的特征, 还能实现在一个模型中完成

端到端的特征提取、选择及分类。卷积神经网络在作物叶片病虫害识别相关领域已得到广泛应用。鲍文霞等^[13]利用深度语义分割网络 U-Net 分割出麦穗图像, 并设计了多路卷积神经网络进行小麦赤霉病的识别, 识别精度超过 98%; Zhang 等^[14]提出了全局池化扩张卷积网络(global pooling dilated convolutional neural network, GPDCNN)用于植物病害识别, 对 6 种常见黄瓜叶片病害的识别精度达 94.65%; Liu 等^[15]通过基于全局对比度的显著性区域检测算法计算用于定位害虫的显著性图, 提取目标的边界框, 通过卷积神经网络对水稻害虫进行识别, 平均精度达到 95.1%。上述研究基于 CNN 的方法通过提取农作物病害图像的深层特征, 在一定程度上提高了病害识别的准确率。由于荔枝叶片病虫害的纹理形状等相近, 为了能够更加准确地识别荔枝叶片病虫害类型, 需要在深层特征基础上进一步利用注意力机制(attention mechanism)选择关注对目标更为关键的信息^[16]。深度学习中的注意力机制源于对人类视觉的研究, 通过学习的方式自适应对每张特征图的重要性进行选取, 从众多特征信息中选择出对当前目标任务更为关键的信息^[17]。例如, 通道注意力模块 SE 中通过挤压(squeeze)将每一张特征图挤压成一个具有全局信息的实数, 接着经过激励(excitation)将此信息通过全连接层与激活函数放大对识别任务重要的特征图权重, 缩小不重要的特征图权重^[18]。坐标注意力模块(coordinate attention, CA)通过在像素坐标系上的有效定位, 将位置信息嵌入到通道注意力从而捕获方向感知和位置敏感信息, 最终达到更好的目标分类效果^[19]。双注意力机制(DA-Net)通过综合通道和空间的特征信息对关键特征权重进行更新, 将更多的注意力放在更易于进行区分的关键特征之上^[20]。Junde 等^[21]在轻量级神经网络模型 MobileNet V2 中结合了通道注意力机(CAM), 增强了对复杂背景害虫的识别能力, 平均准确率达 92.79%; Yun 等^[22]

将改进的通道和空间注意力模块嵌入到 ResNet 中,减少了通道之间的信息冗余,并将注意力集中在特征图中信息最丰富的区域,平均准确率超过 95.37%;王美华等^[23]对 CBAM 注意力模块进行改进并嵌入到 ResNet、AlexNet 等网络中,平均准确率均得到了 0.5%的提升。

上述研究通过下采样的方式逐步降低特征图尺寸(典型尺寸 7×7),虽然下采样的方式能够降低运算量,但同时也伴随着分辨率的降低导致很多小特征信息的丢失,例如本研究中叶片较小面积的病斑特征。笔者通过对 ShuffleNet V2 模型添加注意力模块 ECA 自适应对特征图重要信息进行提取;采用混合空洞卷积获得大感受野同时避免下采样造成的

小特征信息丢失;采用分组卷积等设计手段对荔枝 5 种叶片病虫害进行训练及测试,结果显示在提高了对荔枝叶片病虫害识别率的同时进一步减少了网络模型的参数量及浮点运算量,为小病斑识别及模型移植到可移动设备提供思路。

1 材料和方法

1.1 试验数据获取

图像数据集在 2020 年 6—10 月采集于广东省广州市柯木塱农业技术推广中心的荔枝园,分别采集了荔枝叶片毛毡病、炭疽病、煤烟病、叶瘿蚊、藻斑病 5 种病虫害图片数据共 1160 张,5 种病虫害数据图片如图 1 所示。图中显示了叶片病虫害存在着病斑面



图 1 不同荔枝病虫害的图像

Fig. 1 Images of different litchi diseases

积大小不一且分布区域广的特点。

1.2 试验数据扩充

由于卷积神经网络的训练属于监督学习,需要大量的样本进行训练,为避免网络过拟合并增强模型的鲁棒性,因此使用 Python 中的开源工具库 imgaug 对原始数据集进行数据增强。采用水平镜像翻转、垂直镜像翻转、随机比例裁剪、添加高斯噪声、改变亮度和对比度等随机 2~4 种方法叠加增强,增加数据的多样性并将图像尺寸归一化之 224×224 像素。增强后数据集共 6523 张。随机选取增强后数据集中的 80%作为训练集,其余 20%作

为测试集,扩充后的叶片病虫害种类和数目如表 1 所示。

表 1 荔枝叶片病虫害数据分布

Table 1 Distribution of litchi leaf disease data

病虫害类型 Types of diseases	原始数目 Originalnumber	增强后数目 Enhancednumber	
		训练集 Train	测试集 Test
毛毡病 Aceria litchi	225	1080	187
炭疽病 Anthraconse	231	1059	174
煤烟病 Sooty mold	228	1163	198
叶瘿蚊 Dasineura sp	235	1184	194
藻斑病 Algal leaf spot	245	1098	186

1.3 方法

试验运行环境为 Windows10(64 位)操作系统,训练基于 PyTorch 深度学习框架,编程语言为 Python3.8, GPU 为 GeForce GTX 1080Ti, CPU 为 Intel (R) Core(TM)i7-6800K, CUDA11.0 版本。试验中选取所有图像样本的 80%用于训练,剩余 20%作为测试集对模型效果进行验证并输出病虫害类别的识别情况。超参数根据文献[24]和本研究对比试验情况进行设置,批处理大小(Batch-size)设置为 32,模型的训练轮次为 100 epoch,模型训练的优化器选用动量梯度下降算法(gradient descent with momentum),学习率设置为 0.01,动量取 0.9,为避免模型发生过拟合,在全连接层采用了随机失活(dropout)技术,其参数值设置为 0.5。在特征提取模块采用混合空洞卷积,分别设置采用普通卷积、空洞卷积、混合空洞卷积 3 种不同方式的对照组。通过增加网络感受野方法,在保持图像分辨率的同时能够提供更多的

全局信息。空洞卷积(dilated convolution)是在标准的卷积核中注入空洞(图 2-a),以此来增加模型的感受野尺寸,见公式(1)。相比原来的普通卷积操作,它以一定的步长跳过输入值,从而覆盖更大的输入^[25]。

$$s = (d - 1) \times (k - 1) + k \tag{1}$$

式中 s 表示感受野的大小, d 表示空洞率, k 表示卷积核大小。

从公式(1)的定义可以看出,相比于标准卷积,空洞卷积可以在不增加参数的情况下获取更大的感受野。如图 2-b 所示, 3×3 的标准卷积的感受野为 3×3 , 空洞率为 2 的 3×3 空洞卷积的感受野为 5×5 。

空洞卷积虽然能够增加感受野,但是如果相同空洞率的卷积层堆叠会造成局部信息的丢失(图 2-c)。stage 模块中的分支 2 采用混合空洞卷积(hybrid dilated convolution)^[26](图 2-d)。

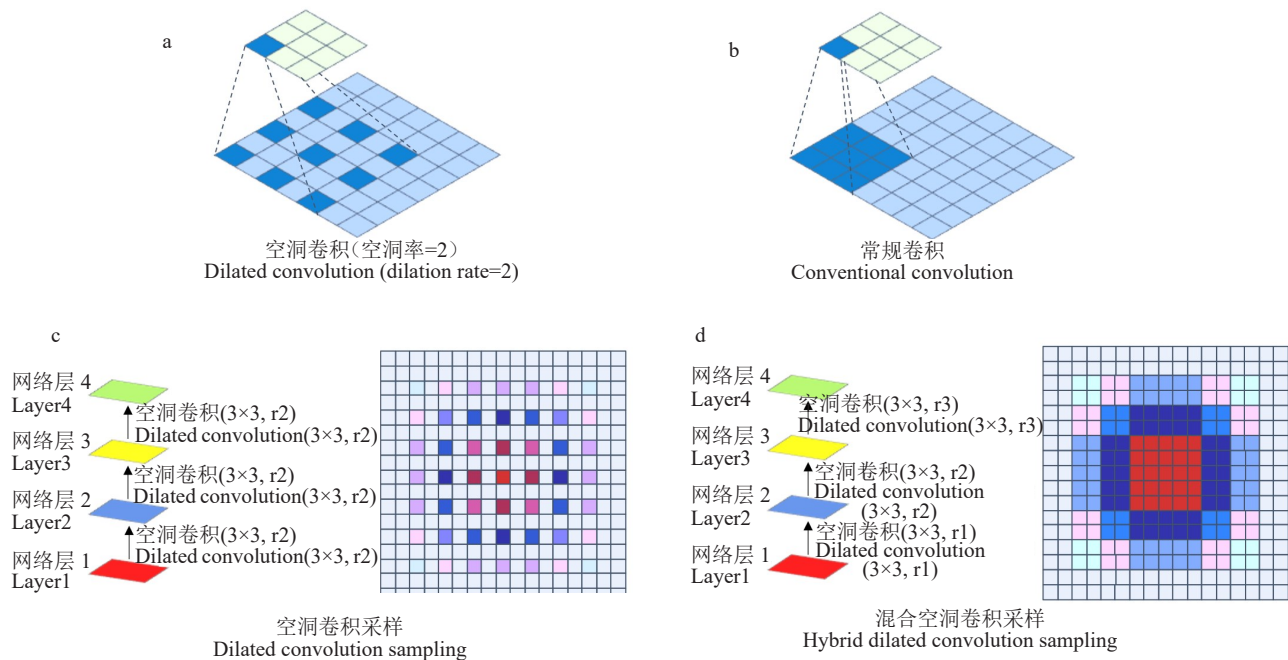


图 2 空洞卷积及其采样示意图

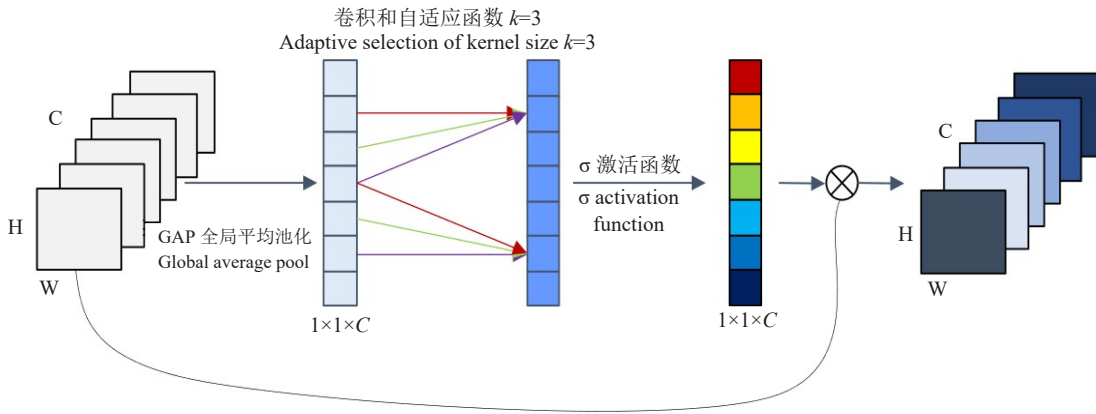
Fig. 2 Schematic diagram of dilated convolution and its sampling

通过嵌入轻量型通道注意力模块 ECA, 并分别设置了 4 个对照组, 对模型中嵌入 ECA 注意力模块^[27]、BAM 注意力模块、SE 注意力模块、CBAM 注意力模块 4 种不同注意力模块进行了对比试验。ECA 的结构如图 3 所示。

为了达到模型轻量化的目的, 将 ShuffleNet V2^[28-29] 结构中的 stage2、stage3、stage4 中模块个数调

整为 3、6、3 个, 通道数调整为 16、24、48、96、192。模型结构如图 4 所示。

设置 7 个对照组, 采用控制单一变量的方式, 对混合空洞卷积、ECA 注意力模块、模型结构调整 3 个因素对模型的影响进行了消融试验; 最后将模型分别与 AlexNet、ResNet-18、DenseNet 和 MobileNet V2 4 种不同模型进行对比。



H. 特征图高度;W. 特征图宽度;C. 特征图通道数;⊗. 特征图对应元素相乘。

H. Feature map height; W. Feature map width; C. Number of feature map channels; ⊗. Multiply the corresponding elements of the feature map.

图3 ECA 注意力模块结构

Fig. 3 ECA attention mechanism structure

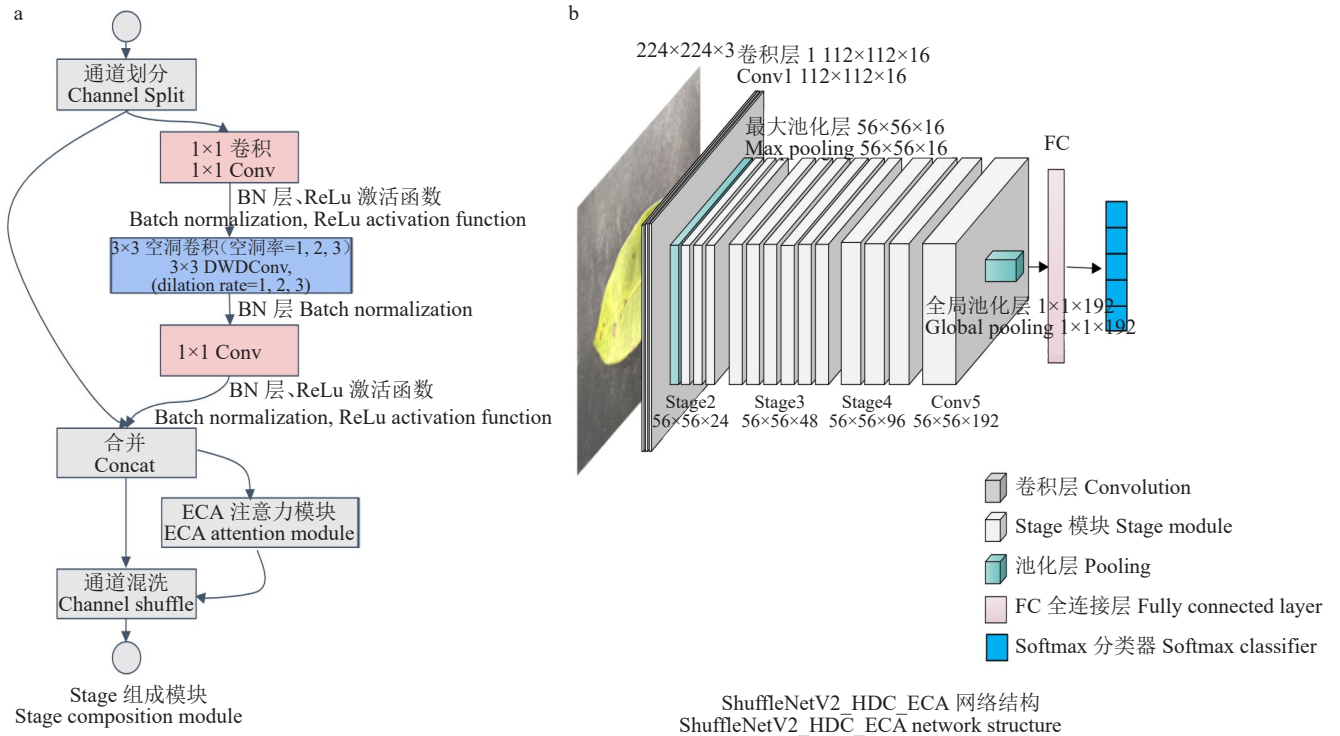


图4 ShuffleNet V2_HDC_ECA 模型结构

Fig. 4 ShuffleNet V2_HDC_ECA model structure diagram

1.4 评价指标

为了客观评价改进模型对荔枝病虫害检测分类的性能,本研究中采用5种分类评价指标进行衡量。

1.4.1 查准率(Precision) 是分类器预测的正样本中预测正确的比例,计算公式如下:

P = TP / (TP + FP) * 100% . (2)

式中 TP 表示实际为正样本预测为正样本的数量;FP 表示实际为负样本预测为正样本的数量,P 的

值越大表示模型预测结果为正例中实际正例的占比越大。

1.4.2 查全率(Recall) 是分类器预测的正样本占所有正样本的比例,计算公式如下:

R / (%) = TP / (TP + FN) * 100 . (3)

式中 FN 表示实际为正样本预测为负样本的数量,R 取值越大表示模型真实情况为正例中预测正例的占比越大。

1.4.3 综合分类率(F1-score) 是查准率和查全率的调和平均数,计算公式如下:

$$F1 = 2 \times \frac{P \times R}{P + R} \quad (4)$$

式中 P 表示模型的查准率; R 表示模型的查全率, $F1$ 取值越大表示模型的预测能力越强。

1.4.4 模型参数量(pares) 是模型结构中包含的参数数量,参数量越小模型运算所需的内存空间就越小。

1.4.5 浮点运算量(FLOPs) 用来衡量模型的复杂度, FLOPs 越小模型越容易移植至移动端。

2 结果与分析

2.1 空洞卷积对模型性能的影响

为了验证空洞卷积对 ShuffleNet V2 模型识别

性能的影响,选择 ShuffleNet V2 进行荔枝病虫害识别对比试验, ShuffleNet V2_DC 特征提取模块采用同一尺寸的空洞卷积设计, ShuffleNetV2_HDC 特征提取模块采用混合空洞卷积设计, ShuffleNet V2_HDC_ECA 在特征提取模块结合了混合空洞卷积和 ECA 注意力模块。3 种不同结构模型的训练结果如图 5 所示。

在相同的试验条件下,特征提取模块采用混合空洞卷积搭建的 ShuffleNet V2_HDC 模型在识别准确率上优于原 ShuffleNet V2 和 ShuffleNet V2_DC 模型,其中 ShuffleNet V2_HDC 的最终识别率达到 97.55%; 而原 ShuffleNet V2 模型的最终识别率只有 95.53%, 说明传统下采样造成了信息损失, 识别效

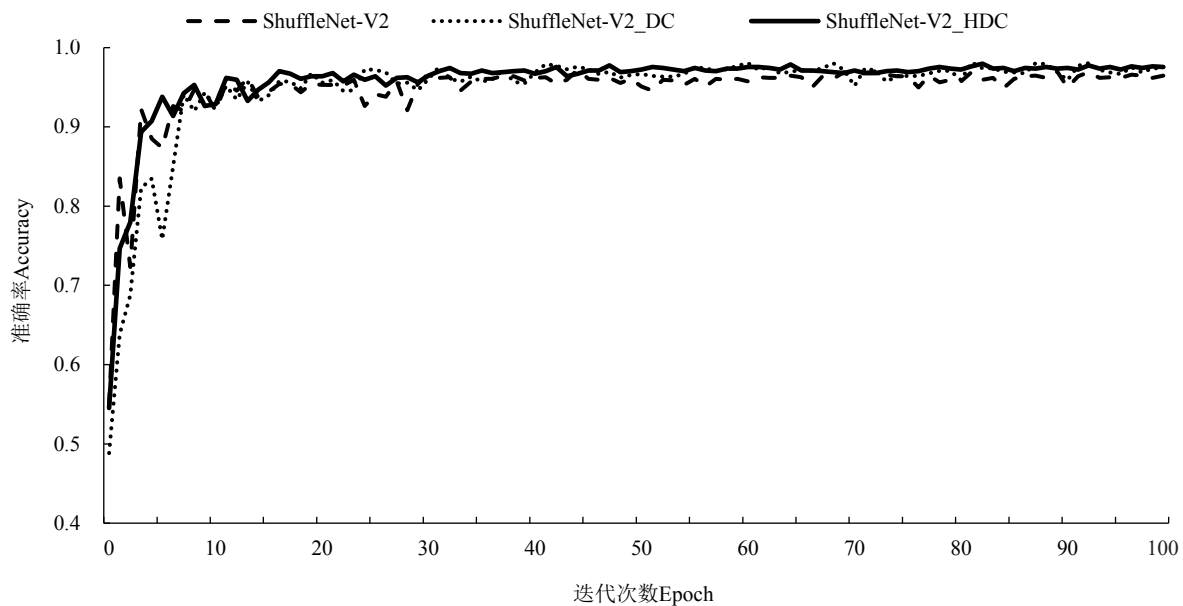


图 5 不同模型的测试分类准确率曲线

Fig. 5 Test classification accuracy curves under different models

果低于采用空洞卷积设计的模型,同时高于 ShuffleNet V2_DC 模型的 96.61%, 说明在使用统一尺寸空洞卷积时产生的网格效应造成了部分信息丢失, 降低了模型识别效果。

2.2 注意力机制对模型性能的影响

为了验证 ECA 注意力模块相比于其他注意力模块对模型的提升效果,在相同试验条件下,使用 BAM 注意力模块、SE 注意力模块和 CBAM 注意力模块替换本研究使用的 ECA 模块进行对比试验。得到网络模型识别结果和模型参数量及运算量,如表 2 所示。

从表 2 可以看出,与其他注意力模块对比,本研究模型添加的 ECA 注意力模块参数量和运算量最少,性能表现最好。其中 SE、BAM 以及 CBAM

表 2 不同注意力模块性能验证

Table 2 Performance verification test results of different attention mechanisms

注意力模块 Attention mechanism	准确率 Accuracy/%	参数量 Parameters/10	运算量 FLOPs/10
BAM	98.51	0.073	0.206
SE	97.55	0.064	0.183
CBAM	98.62	0.065	0.187
ECA	99.04	0.059	0.183

3种注意力模块都存在压缩通道降维操作,不利于学习通道之间的依赖关系,缺少相邻通道之间的信息交互,适当的跨通道信息交互可以显著降低模型复杂度的同时保持性能。研究采用的 ECA 注意力模块在不压缩维度的情况下进行跨通道交互学习,有助于模型学习荔枝病虫害特征通道之间的信息交互和依赖关系,进一步提升了模型识

别性能。

2.3 模型的消融实验结果

为了探究使用 HDC 混合空洞卷积、ECA 注意力模块和结构调整的改进方式对 ShuffleNet V2 模型带来的性能提升效果,进行消融试验,结果如表 3 所示。

从表 3 中可以看出,在 ShuffleNet V2 中采用

表 3 ShuffleNet V2_HDC_ECA 模型的消融试验
Table 3 Ablation experiment of ShuffleNet V2_HDC_ECA model

编号 No.	混合空洞卷积 Hybrid Dilated Convolution (HDC)	注意力模块 Attention mechanism (ECA)	结构调整 Structure adjustment	准确率 Accuracy/%	F1 分数 F1-score/%	参数量 Parameters/10	运算量 FLOPs/10
0	×	×	×	96.49	96.50	0.331	0.141
1	√	×	×	98.08	98.06	0.365	1.111
2	×	√	×	97.02	97.27	0.331	0.145
3	√	×	√	97.55	95.57	0.057	0.179
4	×	√	√	96.59	96.62	0.300	0.035
5	√	√	×	98.94	98.96	0.366	1.117
6	√	√	√	99.04	99.03	0.059	0.183

注: ×表示未使用该改进方法;√表示使用了该改进方法。

Note: × indicates that the improved method is not used;√ indicates that the improved method is used.

HDC 混合空洞卷积代替传统卷积,模型的准确率提升了 1.59 个百分点,F1 分数提升了 1.56 个百分点,模型的参数量减少了 1/3,模型的运算量增加较多。将 ECA 注意力模块添加到 ShuffleNet V2 的基本单元中,模型的准确率和 F1 分数均得到了提升,分别为 0.53 个百分点和 0.77 个百分点,模型参数量和运算量变化较小。综合 HDC 混合空洞卷积、ECA 注意力模块的改进,ShuffleNet V2 的性能得到了较大的提升,准确率和 F1 分数分别提升了 2.45 个百分点和 2.46 个百分点,但是参数量增加了,运算量增加了。通过对模型的结构进行调整,显著降低了模型的参数量和运算量。最终提出的 ShuffleNet V2_HDC_ECA 轻量化模型的识别率为 99.04%,F1 分数为 99.03%,较改进前的 ShuffleNet V2 模型提高了 2.55 个百分点和 2.53 个百分点,而参数量和浮点运算量分别为 0.059 G 和 0.183 M。

为了更直观地展示本研究提出的 ShuffleNet V2_HDC_ECA 轻量化模型的有效性,使用类激活图 (CAM)对病虫害进行可视化分析。CAM 能够以热力图的形式显示模型在训练过程中权重或重心在何处、如何转移,分类模型是根据哪一部分的特征进行判断^[30]。图 6 给出了叶瘿蚊图像在不同模型下的类激活特征热力图及中间特征图。在 ShuffleNet V2

模型中,下采样的分辨率下降导致底层的病斑小特征点的信息丢失,提取到了叶片的边缘但没有提取到病斑特征;使用统一尺寸的空洞卷积设计的 ShuffleNet V2_DC_ECA 模型,对叶片病斑的特征提取存在不足,由于网格效应造成的部分信息丢失,不能够提取到所有的病斑信息;研究模型 ShuffleNet V2_HDC_ECA,在增大感受视野的同时避免了下采样和网格效应造成的小特征信息丢失,准确提取到了病斑特征点。针对荔枝病虫害的病斑分布范围广、面积大小不一的特点,本研究改进的特征提取模块能够更好地学习到全局信息,从而提升了模型分类性能。

2.4 与其他模型性能的对比

为进一步验证所设计的模型识别荔枝病虫害的效果,基于模型竞争力角度选择 AlexNet、ResNet-18、DenseNet 和 MobileNetV2 作为对比模型进行对比试验,在相同的试验条件下进行对比(表 4)。

从表 4 可以看出,不同的网络架构对荔枝叶片病虫害识别效果不同,ShuffleNet V2_HDC_ECA 的识别率达到了 99.04%,参数量和运算量远远小于 AlexNet、ResNet-18、DenseNet,因此改进的模型在识别准确率和参数量、运算量之间取得了良好的平衡,能够应用于移动端等算力有限设备对荔枝叶片

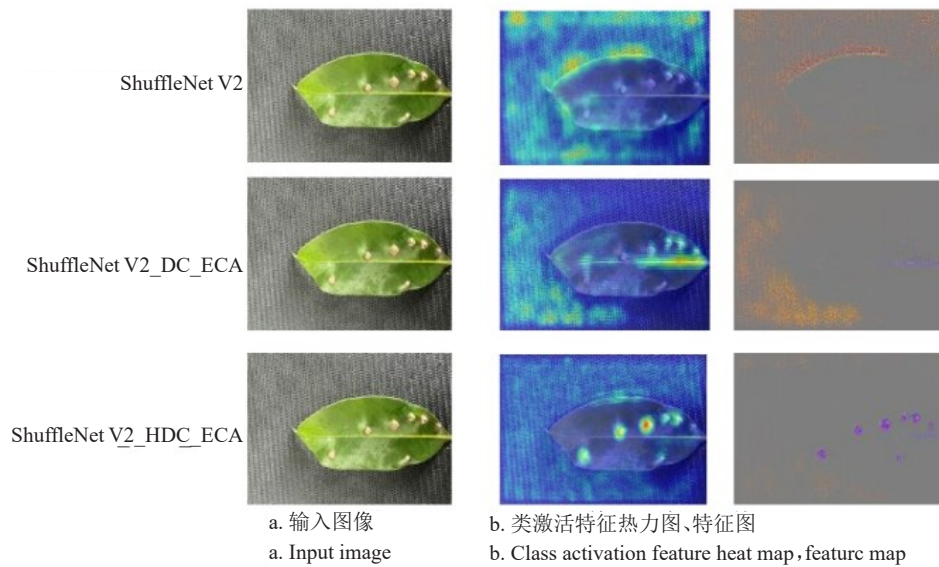


图 6 特征提取模块 stage3 中特征图可视化结果

Fig. 6 Feature map visualization results in the feature extraction module stage3

表 4 不同网络模型的性能对比

Table 4 Performance comparison of different network models

模型 Model	准确率 Accuracy/%	查准率 Precision/%	查全率 Recall/%	F1 分数 F1 Score/%	参数量 Parameters/10	运算量 FLOPs/10
AlexNet	95.53	95.76	95.43	95.59	54.383	0.678
ResNet-18	98.83	98.85	98.84	98.85	10.654	6.416
DenseNet	98.94	98.97	98.95	98.96	6.637	2.668
MobileNetV2	98.08	98.14	98.05	98.10	1.200	0.141
ShuffleNet V2_HDC_ECA	99.04	99.04	99.01	99.03	0.059	0.183

病虫害的识别。并且图 7 的混淆矩阵显示, 5 种病虫害中炭疽病和烟煤病容易被误识别, 本研究通过保留更多底层微小病斑特征而能够有着较低的误识别率。

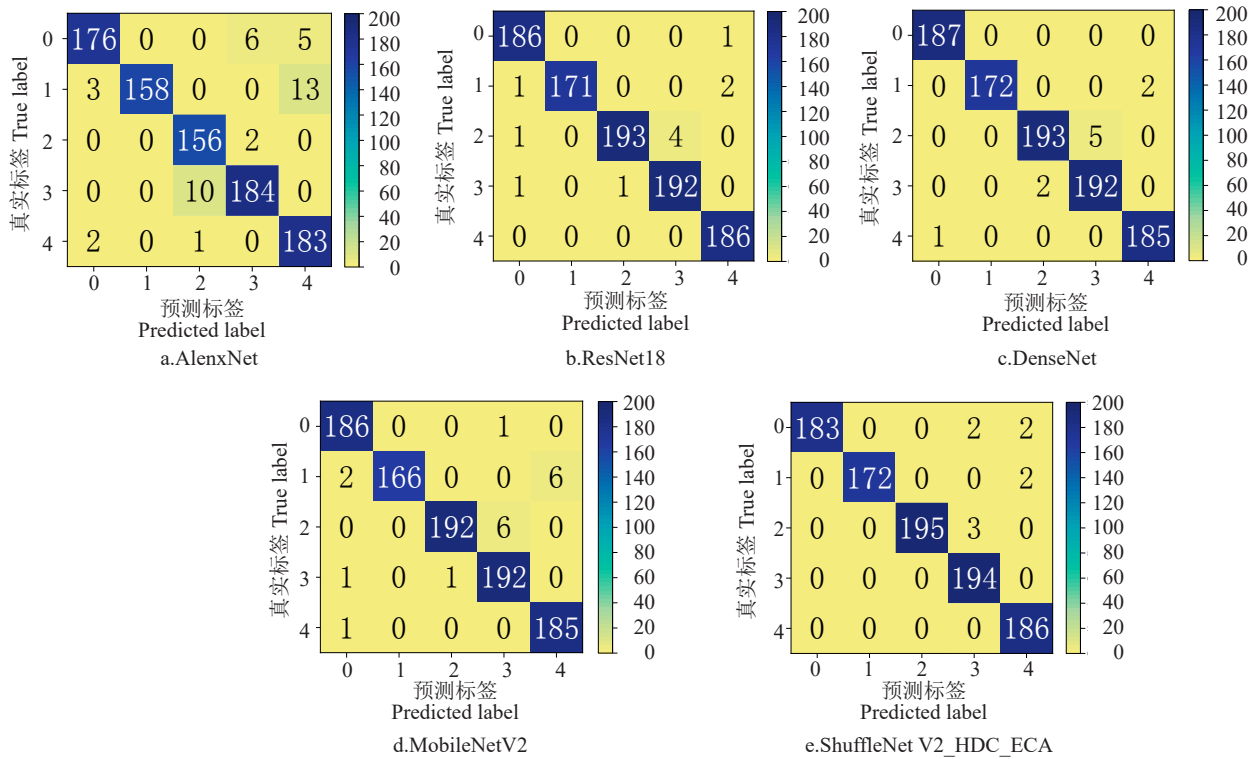
3 讨 论

研究在原 ShuffleNet V2 模型结构的基础上, 采用混合空洞卷积替换了模型中的常规卷积、嵌入 ECA 注意力模块和缩减及删除了不必要的网络层数及通道数等方法并进行了对比试验。

笔者在本研究中分析了普通卷积、空洞卷积和混合空洞卷积对网络性能的影响, 分别使用不同的下采样方式设置对照组进行了验证。发现在图像分类的任务中, 感兴趣区域往往分布在图像的多个区域, 需要更多的全局信息及更高层的特征信息才能更好地识别目标。感受野越小, 它对应的原始图像的范围就越小, 意味着它包含的特征更加趋近于局部和细节, 而用来处理复杂任务

的高级语义信息很难被网络捕捉; 感受野越大, 它对应的原始图像的范围就越大, 意味着它蕴含更多全局的、语义层次更高的特征。在实际环境中, 荔枝叶片病虫害存在着病斑面积大小不一、病虫害分布范围广等问题。采用混合空洞卷积能够提高特征信息的提取, 避免了下采样分辨率下降而导致的小特征信息丢失问题。本研究与孙俊等对网络中不同尺寸的特征图采取上采样融合操作, 构建兼具网络高、低层信息的新特征图避免了特征信息丢失的作用一样^[1], 但本研究采用混合空洞卷积在不需要增加卷积层的情况下便得到更大的感受野不会增加网络的参数量及计算量。

注意力机制能够更好地聚集网络模型对待识别目标的特征信息, 减少无关背景的影响。常用的注意力模块 SE 模块中, 主要由全局平均池化层 (global average pooling, GAP)、全连接层 (fully connected, FC) 和 Sigmoid 函数组成。两个 FC 层的设计是为了捕捉非线性的跨通道交互, 其中包括采用降维操作来控制



0, 1, 2, 3, 4 分别代表荔枝毛毡病、炭疽病、煤烟病、叶瘿蚊、藻斑病。

0, 1, 2, 3, and 4 represent Aceria litchi, Anthraconse, Sooty mold, Dasineura sp, Algal leaf spot, respectively.

图7 不同网络模型识别结果的混淆矩阵

Fig. 7 Confusion matrix of recognition results of different network models

模型的复杂度,在SE模块中进行降维操作对通道预测效果造成负面影响,捕获所有通道之间的依赖是低效的。因此,本研究选用由SE变化改进的ECA注意力模块,该注意力模块采用了一种不降维的局部跨通道交互策略和自适应选择一维卷积核大小的方法,通过一维卷积层汇总跨通道信息的方法获取更加精确的注意力信息。该注意力模块首先将输入特征图进行全局平均池化操作,接着进行卷积核大小为K的一维卷积进行卷积操作并经过sigmoid激活函数得到各个通道的权重,最后将各通道权重与对应原始通道权重相乘得到输出特征图。由于荔枝叶片病虫害的识别主要依据病斑的颜色、形状和纹理,在传统的网络模型中加入ECA注意力模块能够增加网络权重分配给病斑这一分类的关键信息量。

使用混合空洞卷积虽然避免了下采样造成的信息损失,但这同时产生了巨大的运算量。在引入混合空洞卷积后保证了分辨率不变且能够很好地提取到叶片病斑的特征,以及嵌入了ECA注意力模块能够更加有效地使用通道信息,图像数据不需要更深的网络层数和更多的通道数即可获得最终的分类结

果,可以进一步对网络模型进行结构调整减少模型的参数量及浮点计算量。在轻量型网络的研究方面,平衡了准确率、模型参数量和浮点计算量这3个指标,部署在移动终端等嵌入式资源受限设备上,是后续研究的内容之一。

4 结论

针对荔枝叶片病虫害的识别,以5种不同荔枝病虫害图像作为研究对象,通过实地采集的1160张图像构建了荔枝叶片病虫害种类识别图像数据集。在原ShuffleNet V2模型结构的基础上,采用混合空洞卷积替换了模型中的常规卷积、嵌入ECA注意力模块、缩减或删除了不必要的网络层数及通道数等方法并进行了对比试验。

(1)相对于传统分类模型仅依靠网络高层信息进行分类,采用混合空洞卷积进行采样避免了下采样分辨率下降而导致的小特征信息丢失问题,表明本研究的方法能够针对荔枝叶片病虫害特点强化对病斑的特征提取,识别准确率达到了99.04%。

(2)通过与不同的注意力模块进行对比试验,采

用ECA注意力模块自适应提取对识别目标重要的通道信息,在增加少量参数的情况下提升模型识别准确率。

(3)识别模型的试验结果表明,在特征提取中保持高分辨率能够获得更好的分类效果,并且能够在达到识别效果的前提下进一步删减模型网络层数及通道数,有利于将卷积神经网络模型部署在移动端等嵌入式资源受限设备上。

参考文献 References:

- [1] 庄丽娟,邱泽慧. 2019年中国荔枝产业发展特征与政策建议[J]. 中国南方果树,2021,50(4):184-188.
ZHUANG Lijuan, QIU Zehui. Development characteristics and policy recommendations of China's lychee industry in 2019[J]. South China Fruits, 2021, 50(4): 184-188.
- [2] 向旭. 广东荔枝产业发展瓶颈与产业技术研发进展[J]. 广东农业科学, 2020, 47(12): 32-41.
XIANG Xu. Bottleneck of litchi industry development and progress of technology research and development[J]. Guangdong Agricultural Sciences, 2020, 47(12): 32-41.
- [3] 黄康泉. 荔枝主要病虫害的危害症状及防治方法[J]. 现代园艺, 2021, 44(22): 40-41.
HUANG Kangquan. Symptoms and control methods of main diseases and insect pests of litchi[J]. Contemporary Horticulture, 2021, 44(22): 40-41.
- [4] 谢堂胜,刘德营,陈京,蔡俊. 白背飞虱智能识别技术研究[J]. 南京农业大学学报, 2016, 39(3): 519-526.
XIE Tangsheng, LIU Deying, CHEN Jing, CAI Jun. Automatic identification of *Sogatella furcifera*[J]. Journal of Nanjing Agricultural University, 2016, 39(3): 519-526.
- [5] BARMAN U, CHOUDHURY R D, SAHU D, BARMAN G G. Comparison of convolution neural networks for smartphone image based real time classification of citrus leaf disease[J]. Computers and Electronics in Agriculture, 2020, 177: 105661.
- [6] SINGH V, MISRA A K. Detection of plant leaf diseases using image segmentation and soft computing techniques[J]. Information Processing in Agriculture, 2017, 4(1): 41-49.
- [7] HLAING C S, MAUNG ZAW S M. Tomato plant diseases classification using statistical texture feature and color feature[C]// 2018 IEEE/ACIS 17th International Conference on Computer and Information Science (ICIS). IEEE, 2018: 439-444.
- [8] 党满意,孟庆魁,谷芳,顾彪,胡耀华. 基于机器视觉的马铃薯晚疫病快速识别[J]. 农业工程学报, 2020, 36(2): 193-200.
DANG Manyi, MENG Qingkui, GU Fang, GU Biao, HU Yao-hua. Rapid recognition of potato late blight based on machine vision[J]. Transactions of the Chinese Society of Agricultural Engineering, 2020, 36(2): 193-200.
- [9] 翟肇裕,曹益飞,徐焕良,袁培森,王浩云. 农作物病虫害识别关键技术研究综述[J]. 农业机械学报, 2021, 52(7): 1-18.
ZHAI Zhaoyu, CAO Yifei, XU Huanliang, YUAN Peisen, WANG Haoyun. Review of key techniques for crop disease and pest detection[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(7): 1-18.
- [10] 卜俊怡,孙国祥,王迎旭,魏天翔,汪小昂. 基于诱虫板图像的温室番茄作物害虫识别与监测方法[J]. 南京农业大学学报, 2021, 44(2): 373-383.
BU Junyi, SUN Guoxiang, WANG Yingxu, WEI Tianxiang, WANG Xiaochan. Identification and monitoring method of tomato crop pests in greenhouse based on trapping board image[J]. Journal of Nanjing Agricultural University, 2021, 44(2): 373-383.
- [11] LECUN Y, BENGIO Y, HINTON G. Deep learning[J]. Nature, 2015, 521: 436-444.
- [12] LECUN Y, KAVUKCUOGLU K, FARABET C. Convolutional networks and applications in vision[C]//Proceedings of 2010 IEEE international symposium on circuits and systems. IEEE, 2010: 253-256.
- [13] 鲍文霞,孙庆,胡根生,黄林生,梁栋,赵健. 基于多路卷积神经网络的大田小麦赤霉病图像识别[J]. 农业工程学报, 2020, 36(11): 174-181.
BAO Wenxia, SUN Qing, HU Gensheng, HUANG Linsheng, LIANG Dong, ZHAO Jian. Image recognition of field wheat scab based on multi-way convolutional neural network[J]. Transactions of the Chinese Society of Agricultural Engineering, 2020, 36(11): 174-181.
- [14] ZHANG S W, ZHANG S B, ZHANG C L, WANG X F, SHI Y. Cucumber leaf disease identification with global pooling dilated convolutional neural network[J]. Computers and Electronics in Agriculture, 2019, 162: 422-430.
- [15] LIU Z Y, GAO J F, YANG G G, ZHANG H, HE Y. Localization and classification of paddy field pests using a saliency map and deep convolutional neural network[J]. Scientific Reports, 2016, 6(1): 1-12.
- [16] VASWANI A, SHAZEER N, PARMAR N, USZKOREIT J, JONES L, GOMEZ A N, KAISER L, POLOSUKHIN I. Attention is all you need[C]. In Advances in Neural Information Processing Systems, 2017: 6000-6010.
- [17] 张宸嘉,朱磊,俞璐. 卷积神经网络中的注意力机制综述[J]. 计算机工程与应用, 2021, 57(20): 64-72.
ZHANG Chenjia, ZHU Lei, YU Lu. Review of attention mechanism in convolutional neural networks[J]. Computer Engineering and Applications, 2021, 57(20): 64-72.
- [18] HU J, SHEN L, SUN G, ALBANIE S, WU E H. Squeeze and excitation networks[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. IEEE, 2018: 7132-7141.
- [19] HUANG Z L, WANG X G, WEI Y C, HUANG L C, SHI H, LIU W Y, HUANG T S. CCNet: Criss-cross attention for semantic segmentation[C]//Proceedings of the IEEE/CVF International

- Conference on Computer Vision. IEEE, 2019: 603-612.
- [20] FU J, LIU J, TIAN H J, LI Y, BAO Y J, FANG Z W, LU H Q. Dual attention network for scene segmentation[C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. IEEE, 2020: 3141-3149.
- [21] CHEN J D, CHEN W R, ZEB A, ZHANG D F, NANEH-KARAN Y A. Crop pest recognition using attention-embedded lightweight network under field conditions[J]. Applied Entomology and Zoology, 2021, 56(4): 427-442.
- [22] ZHAO Y, CHEN J G, XU X, LEI J S, ZHOU W J. SEV-Net: residual network embedded with attention mechanism for plant disease severity detection[J]. Concurrency and Computation: Practice and Experience, 2021, 33(10): e6161.1-e6161.18.
- [23] 王美华, 吴振鑫, 周祖光. 基于注意力改进 CBAM 的农作物病虫害细粒度识别研究[J]. 农业机械学报, 2021, 52(4): 239-247.
- WANG Meihua, WU Zhenxin, ZHOU Zuguang. Fine-grained identification research of crop pests and diseases based on improved CBAM via attention[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(4): 239-247.
- [24] [MISHKIN D, NIKOLAY S, JIRI M. Systematic evaluation of convolution neural network advances on the imagenet[J]. Computer Vision and Image Understanding, 2017, 161: 11-19.
- [25] YU F, KOLTUN V, FUNKHOUSER T. Dilated residual networks[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. IEEE, 2017: 636-644.
- [26] WANG P Q, CHEN P F, YUAN Y, LIU D, HUANG Z H, HOU X D, COTTRELL G. Understanding convolution for semantic segmentation[C]//2018 IEEE winter conference on applications of computer vision (WACV). IEEE, 2018: 1451-1460.
- [27] WANG Q L, WU B G, ZHU P F, LI P H, ZUO W M, HU Q H. ECA-Net: Efficient channel attention for deep convolutional neural networks[C]//Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. IEEE, 2020: 11531-11539.
- [28] NINGNING M, ZHANG X Y, ZHENG H T, SUN J. ShuffleNet V2: Practical guidelines for efficient CNN architecture design[C]//European Conference on Computer Vision. Cham, Springer, 2018: 122-138.
- [29] 张珂, 冯晓晗, 郭玉荣, 苏昱坤, 赵凯, 赵振兵, 马占宇, 丁巧林. 图像分类的深度卷积神经网络模型综述[J]. 中国图象图形学报, 2021, 26(10): 2305-2325.
- ZHANG Ke, FENG Xiaohan, GUO Yurong, SU Yukun, ZHAO Kai, ZHAO Zhenbing, MA Zhanyu, DING Qiaolin. Overview of deep convolutional neural networks for image classification[J]. Journal of Image and Graphics, 2021, 26(10): 2305-2325.
- [30] ZHOU B L, KHOSLA A, LAPEDRIZA A, OLIVA A, TORRALBA A. Learning deep features for discriminative localization[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. IEEE, 2016: 2921-2929.
- [31] 孙俊, 朱伟栋, 罗元秋, 沈继锋, 陈义德, 周鑫. 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.
- SUN Jun, ZHU Weidong, LUO Yuanqiu, SHEN Jifeng, CHEN Yide, ZHOU Xin. Recognizing the diseases of crop leaves in fields using improved Mobilenet-V2[J]. Transactions of the Chinese Society of Agricultural Engineering, 2021, 37(22): 161-169.