

基于注意力机制和特征融合的葡萄病害识别模型

贾璐 叶中华

(中国农业大学信息与电气工程学院, 北京 100083)

摘要: 植物病害是造成农作物减产的主要原因之一。针对传统的人工诊断方法存在成本高、效率低等问题,构建了一个自然复杂环境下的葡萄病害数据集,该数据集中的图像由农民在实际农业生产中拍摄,同时提出了一个新的网络模型 MANet,该模型可以准确地识别复杂环境下的葡萄病害。在 MANet 中嵌入倒残差模块来构建网络,这极大降低了模型参数量和计算成本。同时,将注意力机制 SENet 模块添加到 MANet 中,提高了模型对病害特征的代表能力,使模型更加注意关键特征,抑制不必要的特征,从而减少图像中复杂背景的影响。此外,设计了一个多尺度特征融合模块 (Multi-scale convolution) 用来提取和融合病害图像的多尺度特征,这进一步提高了模型对不同病害的识别精度。实验结果表明,与其他先进模型相比,本文模型表现出了优越的性能,该模型在自建复杂背景病害数据集上的平均识别准确率为 87.93%,优于其他模型,模型参数量为 2.20×10^6 。同时,为了进一步验证该模型的鲁棒性,还在公开农作物病害数据集上进行了测试,该模型依然表现出较好的识别效果,平均识别准确率为 99.65%,高于其他模型。因此,本文模型具有实际应用潜力。

关键词: 葡萄病害; 图像识别; 卷积神经网络; 多尺度特征融合; 注意力机制

中图分类号: TP391 文献标识码: A 文章编号: 1000-1298(2023)07-0223-11

OSID: 

Grape Disease Recognition Model Based on Attention Mechanism and Feature Fusion

JIA Lu YE Zhonghua

(College of Information and Electrical Engineering, China Agricultural University, Beijing 100083, China)

Abstract: Plant diseases are one of the main causes of crop yield reduction, however, traditional manual diagnosis methods are costly and inefficient, which are difficult to adapt to the demands of modern agricultural production. Recognizing crop diseases automatically and accurately is hence of great importance. Currently, most studies have focused on images taken by professionals for academic purposes, rather than by farmers in actual agricultural production. However, images taken in real applications by farmers are with far more complex backgrounds and hence alleviating the performance of many state-of-art methods. A grape leaf disease dataset were construted under natural complex environments where images were taken by farmers in actual agricultural production. And a network architecture named MANet was proposed for efficient recognition of grape leaf diseases under natural complex environment. The inverted residual module was embedded to build the model, which significantly lowered the number of model parameters. Moreover, the attention mechanism SENet module was used to improve the ability of the model to extract key disease features from complex background images and suppress other irrelevant information. In addition, a multi-scale convolution (MConv) module was designed to extract and fuse multi-scale features of disease images. The experimental results indicated that the proposed model presented a superior performance relative to other most advanced methods. On the public crop disease dataset, MANet achieved the highest average recognition accuracy of 99.65%. And even on the complex background crop disease dataset of the construction, the average recognition accuracy of grape diseases reached 87.93%, which was still better than other state-of-the-art

收稿日期: 2022-10-18 修回日期: 2023-01-14

基金项目: 国家自然科学基金项目(61502500)

作者简介: 贾璐(1983—),女,副教授,博士生导师,主要从事数据挖掘和复杂网络分析研究,E-mail: Ljia@cau.edu.cn

models. Therefore, the proposed model can effectively recognize grape leaf diseases and has certain potential for practical applications.

Key words: grape disease; image recognition; convolutional neural network; multi-scale feature fusion; attention mechanism

0 引言

葡萄是世界上最重要的果树树种之一,它是一种广泛种植在世界各地的作物,并且具有丰富的营养价值、药用价值和经济价值等。然而,每年包括葡萄在内的农作物由于病害侵袭,对其产量造成了严重的损失。最近研究发现,植物病害是造成世界各地农作物产量下降的主要原因之一,这对全球粮食安全也造成了严重的威胁^[1-3]。此外,传统的农作物病害识别方法大多依靠人工观察和经验判断或者在实验室依赖特定的实验分析来识别病害的类别,非常低效并且操作复杂,难以满足实际农业生产的需求。因此研究农作物病害的智能识别,具有重要意义。

随着人工智能技术的飞速发展,许多研究人员使用机器视觉和图像处理的方法来识别农作物病害^[4-7]。传统的机器视觉和图像处理的病害识别方法是从颜色、形状、纹理和小波特征或组合中提取外部特征^[8-11],然后利用分类器对提取的特征进行分类。最后,实现农作物病害的分类。其中常用的分类器包括支持向量机^[12-13]、贝叶斯^[14]、主成分分析^[15]、随机森林^[16-17]和决策树^[18]等。例如,MAJUMDAR等^[5]首先采用模糊 c-means 聚类算法来提取小麦叶片的病害特征,然后利用人工神经网络(ANN)对提取的病害特征进行分类,最终实现对小麦病害的识别。PANTAZI等^[6]采用局部二值模式(LBPs)进行病害特征的提取,并实现了对不同作物病害的自动识别。张燕等^[19]将农作物的颜色纹理等特征结合支持向量机(SVM)提出了一种有效的病害识别模型。然而,这些研究大都是实验室环境下进行的,采集的病害图像数据噪声较少,这与自然环境下的病害图像差距较大,难以满足实际农业生产的需求。

近年来,深度学习方法被应用到各个领域,并且取得了较大的进展,其中卷积神经网络(CNNs)在图像识别领域取得了显著的效果。目前有许多研究者已经提出了一系列经典的 CNN 模型,比如 AlexNet^[20]、VGG^[21]、ResNet^[22]和 DenseNet^[23]等。同时,随着移动设备的普及,一些学者也提出了一些轻量级 CNN 模型,比如 MobileNetV1^[24]、MobileNetV2^[25]、ShufftNetV1^[26]、ShufftNetV2^[27]和 GhostNet^[28]等。此外,许多研究者已将深度学习应用到农作物病虫害识别领域。MOHANTY等^[29]训

练了一个深度卷积神经网络模型,用来识别 14 种作物的 26 种病害,在测试集上的准确率达到 99.35%。LU等^[30]提出了一种基于深度卷积神经网络的模型,并且基于自然环境下采集的水稻病害数据集进行实验,在 10 倍交叉验证策略下,提出的模型准确率达到 95.48%。PRIYADHARSHINI等^[31]提出了一种基于深度卷积神经网络的玉米叶片病害识别模型,然后使用公开数据集 PlantVillage 中的玉米叶片图像进行实验,最终模型在测试集上的识别准确率达到 97.89%。ZHANG等^[32]对 AlexNet 模型进行改进,提出了一种新的植物病害识别模型 GPDCNN,并且在黄瓜病害数据集上展开实验,实验结果表明该模型可以有效识别黄瓜病害。CHEN等^[33]改进了 DenseNet 模型,并设计了一个新的玉米病害识别模型 Mobile-DANet,在其自建数据集上的平均准确率为 95.86%。李书琴等^[34]设计了一种轻量级卷积神经网络模型 Scale-Down ResNet,该模型对植物病害叶片具有较好的识别效果。以上深度学习模型,虽然在植物病害识别领域已经取得了一定的进展,但是经典的 CNN 模型依然存在模型参数量和计算量较大的问题,轻量级的 CNN 模型虽然参数量和计算量大幅降低,但是当应对自然状态下复杂背景病害图像时,其识别准确率仍有待进一步提高。上述研究大多在实验室环境下进行,或者是专业人员为了研究而专门拍摄,并不是在实际生产中拍摄,所以采集到的病害图像噪声较小,与在自然复杂环境中拍摄的病害图像有较大差距,难以满足实际农业生产的需要。

为解决上述问题,本文提出一种新的网络模型 MANet,用于自动识别自然状态下的葡萄病害图像,如白腐病、褐斑病和霜霉病。在该模型中嵌入倒残差模块^[25]来构建网络,以降低模型参数量和计算成本。同时,提出一种多尺度特征融合模块(Multi-scale convolution, MConv),可以提取病害图像多尺度的特征并进行融合,提高模型对不同病害的识别准确率。此外,在 MANet 中嵌入注意力机制 SENet^[35]模块,以增强模型在复杂背景的图像中提取关键病害特征的能力,并抑制其他无关信息。

1 材料和方法

1.1 数据获取

本文使用 2 个植物病害数据集进行实验,一个

是本文构建的复杂背景葡萄病害数据集 (命名为 Grape_Complex), 另一个是公开的 PlantVillage 数据集 (<http://www.plantvillage.org>)。

PlantVillage 是一个公开的农作物病害数据集, 共有 54 305 幅图像, 包含 13 种作物, 涉及 26 类作物病害叶片和 12 类健康叶片, 共计 38 种类别^[36]。这

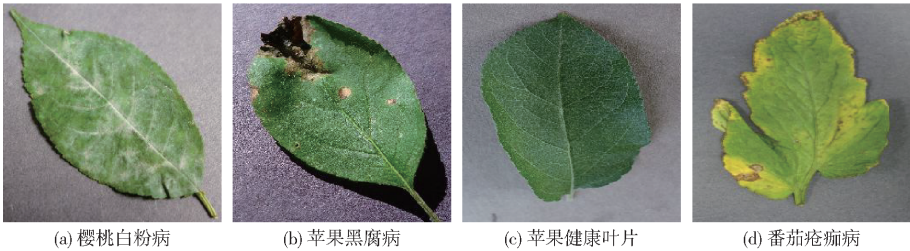


图 1 PlantVillage 数据集中的部分样本图像

Fig. 1 Sample images from a portion of PlantVillage dataset

Grape_Complex 数据集包括 3 种病害 (白腐病、褐斑病、霜霉病) 和健康叶片, 数据主要来源于农管家在线农技问答平台 (<http://www.laodao.so/>)。农管家是一种为农民在实际生产中提供交流的在线农技问答平台, 其中问答板块主要是农民在种植农作物过程中遇到的各种作物病害、虫害等问题, 农民提出问题后会有相关的专家进行解答。基于此, 收集该在线农技问答平台的相关数据, 然后根据专家领域的知识对图

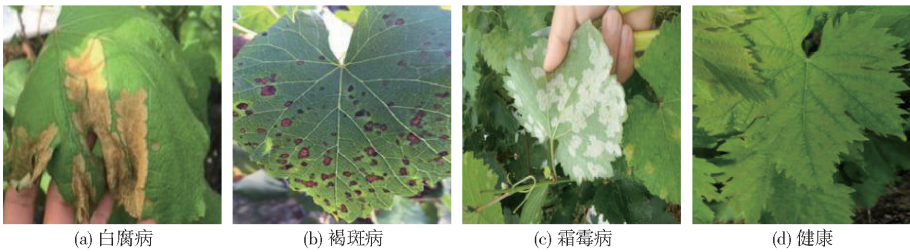


图 2 Grape_Complex 数据集中的样本图像

Fig. 2 Sample images from Grape_Complex dataset

1.2 数据预处理与数据增强

由于收集的 Grape_Complex 数据集中的图像分辨率和尺寸不同, 首先将图像尺寸调整为 $3 \times 224 \times 224$, 以适应模型的输入尺寸, 然后, 将数据集按照比例 8:2 分为训练集和测试集。其中训练集有 1 616 幅图像, 包括 438 幅白腐病图像、573 幅褐斑病图像、331 幅霜霉病图像以及 274 幅健康图像。测试集有 406 幅图像, 包括 110 幅白腐病图像、144 幅褐斑病图像、83 幅霜霉病图像以及 69 幅健康图像。由以上统计可以发现, 自建数据集中样本分布存在不平衡的现象, 这种不平衡的训练数据会干扰到深度学习模型的识别效果^[37-39], 所以本文对原始数据进行数据增强, 扩充训练样本的数量, 以增加训练图像中样本的多样性, 防止模型出现过拟合现象, 同时使模型对自然状态下的病害图像具有较强的泛化能

个数据集中的图像是由专业人士出于科学研究目的而拍摄的, 因此其病害图像的背景比较简单。实验中, 按照比例 8:2 将图像随机分成训练集和测试集。同时将数据集中的原始图像归一化为 $3 \times 224 \times 224$, 使其适应模型的输入。PlantVillage 数据集中的部分样本图像如图 1 所示。

像标记病害类别, 并保存为 JPG 格式。Grape_Complex 数据集中的各类别样本示例如图 2 所示。明显可以看出, 这些图像是农民在实际生产中拍摄的, 比为科学研究目的而去专门拍摄的图像背景更加复杂。本研究收集的病害图像数据更加符合实际农业生产的情况。最终, Grape_Complex 数据集共收集 2 022 幅图像, 包括 548 幅白腐病图像, 717 幅褐斑病图像, 414 幅霜霉病图像, 343 幅健康图像。

力。本文主要通过随机旋转、水平翻转和垂直翻转来扩充训练样本的数量。数据增强后的样本分布如表 1 所示。

表 1 Grape_Complex 数据集数据增强后的样本分布

Tab.1 Sample distribution of Grape_Complex dataset

after data augmentation		幅
病害类别	训练集图像数量	验证集图像数量
白腐病	1 314	110
褐斑病	1 146	144
霜霉病	1 324	83
健康	1 096	69
总计	4 880	406

1.3 病害识别总体流程

葡萄叶片病害识别的总体流程如图 3 所示。首先进行病害图像数据的收集, 并通过人工方式剔除无法使用的图像。其次对构建的数据集进行预处理

操作。然后按照比例 8:2 将数据集划分为训练集和测试集,并且通过随机旋转、水平或者垂直翻转等方式对原始数据集进行数据扩充,这样原有数据的多样性得以增加,使得最终训练得到模型的泛化能力

更强。最后用数据增强过的数据集训练 MANet 模型,保存训练过程中性能最佳的模型权重。基于以上训练的 MANet 模型用测试集中的图像进行测试,得到测试样本的预测类别,以实现葡萄叶片病害的识别。

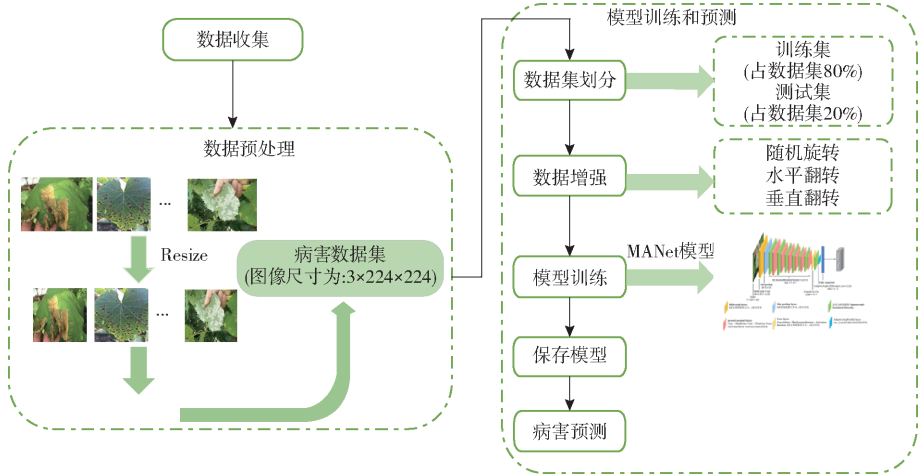


图 3 葡萄叶片病害识别的整体流程

Fig. 3 Overall recognition process of grape leaf diseases

1.4 MANet 模型

1.4.1 模型结构

本研究提出了一种对葡萄叶片病害具有更高识别准确率的 CNN 模型,将其命名为 MANet。MANet 模型的整体结构如图 4 所示。从图 4 可看出,MANet 模型主要由多尺度特征融合模块 (Multi-scale convolution, MConv)、最大池化层、SENet-InvertedResidual 模块 (由 SENet 模块和倒残差模块^[25]交替连接构成)、注意力机制模块、自适应平均池化层和全连接层组成。MANet 首先利用一个多

尺度特征融合 MConv 模块来提取病害多尺度的特征信息,获得图像的丰富特征表示。其次,连接一个最大池化层用于降低图像的尺寸,然后,在最大池化层后将 SENet 和倒残差模块交替连接,用少量参数增加网络的深度,从而提高模型的非线性表示能力。同时,在 MANet 中加入了注意力机制 SENet 模块,使其能够更加关注图像中葡萄叶片的病害区域,从而减少图像中复杂背景产生的噪声对识别效果的影响。自适应平均池化层主要起到降低特征图的空间维度的作用,并在一定程度上增加模型的感受野。

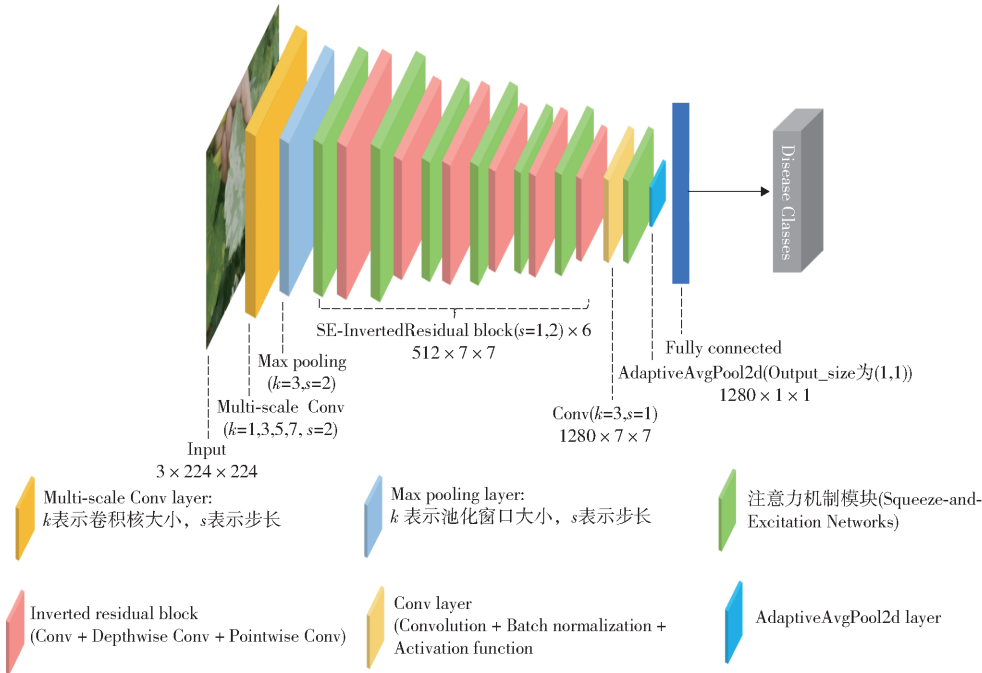


图 4 MANet 模型结构

Fig. 4 Overall structure of MANet model

最后通过一个分类器模块使用全连接层和 Softmax 函数来实现病害的分类。

1.4.2 多尺度特征融合模块

由于在自然状态下采集的葡萄叶片病害图像往往包含较多噪声,并且图像背景较为复杂,所以这使得病害识别难度更大。为此,本文提出了一种多尺度特征融合模块(Multi-scale convolution, MConv),在不同感受野和多尺度空间下尽可能多地提取病害的关键特征。MConv 模块结构如图 5 所示。在该模块中,输入特征图经过卷积核大小分别为 1×1 、 3×3 、 5×5 和 7×7 的卷积层进行并行处理,以获取病害图像更加丰富的特征信息。然后将各个支路卷积得到的特征图合并为一个新的特征图,并利用一个 1×1 的卷积对新的特征图进行特征融合。综上所述,本文提出的 MConv 模块可以捕获多尺度的病害信息,从而使得对于复杂背景的病害图像,获得更有辨别力的病害特征信息,有助于更加准确地识别葡萄叶片病害。

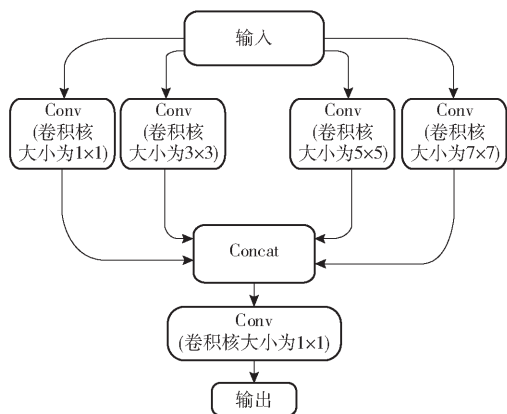


图 5 MConv 模块结构

Fig.5 Structure of MConv module

1.4.3 倒残差结构模块

本文提出的 MANet 中使用倒残差结构^[25]来减少模型的参数量,其主要亮点在于深度可分卷积。倒残差模块结构如图 6 所示。从图 6 中可以看出,倒残差模块主要由普通卷积、Depthwise convolution 和 Pointwise convolution 组成。当输入特征图进入倒

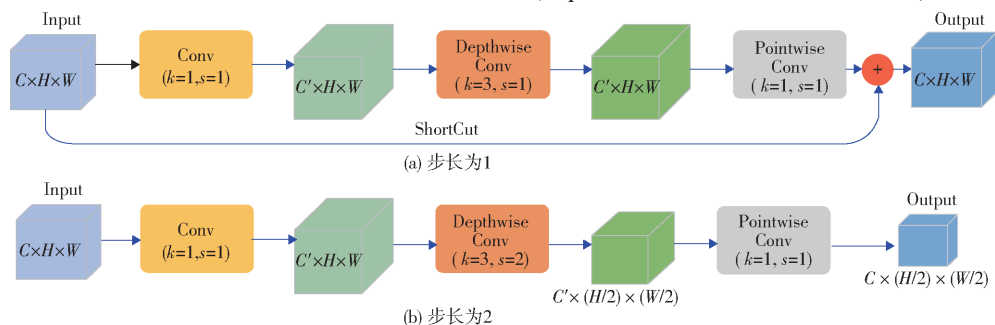


图 6 倒残差模块结构

Fig.6 Overall structure of inverted residual module

残差结构时,首先由卷积核大小为 1×1 的普通卷积进行升维操作,然后 Depthwise convolution 提取特征图的深度特征,并且特征图的通道数不会发生改变,最后由 Pointwise convolution 对特征图进行降维操作。同时当 Depthwise convolution 的步长为 1 时,即输入输出特征图的形状一致,则通过 ShortCut 支路将 Pointwise convolution 的输出特征图和输入特征图进行相加操作,得到倒残差模块最终的输出特征图,如图 6a 所示。当 Depthwise convolution 中的步长为 2 时,倒残差模块没有 ShortCut 支路,Depthwise convolution 在提取深度特征的同时进行下采样操作,扩大网络的感受野,最终倒残差模块输出特征图的空间维度 H 和 W 都将减半,如图 6b 所示。

与普通卷积相比,使用深度可分卷积可以大大减少模型的参数量和成本。普通卷积和深度可分卷积的卷积操作如图 7 所示。在图 7a 中假设输入特征图的宽和高都为 D ,通道数为 M ,卷积核大小为 $K \times K$,输出特征图的通道数为 N ,则普通卷积的参数量为 $K^2 MND^2$ 。同理在图 7b 中深度可分卷积的参数量为 $K^2 MD^2 + MND^2$ 。由此可计算深度可分卷积与普通卷积参数量的比值为

$$\frac{K^2 MD^2 + MND^2}{K^2 MND^2} = \frac{1}{N} + \frac{1}{K^2} \quad (1)$$

在本文提出的 MANet 模型中,Depthwise convolution 使用的卷积核为 3×3 ,将 $K = 3$ 代入式(1)中,可得深度可分卷积与普通卷积参数量的比值为 $\frac{1}{N} + \frac{1}{9}$ 。上述计算结果表明,理论上普通卷积的参数量是深度可分卷积参数量的 8~9 倍。

综上所述,本文在提出的模型中嵌入倒残差结构,这不仅大大减少了模型参数量和成本,而且在一定程度上增加了网络的深度,提高了模型的非线性表示能力。

1.4.4 Squeeze-and-Excitation Networks 模块

本文提出的 MANet 中嵌入了注意力机制模块(Squeeze-and-Excitation Networks, SENet),使其能够

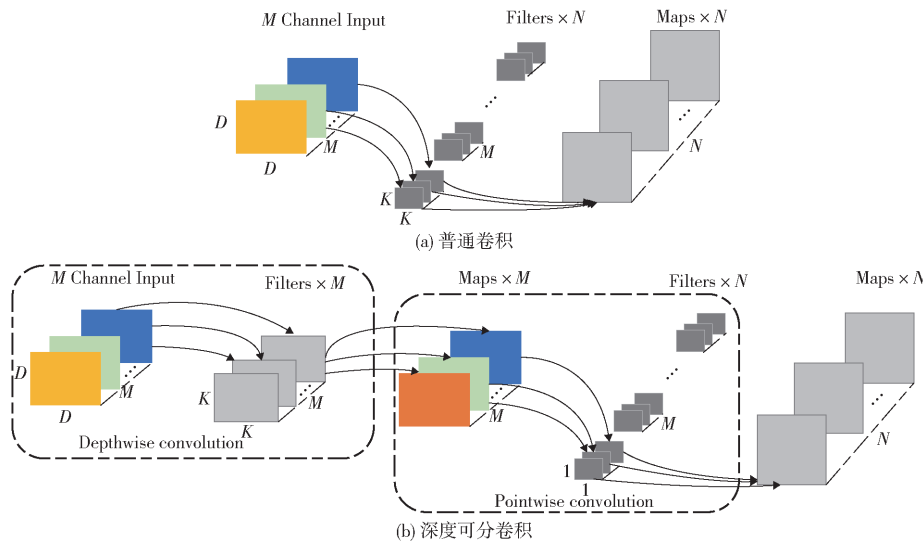


图 7 普通卷积和深度可分卷积的卷积过程比较

Fig. 7 Comparison of convolution process between ordinary convolution and depthwise separable convolution

关注图像中的葡萄叶片病害区域,并减少复杂背景对病害识别的影响。SENet 是一种通道注意力机制,通过网络学习得到不同通道特征的重要程度,并且按照不同的重要程度依次为不同的通道赋予不同的权重,通道的特征重要程度越高被赋予的权重越大,通道的特征重要程度越小被赋予的权重越小。SENet 模块主要分为压缩、激励和重构 3 个步骤,其中前两个是最为关键的步骤。SENet 模块的结构如图 8 所示。

在图 8 中 X 表示输入特征图, F_{tr} 表示特征图 X 到特征图 U 的变换过程,即 $F_{tr}: X \rightarrow U$, $X \in \mathbf{R}^{C_1 \times H_1 \times W_1}$, $U \in \mathbf{R}^{C \times H \times W}$, 其中 C_1 、 H_1 和 W_1 分别表示特征图 X 的通道数、高度和宽度;同理对于特征图 U , C 、 H 和 W 分别表示通道数、高度和宽度。卷积核为 $K = [k_1, k_2, \dots, k_c]$, 其中 k_1 表示第 1 个卷积核, k_c 表示第 c 个卷积核, 由此可得输出特征图为 $U = [u_1, u_2, \dots, u_c]$, 其中 u_c 的计算公式为

$$u_c = k_c X = \sum_{i=1}^{c_1} k_c^i x^i \tag{2}$$

式中 u_c ——第 c 个通道的特征图
 k_c ——第 c 个卷积核
 C_1 ——特征图 X 的通道数
 i ——计数变量
 k_c^i ——3 维卷积核
 x^i ——第 i 个特征图

式(2)中当 k_c^i 表示一个 3 维的卷积核、与原始特征图进行卷积操作时,学习到的通道特征是各个通道的累加值,并且与学习到的局部特征混合在一起。注意力机制 SENet 就是为了解决这种混合的通道特征关系,从而使得模型可以学习到各个通道的特征关系,将更为重要的通道信息突显出来。

在图 8 中, F_{sq} 表示 SENet 的压缩操作,它通过一个全局的平均池化层将每个通道上整个空间的信息压缩编码成一个全局特征。如图 8 所示,具有 C 个通道的特征图 U , 将其每个通道 $H \times W$ 的空间维度压缩为 1×1 的全局特征。 F_{sq} 操作结束后得到 shape 为 $C \times 1 \times 1$ 的特征图,具体的计算公式为

$$z_c = F_{sq}(u_c) = \frac{1}{HW} \sum_{i=1}^H \sum_{j=1}^W u_c(i, j) \tag{3}$$

式中 z_c ——第 c 个通道的特征图
 $F_{sq}(u_c)$ ——压缩操作
 i, j ——计数变量

在图 8 中 F_{ex} 表示 SENet 的激励操作,主要作用是学习到特征图各个通道之间的关系,并且根据每个通道不同的重要程度为其赋予不同的权重。 F_{ex} 操作的具体实现过程为将 F_{sq} 操作得到的所有全局特征传入两个全连接层,并且在两个全连接层后分别用 ReLU 和 Sigmoid 激活函数进行激活,从而使得模型学习到特征图每个通道的重要程度。 F_{sq} 操作的具体计算公式为

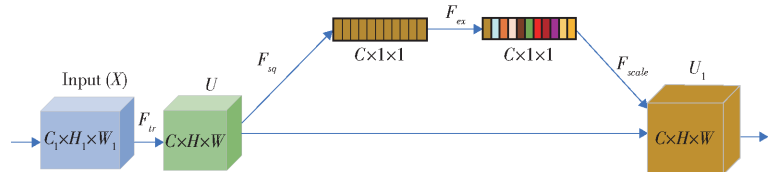


图 8 SENet 模块结构

Fig. 8 Structure of SENet module

$$s = F_{ex}(z, W) = \sigma(g(z, W)) = \sigma(W_2 \text{ReLU}(W_1 z))$$

(4)

式中 s ——权重 $W_1、W_2$ ——权重
 $F_{ex}(z, W)$ ——激励操作
 z ——特征图 σ ——激活函数

在式(4)中, $W_1 \in \mathbf{R}^{\frac{C}{r} \times C}$, $W_2 \in \mathbf{R}^{C \times \frac{C}{r}}$, 其中 r 表示通道的压缩比。在 F_{ex} 操作中为了更好地学习特征图各个通道间的相关性, 特征图在经过第 1 个全连接层时对神经元个数 C 除以 r 进行降维, 然后在经过第 2 个全连接层时再升维恢复到原来的神经元个数, 最终经过以上操作后得到 $C \times 1 \times 1$ 的特征图。在一定程度上来说, 这极大地减少了模型参数量和计算量, 并且增加了模型的非线性拟合效果。同时由于特征图通道之间往往具有一定的相关性, 因此通过网络学习到的关系也往往不是互斥的。为了解决这种现象, F_{ex} 操作在第 2 个全连接层后采用的是 Sigmoid 激活函数。

F_{scale} 表示重构操作, 即将 SENet 预测出的通道特征权重加权到特征图的每个通道空间上, 得出一个新的特征图 U_1 。 F_{scale} 的具体计算公式为

$$U_1 = F_{scale}(u_c, s_c) = s_c u_c$$

(5)

式中 $F_{scale}(u_c, s_c)$ ——重构操作
 s_c ——特征图中第 c 个通道的权重

在式(5)中 $F_{scale}(u_c, s_c)$ 表示在特征映射之间的通道运算, 其中 $u_c \in \mathbf{R}^{H \times W}$ 。综上所述, SENet 模块具有很强的移植性, 可以灵活地嵌入到各种图像识别的深度学习模型中, 并且在很大程度上提升了模型的性能。

2 实验结果与分析

2.1 实验平台

本研究实验平台的硬件环境为 Windows 10 企业版 64 位操作系统, 处理器为 Intel(R) Xeon(R) Silver 4210R, RAM 为 128 GB, 显卡为 NVIDIA GeForce RTX 3090, 显存为 24 GB。同时使用 Python 语言进行编程, 选用深度学习框架 Pytorch 1.7.1 进行模型的搭建和训练。

2.2 超参数设置

综合考虑模型的训练效果, 本文对模型超参数进行以下设置: 模型在训练和测试时采用批处理的方式, 每批数据包含 32 幅图像, 即批训练数据量的大小设置为 32。模型训练的迭代总次数 (Epoch) 设置为 100, 采用随机梯度下降 (Stochastic gradient descent, SGD) 优化器, 初始学习率设置为 0.01, 动量设置为 0.9, 同时在训练过程中, 学习率采用自适应动态调整的策略, 它会随着验证集的准确率而变

化, 损失函数使用交叉熵损失。最后, 为了避免模型出现过拟合现象, 在训练过程中使用 EarlyStopping 策略。

2.3 模型评价指标

本研究中, 使用平均识别准确率和 F1 值来综合评价模型的识别效果。同时为了进一步评价各个深度学习模型的性能, 在实验中还计算每个模型的参数量以及模型的浮点运算量 (Floating point operations, FLOPs)。模型参数量决定各个模型的尺寸, 也影响模型推断时对内存的消耗。模型浮点运算量用于评价各个模型的计算复杂性。

2.4 消融实验

在本实验中, 将不同注意力机制 SENet 模块和 CBAM^[40] (Convolutional block attention module, CBAM) 模块分别嵌入 MANet, 具体实验结果如表 2 所示。MANet_CBAM 表示在 MANet 中采用 CBAM 注意力机制, MANet_SENet 表示在 MANet 中采用 SENet 注意力机制。从表 2 中可以看出, 当向 MANet 中添加 SENet 注意力机制模块时, 模型表现效果更好, 平均识别准确率为 87.93%, 高于 MANet_CBAM 模型 0.74 个百分点, 所以本研究选择将 SENet 模块嵌入到 MANet 中来建立模型。同时为了验证 MANet 模型的 MConv 模块、倒残差模块和注意力机制 SENet 模块的有效性, 在自建 Grape_Complex 数据集展开了实验。消融实验结果如表 3 所示, MANet 的平均识别准确率比 Base、Base_SE、Base_MConv 分别高 5.17、3.45、0.5 个百分点。其中 Base 表示只有倒残差结构而没有 MConv 和 SENet 模块的基准模型, Base_SE 表示向基准模型加入 SENet 模块, Base_MConv 表示向基准模型加入 MConv 模块, MANet 为本文提出的新模型, 此时向基准模型中同时加入 SENet 和 MConv 模块。

表 2 不同注意力机制的实验结果

Tab.2 Experimental results of different attention mechanisms

模型	参数量	浮点运算量	平均准确率/%
MANet_CBAM	2.20×10^6	5.3×10^8	87.19
MANet_SENet	2.20×10^6	5.2×10^8	87.93

表 3 消融实验结果

Tab.3 Results of ablation experiments

模型	Inverted Residual	SENet	MConv	平均准确率/%
Base	✓	—	—	82.76
Base_SE	✓	✓	—	84.48
Base_MConv	✓	—	✓	87.43
MANet	✓	✓	✓	87.93

注: “✓”表示添加该模块, “—”表示不添加该模块。

从表 3 中可以发现,当向基准模型中分别单独引入 SENet 和 MConv 模块时,模型在训练集和测试集上平均识别准确率都得到了明显提升。这是因为当在模型中加入 SENet 时,模型在几乎不增加网络开销的情况下,可以更加专注于葡萄叶片的病害部分信息,抑制其他无关信息的干扰。当在模型中加入 MConv 模块时,模型可以提取葡萄叶片多尺度的病害信息,并进行融合,从而使模型的性能得到了进一步的提升。同时,用 Grad-CAM^[41] 对每个模型识

别病害的结果进行了可视化,如图 9 所示。可以发现图 9b、9c 可能关注了与病害无关的区域,同时也忽略了一些有用的信息。当加入 SENet 模块时,模型可以更加专注于病害区域,如图 9d 所示。与其他基准模型相比,所提出的 MANet 模型可以准确地关注病害区域,而很少关注不相关的复杂背景(图 9e),从而获得比其他模型更高的病害识别精度。综上所述,本文提出的 MANet 模型,当在基准模型中同时引入 SENet 和 MConv 模块时,模型性能表现最佳。

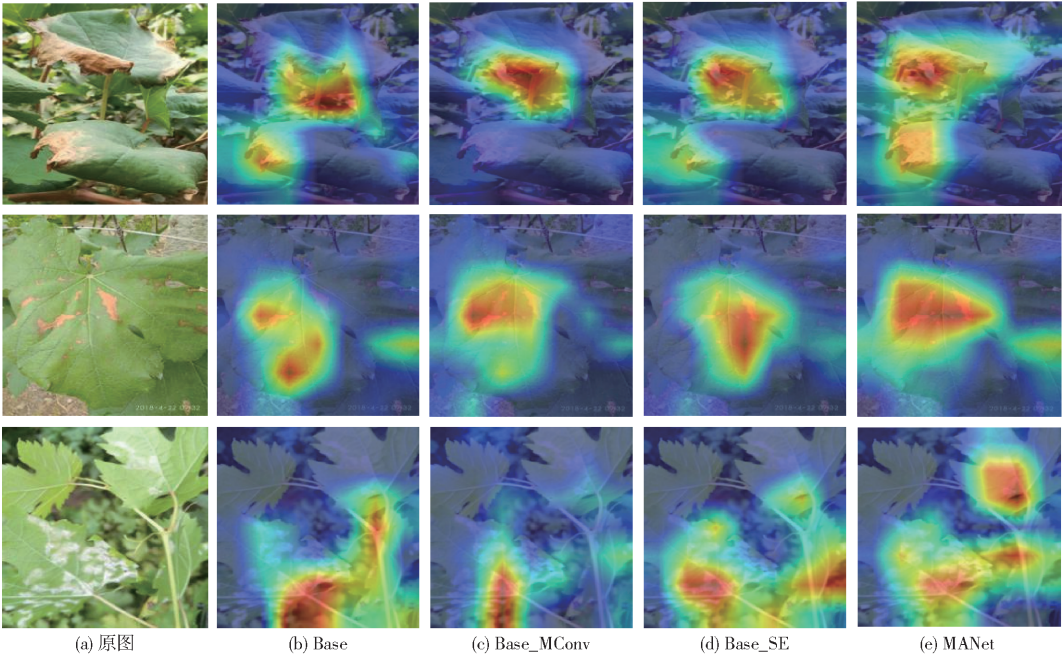


图 9 不同模型识别效果的可视化

Fig. 9 Visualization of recognition effect of different models

2.5 基于自建 Grape_Complex 数据集的模型评估

为了测试 MANet 模型在自然复杂环境下识别病害图像的效果,本实验选择了一些先进的模型作为对比,包括 AlexNet^[20]、VGG16^[21]、ResNet50^[22]、MobileNetV1^[24]、MobileNetV2^[25]、MobileNetV3^[42]、ShuffleNetV2^[27]、GhostNet^[28] 和 Swin Transformer^[43]。Grape_Complex 数据集的预处理、数据增强方法和数据集的分布与 1.2 节一致。不同算法的识别效果如表 4 所示。总体来看,通过与这些模型的比较,MANet 的识别表现最好,它达到了最高的平均识别准确率。MANet 的平均识别准确率和参数量分别为 87.93% 和 2.20×10^6 。除了 MobileNetV3 和 ShuffleNetV2 之外,MANet 的参数量比其他模型少。MANet 模型的缺点是浮点运算量为 5.2×10^8 ,略高于 MobileNetV2、MobileNetV3、ShuffleNetV2 和 GhostNet。然而,与 MobileNetV2、MobileNetV3、ShuffleNetV2 和 GhostNet 相比,MANet 的平均识别准确率分别高 4.68、3.45、10.10、5.17 个百分点。除了这4个模型外,MANet模型的浮点运算量最小,

表 4 不同模型基于 Grape_Complex 数据集的识别效果

Tab. 4 Recognition effects of different models based on Grape_Complex dataset

模型	参数量	浮点运算量	平均准确率/%
AlexNet	5.702×10^7	7.1×10^8	80.79
VGG16	1.3428×10^8	1.548×10^{10}	82.76
ResNet50	2.352×10^7	4.11×10^9	79.06
MobileNetV1	3.21×10^6	5.4×10^8	76.60
MobileNetV2	2.23×10^6	3.1×10^8	83.25
MobileNetV3	1.52×10^6	6.0×10^7	84.48
ShuffleNetV2	1.26×10^6	1.5×10^8	77.83
Swin Transformer	2.752×10^7	4.35×10^9	79.56
GhostNet	3.91×10^6	1.5×10^8	82.76
MANet	2.20×10^6	5.2×10^8	87.93

准确率最高。

不同模型对每种病害的 F1 值如表 5 所示。与其他模型相比,MANet 对每种病害都有相对较高的 F1 值。白腐病的病害特征较为明显,各个模型识别效果都比较好,但是褐斑病和霜霉病的识别效果要略微差一些,这是因为自建数据集是

自然状态下收集的图像,其背景较为复杂,这也增加了识别的难度。同时褐斑病和霜霉病在发病早期时,其病斑极其微小,即使人眼也很难辨别,所以模型容易将其误判为健康类别。而且褐斑病在发病晚期时,其枯叶病斑较大,模型也容易将其误判为白腐病。MobileNetV1 和 ShuffleNetV2 模型对葡萄叶片褐斑病和霜霉病的 F1 值相对较低,这可能是因为模型错误地将图像背景判断为识别目标,从而造成了干扰。MANet 对葡萄叶片褐斑病和霜霉病的 F1 值分别为 85.83% 和 85.87%, 高于 MobileNetV1 和 ShuffleNetV2。综上所述,与这些模型相比,虽然 MANet 的参数量和浮点运算量略高于一些轻量级的模型,但它取得了最高的平均识别准确率,模型识别效果也优于上述对比模型。

表 5 不同模型对每种病害识别的 F1 值

Tab.5 F1 values of different models for each disease

模型	白腐病	褐斑病	霜霉病	健康
AlexNet	87.84	77.31	79.46	77.87
VGG16	86.13	81.04	78.19	86.51
ResNet50	84.09	75.63	75.64	81.32
MobileNetV1	79.84	74.45	72.11	81.02
MobileNetV2	83.74	82.02	83.73	84.38
MobileNetV3	86.90	84.03	81.40	85.35
ShuffleNetV2	81.39	75.08	74.39	81.36
Swin Transformer	83.75	76.22	78.80	79.74
GhostNet	86.90	81.31	81.62	80.56
MANet	91.12	85.83	85.87	88.89

2.6 基于 PlantVillage 数据集的模型评估

为了进一步评估 MANet 的鲁棒性,本研究还在公开农作物病害数据集 PlantVillage 上对其进行了测试。与 2.5 节类似,选择一些模型进行对比实验,实验结果如表 6 所示。与这些模型相比,MANet 在测试集的平均识别准确率最高,达到 99.65%。实验结果表明,本文设计的 MANet 模型具有良好的鲁棒性,在公共农作物病害 PlantVillage 数据集上也表现出较好的识别效果。

表 6 不同模型基于 PlantVillage 数据集的识别效果

Tab.6 Recognition effects of different models based on PlantVillage dataset

模型	平均准确率
AlexNet	98.29
VGG16	97.81
ResNet50	99.18
MobileNetV1	98.69
MobileNetV2	99.52
MobileNetV3	99.40
ShuffleNetV2	99.31
Swin Transformer	98.86
GhostNet	99.24
MANet	99.65

3 结论

(1)提出了一种新的网络模型 MANet,其模型的参数量和浮点运算量相对较小,并且对葡萄叶片病害具有更高的识别准确率。在 MANet 中嵌入倒残差模块来构建网络,这极大降低了模型参数量和计算成本。同时,在模型中采用设计的 MConv 模块来提取病害图像多尺度的特征并进行融合,提高了模型对不同病害的识别准确率。此外,在 MANet 中嵌入注意力机制 SENet 模块,增强了模型在复杂背景的图像中提取关键病害特征的能力,并抑制其他无关信息,从而使得模型的平均识别准确率得到进一步的提高。

(2)为了验证该模型的有效性,在自建 Grape_Complex 数据集和 PlantVillage 数据集上进行了实验。实验结果表明,MANet 的平均识别准确率为 87.93%,明显高于经典 CNN 模型,如 ResNet50、AlexNet、VGG16,以及轻量级 CNN 模型,如 MobileNetV1、MobileNetV2、MobileNetV3、ShuffleNetV2 和 GhostNet 等。此外,为了进一步证明模型的鲁棒性,还在公开农作物病害数据集 PlantVillage 上进行了测试,实验结果表明,本文提出的模型在这个数据集上仍然取得了较好的识别效果。然而,与一些轻量级模型相比,MANet 的不足之处在于模型的浮点运算量略高。

参 考 文 献

[1] LIANG Q K, XIANG S, HU Y C, et al. (PDSE)-S-2-Net: computer-assisted plant disease diagnosis and severity estimation network[J]. Computers and Electronics in Agriculture, 2019, 157: 518-529.

[2] LI Z, PAUL R, TIS T B, et al. Non-invasive plant disease diagnostics enabled by smartphone-based fingerprinting of leaf volatiles[J]. Nature Plants, 2019, 5(8): 856-866.

[3] GANDHI R, NIMBALKAR S, YELAMANCHILI N, et al. Plant disease detection using CNNs and GANs as an augmentative approach[C]//2018 IEEE International Conference on Innovative Research and Development (ICIRD). IEEE, 2018: 1-5.

[4] HAN L X, HALEEM M S, TAYLOR M, et al. A novel computer vision-based approach to automatic detection and severity assessment of crop diseases[C]//Science and Information Conference (SAI), 2015:638-644.

- [5] MAJUMDAR D, KOLE D K, CHAKRABORTY A, et al. An integrated digital image analysis system for detection, recognition and diagnosis of disease in wheat leaves[C]//3rd International Symposium on Women in Computing and Informatics (WCI), 2015:400–405.
- [6] PANTAZI X E, MOSHOU D, TAMOURIDOU A A. Automated leaf disease detection in different crop species through image features analysis and one class classifiers[J]. Computers and Electronics in Agriculture, 2019, 156: 96–104.
- [7] PYDIPATI R, BURKS T F, LEE W S. Identification of citrus disease using color texture features and discriminant analysis[J]. Computers and Electronics in Agriculture, 2006, 52(1–2): 49–59.
- [8] ZHANG C L, ZHANG S W, YANG J C, et al. Apple leaf disease identification using genetic algorithm and correlation based feature selection method[J]. International Journal of Agricultural and Biological Engineering, 2017, 10(2): 74–83.
- [9] PRASAD S, PEDDOJU S K, GHOSH D. Multi-resolution mobile vision system for plant leaf disease diagnosis[J]. Signal, Image and Video Processing, 2016, 10(2): 379–388.
- [10] SHRIVASTAVA V K, PRADHAN M K. Rice plant disease classification using color features: a machine learning paradigm [J]. Journal of Plant Pathology, 2021, 103(1): 17–26.
- [11] GUO A T, HUANG W J, YE H C, et al. Identification of wheat yellow rust using spectral and texture features of hyperspectral images[J]. Remote Sensing, 2020, 12(9): 1419.
- [12] PU Y, YU C. Automatic identification of crop diseases image based on the dual coding genetic algorithm of support vector machine[J]. Guizhou Agricultural Sciences, 2013(7): 187–194.
- [13] CAMARGO A, SMITH J S. Image pattern classification for the identification of disease causing agents in plants [J]. Computers and Electronics in Agriculture, 2009, 66(2): 121–125.
- [14] 翟治芬, 徐哲, 周新群, 等. 基于朴素贝叶斯分类器的棉花盲椿象危害等级识别[J]. 农业工程学报, 2015, 31(1): 204–211. ZHAI Zhifen, XU Zhe, ZHOU Xinqun, et al. Recognition of hazard grade for cotton blind stinkbug based on Naive Bayesian classifier[J]. Transactions of the CSAE, 2015, 31(1): 204–211. (in Chinese)
- [15] VALIENTE-GONZALEZ J M, ANDREU-GARCIA G, POTTER P, et al. Automatic corn (*Zea mays*) kernel inspection system using novelty detection based on principal component analysis[J]. Biosystems Engineering, 2014, 117: 94–103.
- [16] SANDIKA B, AVIL S, SANAT S, et al. Random forest based classification of diseases in grapes from images captured in uncontrolled environments[C]//13th IEEE International Conference on Signal Processing (ICSP), 2016:1775–1780.
- [17] MAI X C, MENG M. Automatic lesion segmentation from rice leaf blast field images based on random forest [C]//IEEE International Conference on Real-time Computing and Robotics (IEEE RCAR), 2016:255–259.
- [18] XIAO T, LIU H Y, CHENG Y. Corn disease identification based on improved GBDT method [C]//6th International Conference on Information Science and Control Engineering (ICISCE), 2019:215–219.
- [19] 张燕, 田国英, 杨英茹, 等. 基于 SVM 的设施番茄早疫病在线识别方法研究[J]. 农业机械学报, 2021, 52(增刊): 125–133. ZHANG Yan, TIAN Guoying, YANG Yingru, et al. Online detection method of tomato early blight disease based on SVM[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(Supp.): 125–133. (in Chinese)
- [20] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. Imagenet classification with deep convolutional neural networks [J]. Advances in Neural Information Processing Systems, 2012, 25: 1097–1105.
- [21] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[J]. arXiv preprint arXiv: 1409.1556, 2014.
- [22] 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), 2016:770–778.
- [23] HUANG G, LIU Z, VAN DER MAATEN L, et al. Densely connected convolutional networks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017: 4700–4708.
- [24] HOWARD A G, ZHU M, CHEN B, et al. MobileNets: efficient convolutional neural networks for mobile vision applications [J]. arXiv preprint arXiv:1704.04861, 2017.
- [25] SANDLER M, HOWARD A, ZHU M L, et al. MobileNetV2: inverted residuals and linear bottlenecks[C]//31st IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2018:4510–4520.
- [26] ZHANG X, ZHOU X Y, LIN M X, et al. ShuffleNet: an extremely efficient convolutional neural network for mobile devices [Z]. 31st IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2018:6848–6856.
- [27] MA N N, ZHANG X Y, ZHENG H T, et al. ShuffleNet V2: practical guidelines for efficient CNN architecture design[C]//15th European Conference on Computer Vision (ECCV), 2018: 122–138.
- [28] HAN K, WANG Y H, TIAN Q, et al. GhostNet: more features from cheap operations[C]//IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020:1577–1586.
- [29] MOHANTY S P, HUGHES D P, SALATHE M. Using deep learning for image-based plant disease detection[J]. Frontiers in Plant Science, 2016, 7: 1419.
- [30] LU Y, YI S J, ZENG N Y, et al. Identification of rice diseases using deep convolutional neural networks [J]. Neurocomputing, 2017, 267: 378–384.
- [31] PRIYADHARSHINI R A, ARIVAZHAGAN S, ARUN M, et al. Maize leaf disease classification using deep convolutional neural networks[J]. Neural Computing and Applications, 2019, 31(12): 8887–8895.
- [32] ZHANG S W, ZHANG S B, ZHANG C L, et al. Cucumber leaf disease identification with global pooling dilated convolutional neural network[J]. Computers and Electronics in Agriculture, 2019, 162: 422–430.
- [33] CHEN J D, WANG W H, ZHANG D F, et al. Attention embedded lightweight network for maize disease recognition[J]. Plant Pathology, 2021, 70(3): 630–642.

- [34] 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.
LI Shuqin, CHEN Cong, ZHU Tong, et al. Plant leaf disease identification based on lightweight residual network[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(3): 243–250. (in Chinese)
- [35] HU J, SHEN L, SUN G, et al. Squeeze-and-Excitation networks[C]//31st IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2018:7132–7141.
- [36] HUGHES D, SALAATHÉ M. An open access repository of images on plant health to enable the development of mobile disease diagnostics[J]. arXiv preprint arXiv:1511.08060, 2015.
- [37] PULGAR F J, RIVERA A J, CHARTE F, et al. On the impact of imbalanced data in convolutional neural networks performance[C]//12th International Conference on Hybrid Artificial Intelligent Systems (HAIS), 2017: 220–232.
- [38] BUDA M, MAKI A, MAZUROWSKI M A. A systematic study of the class imbalance problem in convolutional neural networks[J]. Neural Networks, 2018, 106: 249–259.
- [39] WANG S J, LIU W, WU J, et al. Training deep neural networks on imbalanced data sets[C]//International Joint Conference on Neural Networks (IJCNN), 2016:4368–4374.
- [40] WOO S H, PARK J, LEE J Y, et al. CBAM: convolutional block attention module[C]//15th European Conference on Computer Vision (ECCV), 2018:3–19.
- [41] SELVARAJU R R, COGSWELL M, DAS A, et al. Grad-cam: visual explanations from deep networks via gradient-based localization[J]. International Journal of Computer Vision, 2020, 128(2): 336–359.
- [42] HOWARD A, SANDLER M, CHU G, et al. Searching for MobileNetV3[C]//IEEE/CVF International Conference on Computer Vision (ICCV), 2019:1314–1324.
- [43] LIU Z, LIN Y, CAO Y, et al. Swin transformer: hierarchical vision transformer using shifted windows[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021: 10012–10022.

(上接第 203 页)

- [19] LU J, EITEL J U, ENGELS M, et al. Improving unmanned aerial vehicle (UAV) remote sensing of rice plant potassium accumulation by fusing spectral and textural information[J]. International Journal of Applied Earth Observation and Geoinformation, 2021, 104: 102592.
- [20] ZHENG H, CHENG T, ZHOU M, et al. Improved estimation of rice aboveground biomass combining textural and spectral analysis of UAV imagery[J]. Precision Agriculture, 2019, 20: 611–629.
- [21] MA Y, MA L, ZHANG Q, et al. Cotton yield estimation based on vegetation indices and texture features derived from RGB image[J]. Frontiers in Plant Science, 2022, 13: 1983.
- [22] LIU Y, FENG H, YUE J, et al. Remote-sensing estimation of potato above-ground biomass based on spectral and spatial features extracted from high-definition digital camera images[J]. Computers and Electronics in Agriculture, 2022, 198: 107089.
- [23] FU Y, YANG G, LI Z, et al. Winter wheat nitrogen status estimation using UAV-based RGB imagery and Gaussian processes regression[J]. Remote Sensing, 2020, 12(22): 3778.
- [24] YUE J, YANG G, TIAN Q, et al. Estimate of winter-wheat above-ground biomass based on UAV ultrahigh-ground-resolution image textures and vegetation indices[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2019, 150: 226–244.
- [25] LIU Y, FENG H, YUE J, et al. Estimation of potato above-ground biomass based on unmanned aerial vehicle red-green-blue images with different texture features and crop height[J]. Frontiers in Plant Science, 2022, 13: 938216.
- [26] FAN Y, FENG H, JIN X, et al. Estimation of the nitrogen content of potato plants based on morphological parameters and visible light vegetation indices[J]. Frontiers in Plant Science, 2022, 13: 1012070.
- [27] HE J, ZHANG N, SU X, et al. Estimating leaf area index with a new vegetation index considering the influence of rice panicles[J]. Remote Sensing, 2019, 11(15): 1809.
- [28] DUBE T, MUTANGA O. Investigating the robustness of the new Landsat–8 operational land imager derived texture metrics in estimating plantation forest aboveground biomass in resource constrained areas[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2015, 108: 12–32.
- [29] YANG W H, PENG S, HUANG J, et al. Using leaf color charts to estimate leaf nitrogen status of rice[J]. Agronomy Journal, 2003, 95(1): 212–217.
- [30] ZHENG H, MA J, ZHOU M, et al. Enhancing the nitrogen signals of rice canopies across critical growth stages through the integration of textural and spectral information from unmanned aerial vehicle (UAV) multispectral imagery[J]. Remote Sensing, 2020, 12(6): 957.
- [31] 刘杨,黄珏,孙乾,等. 利用无人机数码影像估算马铃薯地上生物量[J]. 遥感学报, 2021, 25(9): 2004–2014.
LIU Yang, HUANG Jue, SUN Qian, et al. Estimation of plant height and above ground biomass of potato based on UAV digital image[J]. National Remote Sensing Bulletin, 2021, 25(9): 2004–2014. (in Chinese)