

DOI: 10.13718/j.cnki.xdzk.2024.06.017

曾姣艳, 林思涛, 谢亚君. 基于改进型 EfficientNet 模型的木薯病害识别方法研究 [J]. 西南大学学报(自然科学版), 2024, 46(6): 197-208.

基于改进型 EfficientNet 模型的木薯病害识别方法研究

曾姣艳^{1,2}, 林思涛¹, 谢亚君^{1,2}, 曾美艳³

1. 福州外语外贸学院 大数据学院, 福州 350202; 2. 数据科学与智能计算重点实验室, 福州 350202;

3. 郴州职业技术学院 商贸旅游学院, 湖南 郴州 423000

摘要: 为实现木薯病害图像的快速、准确识别, 提出一种基于 EfficientNet 模型的木薯病害识别方法。首先针对输入样本的分布不平衡问题, 通过 Mixup、CutMix 及 GridMask 这 3 种数据增强方法对数据进行增强, 数据增强后由 EfficientNet-B4 模型提取特征, 然后引入 warmup 结合余弦退火优化学习率防止模型在初期发生过拟合及后期收敛速度慢的情况。实验结果表明, 所采用模型相较于近年来主流的 VGG16 及 ResNet101 模型不仅参数量远小于两者, 在木薯病害图像分类上的表现也优于两者, 且其计算量更少, 模型精度更高, 训练速度更快, 符合实际应用的要求。EfficientNet 模型在木薯病害数据上的分类准确率可达 90%。

关 键 词: 木薯病害图像; 数据增强; EfficientNet 模型;

余弦退火

中图分类号: TP391

文献标志码: A

开放科学(资源服务)标识码(OSID):



文章编号: 1673-9868(2024)06-0197-12

Research on Cassava Disease Recognition Methods Based on Improved EfficientNet Model

ZENG Jiaoyan^{1,2}, LIN Sitao¹, XIE Yajun^{1,2}, ZENG Meiyuan³

1. Big Data Institute, Fuzhou University of International Studies and Trade, Fuzhou 350202, China;

2. Key Laboratory of Data Science and Intelligent Computing, Fuzhou 350202, China;

3. School of Commerce & Tourism, Chenzhou Vocational Technical College, Chenzhou Hunan 423000, China

Abstract: An EfficientNet model based cassava disease recognition method was proposed for fast and accu-

收稿日期: 2023-10-11

基金项目: 国家自然科学基金项目(12371378); 福建省自然科学基金项目(2022J01378).

作者简介: 曾姣艳, 硕士, 副教授, 主要从事数据挖掘和机器学习研究。

通信作者: 谢亚君, 教授。

rate recognition of cassava disease images. The method first addressed the imbalance in the distribution of the input samples, performed the data enhancement by three data enhancement methods of Mixup, Cut-Mix and GridMask. The feature was extracted by EfficientNet-B4 model after data enhancement, then warmup combined with cosine annealing learning rate was introduced to prevent the model from overfitting at early stage and slow convergence speed at late stage. The experimental results show that compared with the mainstream VGG16 and ResNet101 models in recent years, the model used in this paper not only has a much smaller number of parameters than the two models, but also has a better performance on classification of cassava leaf disease image than that of two models. It is less computationally intensive, with higher model accuracy and faster training, which meets the requirements of practical applications. EfficientNet model can achieve 90% classification accuracy on cassava leaf disease data.

Key words: cassava disease images; data enhancement; EfficientNet model; cosine annealing

木薯是热带地区的重要经济作物之一,由于其具有耐旱耐贫等特性,在世界范围内被广泛种植。木薯是一种重要的粮食作物,尤其在一些发展中国家是主要的食物来源之一,但它易受各种病害的侵袭^[1],其中细菌性枯萎病(CBB)、褐条病(CBSD)、花叶病(CMD)、绿螨病(CGM)等多种病害严重影响了木薯的质量和产量。木薯叶是进行光合作用及蒸腾作用的主要场所,叶片面积是影响木薯生长的一个重要因素,它直接影响木薯的产量。叶片面积还会影响木薯的耐旱性、耐寒性、抗病虫害能力及抗高温性能。因此,想要获得更好的木薯产量,就必须保证植株叶片面积合理。然而,木薯生长期面临各种病害的威胁,其中最严重的病害之一就是花叶病^[2],这种病害会导致木薯叶片产生凹陷、脱落,最终影响植株的生长和产量。木薯叶病害识别的意义不仅在于保障农业产量、维护食品安全、降低经济损失、维护生态平衡,同时也为科学的研究和技术创新提供了重要的基础。因此,研究木薯叶片病害的早期检测和预防方法至关重要。

国内外学者将深度学习技术引入农作物病害识别已有较长时间。近年来,通过采用深度学习技术自动识别农作物病害受到越来越多学者的关注。在国外, Picon 等^[3]利用移动设备获取多种作物图像数据集,提出 3 种不同的卷积神经网络架构实现了多种农作物病害识别。Xiao 等^[4]利用主成分分析和 BP 神经网络算法对稻瘟病进行识别,实验数据表明提出的方法能够快速准确地识别稻瘟病株。Fuentes 等^[5]提出了一种基于深度学习的多种病虫害检测方法。在国内,顾博等^[6]结合 SLIC 算法和 GrabCut 自动分割算法较好地分割出了玉米小斑病、大斑病和灰斑病。张善文等^[7]运用 LeNet 模型进行黄瓜病害识别,并对 1 200 幅黄瓜病害图像进行预处理,调整 RGB 颜色通道,实验结果表明其方法高于传统方法,达到了较高精度。方晨晨等^[8]提出一种基于深度 ResNet 网络的方法对番茄病害图像进行分类,该方法不仅减小了数据存储容量,同时提高了算法精度。熊梦园等^[9]提出一种 ResNet50 结合 CBAM 注意力机制模型的方法对玉米枯萎叶、锈病叶、灰斑病叶和健康叶进行精准检测,相比 ResNet50 模型准确率提升了 4.2 个百分点。宋玲等^[10]提出的 CDD 模型是一种基于改进 YOLOX 网络的木薯叶病害检测模型,对田间木薯叶病害具有更强的检测能力,在不增加参数量的情况下提高了检测分类的精确率。

上述方法在对病害图像进行识别时往往通过改变网络模型的深度、宽度和分辨率优化性能,可能会造成模型过拟合,从而导致模型识别效率降低。为此,本文以木薯细菌性枯萎病(CBB)、褐条病(CBSD)、花叶病(CMD)、绿螨病(CGM)这 4 种常见病害及健康叶片为研究对象,提出一种基于 EfficientNet 模型的木薯病害识别方法,以期实现木薯病害的快速、准确识别。

1 材料和方法

1.1 数据收集

本文数据来源于 Kaggle 官网的 Cassava 数据集, 是非洲种植户在田间使用不同分辨率及不同手机随机拍摄获取的, 包含 5 类木薯叶片图像, 共计 21 375 张图像, 其中细菌性枯萎病(CBB)1 087 张、褐条病(CBSD)2 187 张、花叶病(CMD)13 158 张、绿螨病(CGM)2 386 张及健康叶片 2 557 张, 其分布情况如图 1 所示.

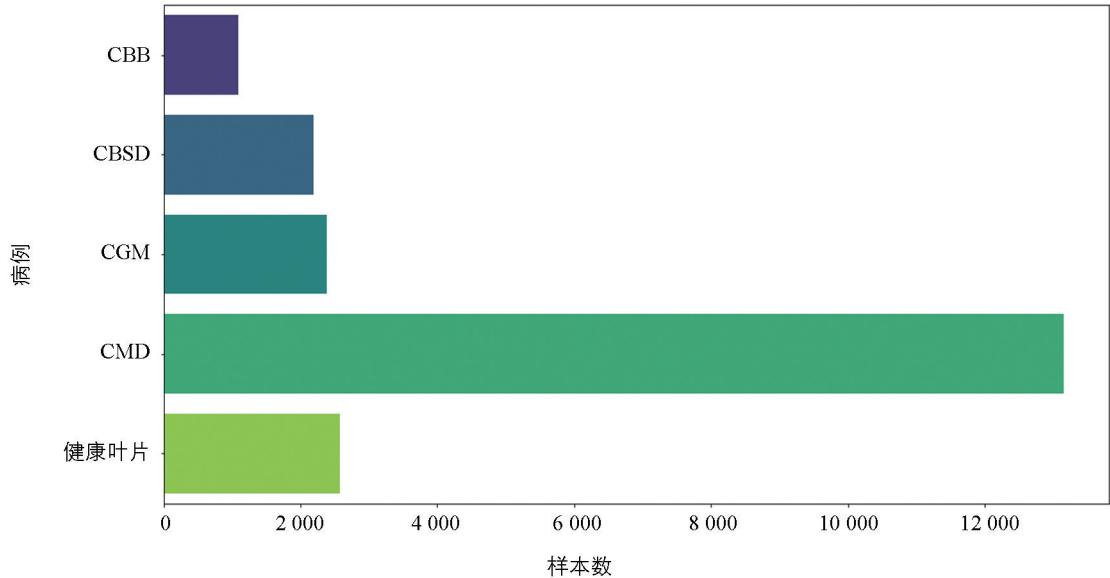


图 1 图像样本分布情况

1.2 数据增强

由图 1 可知, 样本存在严重的不平衡, 绿螨病(CMD)样本占据大部分, 因此需对其进行数据增强, 使模型具有良好的有效性和泛化能力.

1.2.1 Mixup 数据增强

Mixup 是 Zhang 等^[11]提出的一种对图像样本进行混淆的数据增强方法, 其将输入的图像与随机抽取的图像进行融合, 达到训练数据集的扩充. 计算公式为:

$$\tilde{x} = \lambda x_i + (1 - \lambda)x_j \quad (1)$$

$$\tilde{y} = \lambda y_i + (1 - \lambda)y_j \quad (2)$$

式(1)和式(2)中, (x_i, y_i) 和 (x_j, y_j) 是从同一个 batch 中随机抽取的两张图像样本及其对应的标签, λ 是从参数 α , β 的 beta 分布中随机采样的混合系数, $\lambda \in [0, 1]$. 本文使用 Mixup 技术将训练集中的两个图像样本及其相对应标签的线性插值作为扩充数据, 增强了图像样本之间的线性表达, 使模型能更准确地学习更多的信息, 从而提升模型的泛化能力和鲁棒性. Mixup 数据增强样本如图 2 所示, 其中 mean 表示图像像素值的平均值、std 表示图像像素值的标准差、min 表示图像中的最小像素值、max 表示图像中的最大像素值.

1.2.2 CutMix 数据增强

CutMix 是 Yun 等^[12]提出的一种数据增强方法, 具体步骤是通过在训练数据中随机裁剪, 并将一部分图像粘贴到另一图像相同位置来生成新的训练样本, 这样可以有效地提高其鲁棒性和泛化能力. 该方法可以有效地减轻由于样本分布不均、噪声等原因带来的训练偏差问题. 同时, 由于利用了不同样本之间的信息交叉, 也可以提高模型的泛化能力. Mixup 数据增强样本如图 3 所示.

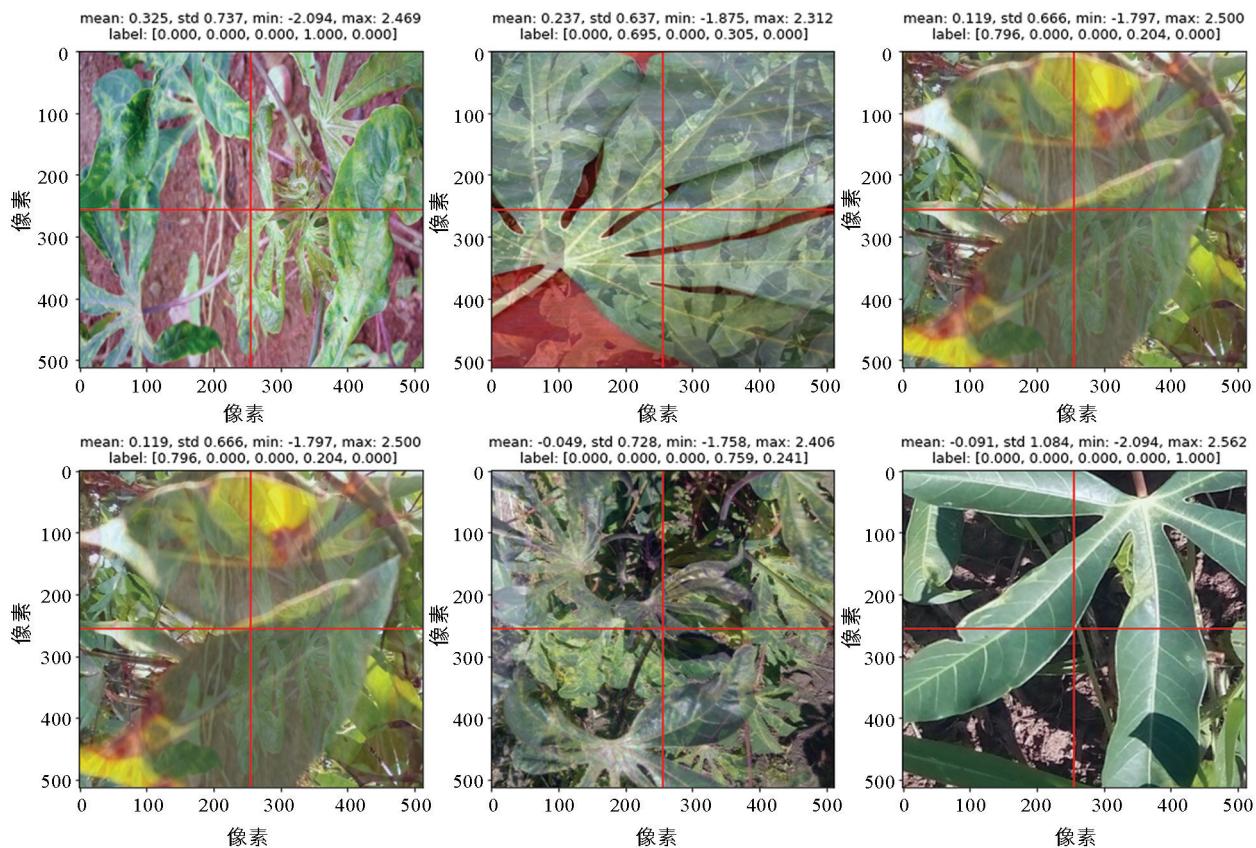


图 2 Mixup 处理结果

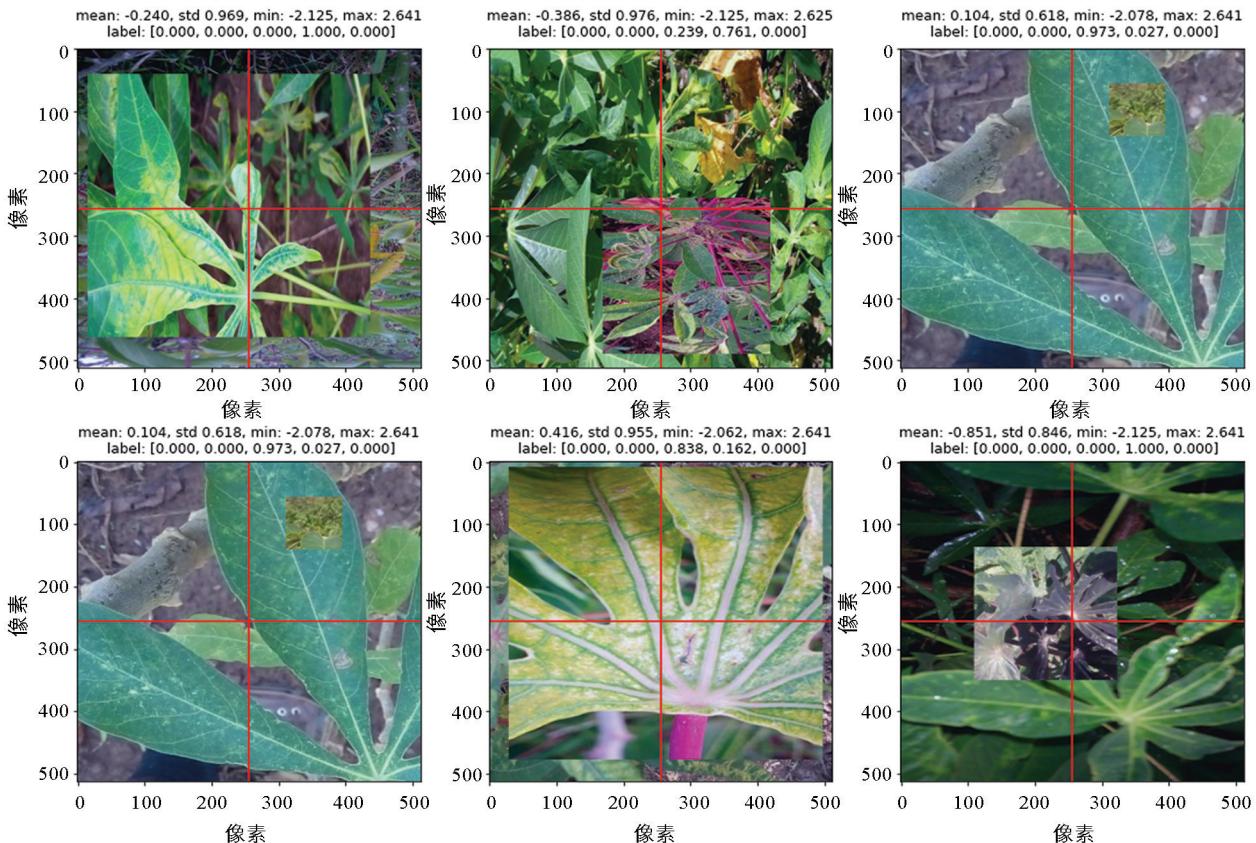


图 3 CutMix 处理结果

1.2.3 Gridmask 数据增强

Gridmask 是 Chen 等^[13]提出的一种性能优越的数据增强方法, 属于 Information Dropping 方法, 如何避免删除过度或保持区域连续是关键问题。首先, 过度删除将导致完整目标被删除或者上下文信息丢失, 使剩余区域无法有效地表达出目标信息, 因为这些区域会受到噪声干扰, 而这些噪声会影响目标的精确表达, 也就无法体现出目标的真实面貌。若保留过多的区域, 则会导致目标在这些区域内不受影响, 从而影响网络的鲁棒性。给定输入样本 x , 则 Gridmask 方法增强后的新样本为:

$$\tilde{x} = x \times M \quad (3)$$

式(3)中, M 为生成的二值掩膜。如图 4 所示, 采用 $(r, d, \delta_x, \delta_y)$ 4 个参数来确定一个唯一的 M 。每个掩膜都是由图 4 所示的单元排列平铺而成。 r 为每个单元短边灰色区域的比例, d 为一个单元的长度, δ_x 和 δ_y 为第一个完整的单元与样本边界的距离。

Gridmask 数据增强样本如图 5 所示。

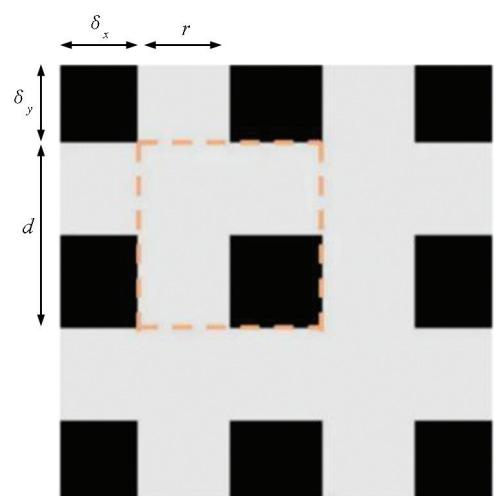


图 4 二值掩膜

Gridmask 处理结果如图 5 所示, 共展示了 6 张木薯叶片的病害识别样本。每张图都显示了原始图像、Gridmask 处理后的二值掩膜以及它们的组合结果。掩膜参数 (mean, std, min, max, label) 分别为:

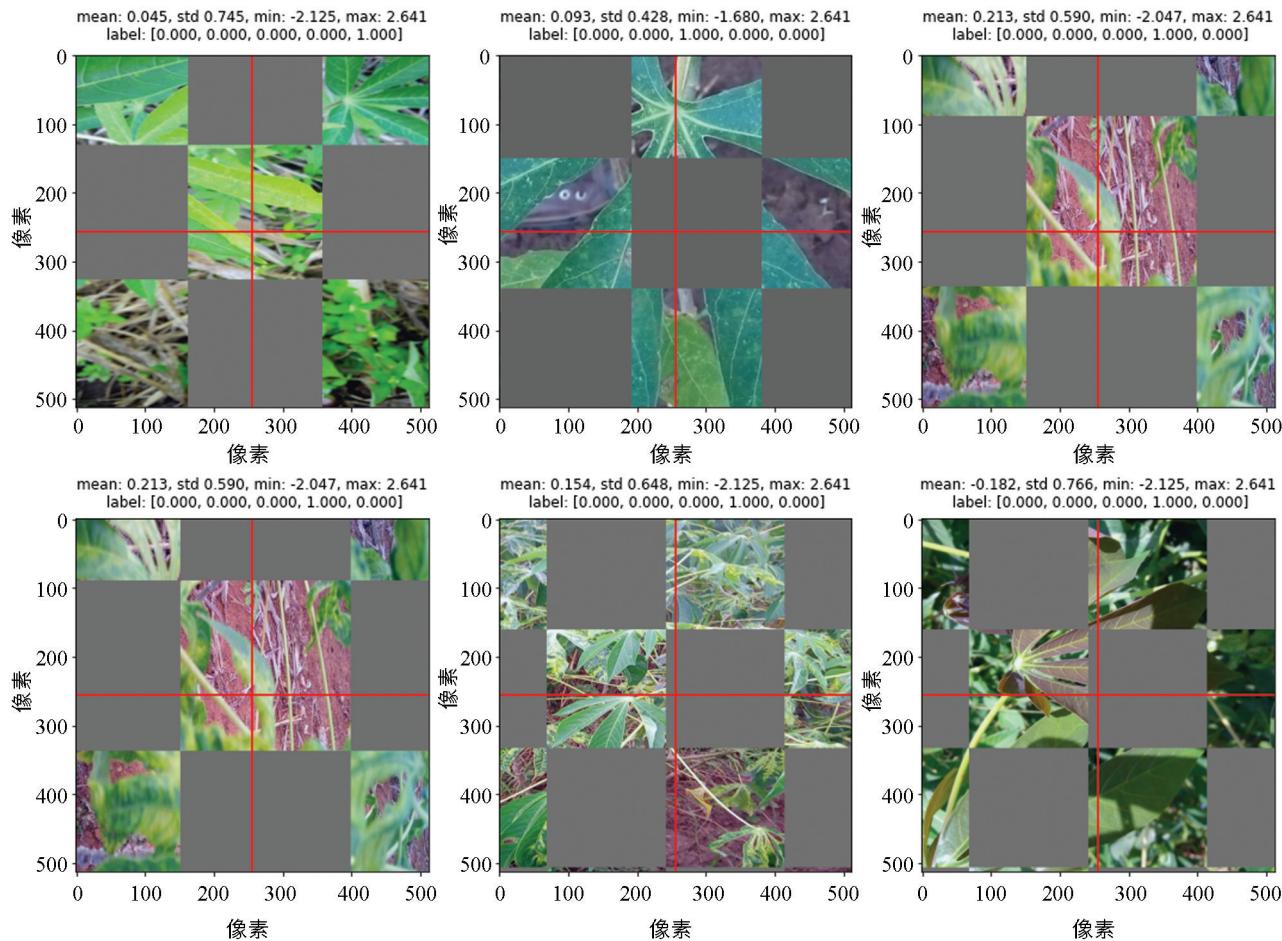


图 5 Gridmask 处理结果

1.3 病害识别模型

1.3.1 卷积神经网络的复合缩放算法

由于计算机性能的不断提升及对深度学习技术的深入研究, 卷积神经网络及其扩展网络得到了快速发

展, 它们在图像识别、语音识别、自然语言处理等方面都有显著的成效, 其中深度、宽度和基数都是影响卷积神经网络性能的主要因素^[14]. 过往的研究中有不少模型扩展的例子, 例如 ResNet^[15]可以通过增加网络层数, 从 ResNet-18 扩展到 ResNet-200; WideResNet^[16] 和 MobileNets^[17]则可以对模型宽度进行调整.

近年来, 学者们在研究中发现仅仅增加网络深度和宽度并不能完全提升模型的性能, 还需考虑其它因素, 比如网络结构、参数调整、训练集大小等. 因此, 在应用中需要综合考虑各种因素, 确保模型具有良好的性能. 本文应用的 EfficientNets^[18]模型, 其原理是基于模型缩放概念, 通过对模型深度、宽度和分辨率进行统一扩展, 实现了高精度和高效性的平衡. 该模型使用卷积神经网络中的普通卷积和深度可分离卷积, 减少整个模型中的计算量和参数数量. 为了进一步提高模型性能, 该模型还使用了一种新颖的复合因子方法, 通过精心挑选深度、宽度和分辨率缩放因子来扩展网络中的所有层. 这种方法通过将各个因素相互融合, 为模型提供不断增强的表示能力, 并在减少计算量和参数数量的同时, 达到超过其他卷积神经网络(CNN)模型的性能表现. 目前, EfficientNet 模型已经在许多图像分类和目标检测任务中取得了很好的效果.

EfficientNets 模型中使用复合系数 φ , 有原则地均匀缩放模型的分辨率、深度及宽度. 故有:

$$\begin{aligned} d &= \alpha^\varphi w = \beta^\varphi r = \gamma^\varphi \\ \text{s. t. } \alpha \cdot \beta^2 \cdot \gamma^2 &\approx 2 \\ \alpha &\geq 1, \beta \geq 1, \gamma \geq 1 \end{aligned} \quad (4)$$

式(4)中, d 表示深度, w 表示宽度, r 表示分辨率, α, β, γ 分别为深度、宽度和分辨率的系数. φ 作为指定的系数, 起到控制可用资源数量并缩放模型的作用; α, β, γ 则起着控制这些额外资源如何分配给模型深度、宽度和分辨率的作用.

1.3.2 EfficientNets 模型参数及结构

由于网络缩放不会影响 CNN 的卷积操作, 因此选择一个良好的基线模型也很重要. 本文使用 EfficientNets 模型的基线模型通过多目标神经网络架构搜索, 该方法由 MBconv 和 SENet 两部分构成. MBconv 通过对输入层的逐点卷积运算, 不断学习不同的信道维数, 从而更好地预测出输出端的信道特征. 同时, 为了进一步提高网络性能, SENet 注意机制模块被加入到 EfficientNet 模型中, 通过对输入数据进行注意力机制计算, 可以更好地获取网络训练过程中的关键信息, 从而有效提升模型的性能. 最后, 利用卷积操作将信道数量还原到初始信道数量, 完成该方法的完整流程. 以 EfficientNet-B0 为例, 其网络结构如表 1 所示, 其中 MBconv1 和 MBconv6 分别表示扩展比例为 1 和 6.

表 1 EfficientNet-B0 网络参数结构

阶段	卷积操作	分辨率	通道	层数
1	Conv 3×3	224×224	32	1
2	MBconv1, $k3 \times 3$	112×112	16	1
3	MBconv6, $k3 \times 3$	112×112	24	2
4	MBconv6, $k5 \times 5$	56×56	40	2
5	MBconv6, $k3 \times 3$	28×28	80	3
6	MBconv6, $k5 \times 5$	14×14	112	3
7	MBconv6, $k5 \times 5$	14×14	192	4
8	MBconv6, $k3 \times 3$	7×7	320	4
9	Conv $\times 1 \&$ pooling $\&$ FC	7×7	1 280	1

本文模型中采用交叉熵损失函数, 该函数是一种非参数化的损失函数, 可以有效地实现对木薯叶片病害分类结果中正确结果与分类结果之间误差的预测. 其数学表达式为:

$$H(p, q) = - \sum_{i=1}^n p(x) \ln q(x_i) \quad (5)$$

式(5)中, p 是正确的标签, 在进行模型训练时如果输入的样本和相应的标记都被设定好了, 则实际的概率分布 p 就被确定好了. q 表示模型预测的标签概率分布, n 表示类别的数目. 当模型预测的概率分布与真实标签一致时, 交叉熵损失函数的值为 0. 否则, 损失值会随着预测误差增大而增大. 本文的目标是通过训练使模型的预测尽可能接近正确的标签分布, 从而达到最小化交叉熵损失函数的目的. 这种方法被广泛应用于深度学习中, 尤其是在图像分类、自然语言处理和语音识别等任务中可以有效地减少训练数据集的数量, 并且可以有效地减少计算量和训练时间.

1.3.3 迁移学习

近年来, 随着机器学习技术的不断更新, 越来越多的相关算法被提出来, 其中最典型的就是深度学习算法^[19]. 深度学习是一种机器学习算法, 它将一个复杂的问题分解为多个较简单的子问题, 然后在每个子问题中通过学习多个参数来实现解决复杂问题的能力. 同时, 它可以处理具有高维度、复杂结构、多尺度和非线性等特点的数据, 并能够有效地处理噪声、模糊和不确定等复杂情况, 从而实现更加准确的预测. 此外, 深度学习算法还可以用于解决高维度、大规模的问题, 并且具有较高的鲁棒性, 但其需要基于大规模的训练数据. 训练样本数量级的大小与网络规模呈线性关系, 而算法需要基于大规模的数据集来理解样本的潜在规律, 且大规模数据集的样本收集及标注成本高昂, 因此迁移学习已经被广泛地用于各领域, 以解决机器学习中训练样本匮乏的核心问题, 其目的在于运用从一个任务或领域中学到的经验, 帮助另一个任务或领域中的学习.

迁移学习有以下两种常见的形式:

- 1) 基于特征的迁移学习: 将源任务中的特征应用到目标任务中, 从而提取目标任务的特征.
- 2) 基于模型的迁移学习: 将源任务中的模型应用到目标任务中, 让目标任务可以通过源任务得到的知识和经验进行更好的学习.

在实践中, 迁移学习可以有多种应用, 如自然语言处理、计算机视觉、推荐系统等, 其具有以下优点:

- 1) 加速学习和提高准确性: 通过将相关任务中学习到的知识应用到新任务中, 可以加速学习过程, 同时提高新任务的准确性.
- 2) 提高模型的鲁棒性: 在迁移学习的过程中, 源任务中已经学习到的知识可以让模型更好地抵御噪声和变异, 从而提高模型的鲁棒性.
- 3) 减少对大量数据的需求: 在进行新任务训练时, 可以利用旧模型已经学到的知识来弥补数据不足的问题.

迁移学习可以帮助我们更好地利用已有的数据和已经训练好的模型, 提高机器学习算法的效率和准确性, 从而更好地解决实际问题. 为了验证 EfficientNet 模型对木薯病害识别的效果, 本文采用基于模型的迁移学习策略, 选择特征提取 EfficientNet-B4 原始的预训练模型进行训练. 迁移学习过程如图 6 所示.

1.3.4 余弦退火学习率

在网络训练时, 学习率会随着训练产生变化. 在模型训练后期, 如果学习率过高会导致损失率振荡, 造成损失函数不稳定, 影响模型的收敛速度; 但如果学习率衰减过快, 会造成模型无法很好地收敛, 使模型变得复杂而难以训练. 由于网络在训练开始时, 对于训练的图像样本是完全未知的, 模型对样本像素信

息的理解分布均匀,因此预测模型训练初期很可能陷入过拟合状态。考虑到以上情形,本文运用学习率预热(warmup)结合余弦退火算法对学习率的衰减进行调整。由于模型训练初始时参数不稳定,且梯度较大,若此时初始学习率设置过大可能造成数值不稳定。使用学习率预热有利于缓解模型在训练初始阶段对 mini-batch 的提早过拟合现象,保持稳定分布,同时也有利于维持模型深层的稳定性。运用余弦退火算法衰减学习率的方法整体参考余弦函数的变化特点,即余弦函数中随着 x 的增加,余弦值的下降速度按照缓慢、快速、缓慢的方式进行变化。将这种下降模式与学习率衰减进行配合是一种十分有效的计算方式,可以让模型轻松跳出局部最优解。其定义如下:

$$\eta_t = \eta_{\min}^i + \frac{1}{2}(\eta_{\max}^i - \eta_{\min}^i) \left(1 + \cos\left(\frac{T_{\text{cur}}}{T_i} \pi\right) \right) \quad (6)$$

式(6)中, i 表示迭代的索引值, η_{\min}^i 和 η_{\max}^i 分别表示学习率的最小值、最大值,这两个变量控制了学习率的变化范围,使学习率在指定范围内衰减。 T_{cur} 表示当前迭代(epoch)的次数,但 T_{cur} 会在每批次(batch)训练后更新,而当前的迭代还未执行结束,因此 T_{cur} 可为小数。 T_i 表示第 i 次训练时总的迭代次数。在本文中,模型学习率的最小值和最大值分别设为 1×10^{-6} 和 2×10^{-4} ,初始学习率为 1×10^{-6} ,其变化如图 7 所示,初始学习率经过前 5 个迭代增加到最大值,然后按照余弦规律先缓慢下降,再加速下降,最后在第 30 个迭代降到 1×10^{-6} 。

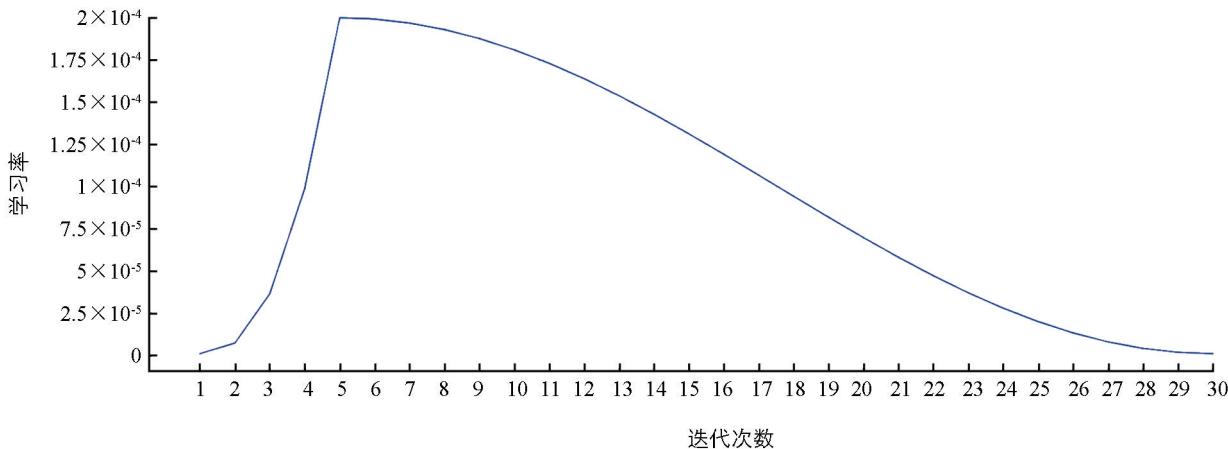


图 7 学习率变化曲线

2 模型训练

2.1 实验环境

实验主要在 Kaggle 提供的张量处理单元(TPU)环境中进行,可以极大地提升模型训练速度,Python 版本为 3.7.9,采用 Tensorflow 2.4.0 深度学习框架。

EfficientNet-B4 参数的选择对于模型效率和准确率具有重要影响,合理地选择参数可使模型不会过分地占用和消耗资源,从而使模型效率和准确率降低。本文采用的 EfficientNet-B4 运行参数配置如表 2 所

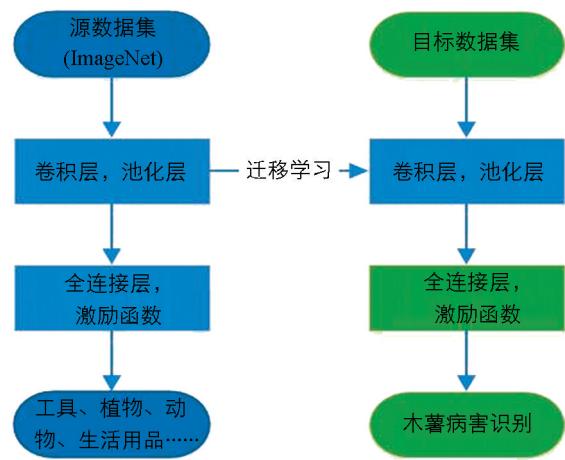


图 6 迁移学习过程

示. 模型迭代次数为 30, 批次大小为 16, 初始学习率为 1×10^{-6} , 利用余弦退火衰减方法优化学习率, 选用 Adam 作为优化器.

表 2 EfficientNet-B4 参数配置

层(类型)	输出形状	参数
Efficientnetb4	(None, 16, 16, 1792)	17 673 816
dropout	(None, 16, 16, 1792)	0
global_average_pooling2d	(None, 1792)	0
dropout_1	(None, 1792)	0
dense	(None, 5)	8 965

2.2 评价指标

2.2.1 准确率

在常用的模型评价指标中, 准确率的采用最为广泛. 在机器学习中, 其定义为预测准确的样本数占全部样本的百分比. 对于二分类模型, 准确率的数学定义为:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (7)$$

式(7)中, TP 表示实际正样本被预测为正样本的数量, FP 表示实际负样本被预测为正样本的数量; TN 表示实际正样本被预测为负样本的数量, FN 表示实际负样本被预测为负样本的数量.

在样本种类数量分布均匀的数据集中, 准确率能对网络的综合性能做出评价. 但是, 实际数据集不同种类样本数量经常极不平衡, 这在木薯病害图像数据中体现得十分明显. 木薯不同病害的患病概率往往不同, 造成采集到的图像样本数量差异较大. 采用准确率作为评价指标或将稀释模型运用于小样本学习中与真实结果会发生较大的偏差, 因此本文使用精确率及召回率作为评价网络性能的指标.

2.2.2 精确率和召回率

1) 以二分类模型为例, 精确率定义为全部预测结果为正样本的样本里, 预测正确所占的百分比, 其数学定义为:

$$Precision = \frac{TP}{TP + FP} \quad (8)$$

式(8)可理解为针对被预测的某类样本进行计算, 得到该类别中出现预测错误的概率.

2) 以二分类模型为例, 召回率定义为正样本中预测结果为正样本的占比, 其数学定义为:

$$Recall = \frac{TP}{TP + FN} \quad (9)$$

式(9)可理解为某类别样本被预测正确的比例, 适用一个数据集中对小样本的预测性能进行评价.

由精确率和召回率定义可知, 在分类模型中两者往往相互矛盾. 当召回率提高时, 精确率往往会降低, 反之亦然. 为了能够更好地综合评价网络性能, 本文引入 $F1-Score$ 评价指标, 其数学定义为:

$$F1-Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (10)$$

$F1-Score$ 结合了精确率和召回率, 可较好地反映网络的预测性能.

3 结果与分析

3.1 识别效果分析

将经过 MixUp、CutMix 和 Gridmask 增强的图像输入到 EfficientNet-B4 模型中, 经过 30 次迭代后的

训练曲线如图 8 所示。由图 8 可知, 模型迭代到 13 次时, 准确率和损失率同时趋于稳定, 模型达到最优。

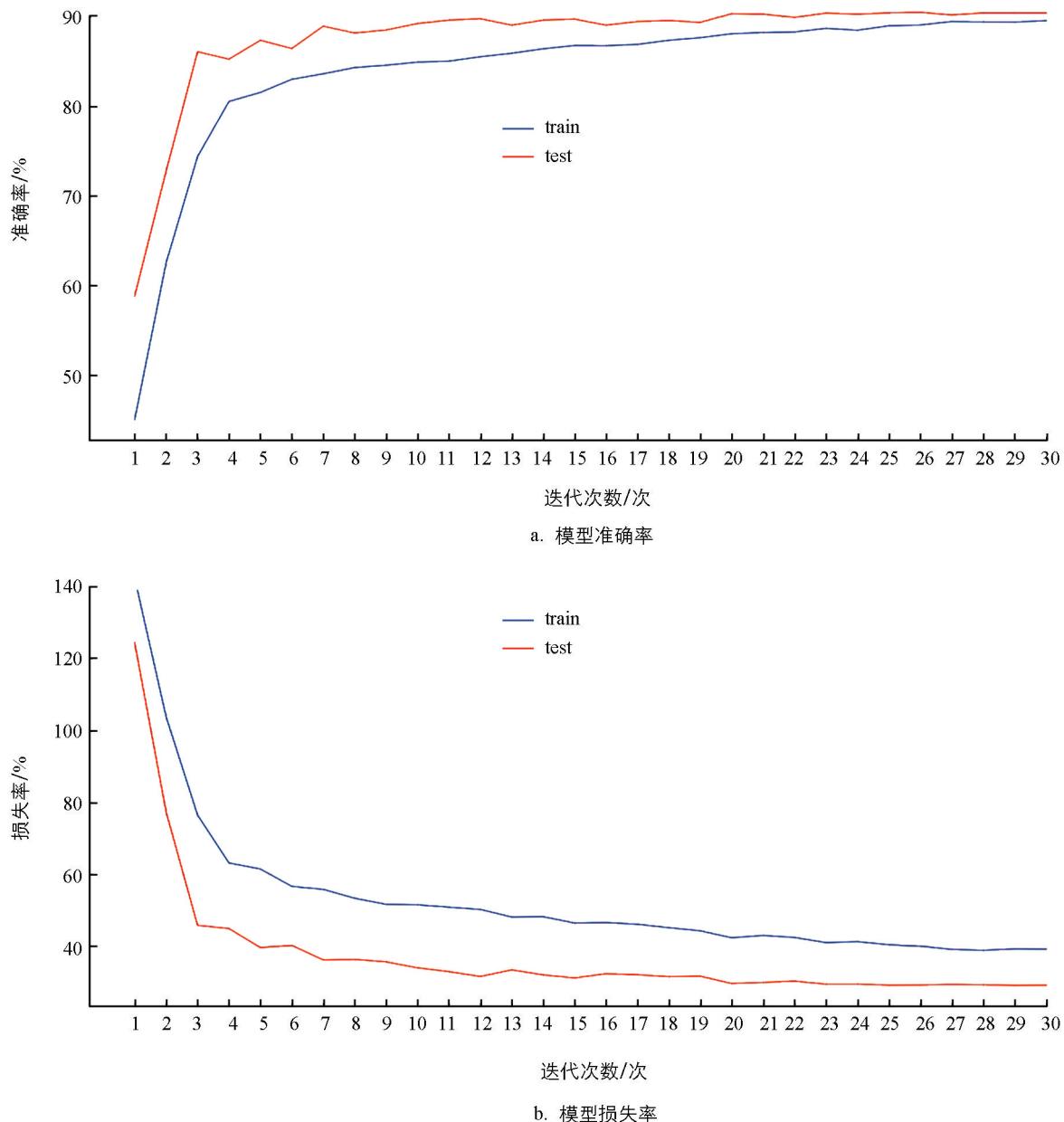


图 8 准确率曲线与损失率曲线

由表 3 和图 9 可知, 模型对花叶病(CMD)的预测准确率最高达到 96%; 模型识别错误主要在细菌性枯萎病(CBB)上, 该类别准确率仅为 69%, 主要原因在于数据集中样本在各类别分布不均衡, 造成模型对花叶病(CMD)学习的权重较多, 而对细菌性枯萎病(CBB)学习的权重较少。模型的平均 $F1$ -Score 达到 90%。

表 3 模型对各类样本的分类性能

病例	精确率/%	召回率/%	$F1$ -Score/%	样本量	平均 $F1$ -Score/%
CBB	69	74	71	246	90
CBSD	88	85	86	604	
CGM	85	81	83	526	
CMD	96	98	97	2591	
健康叶片	80	76	78	481	

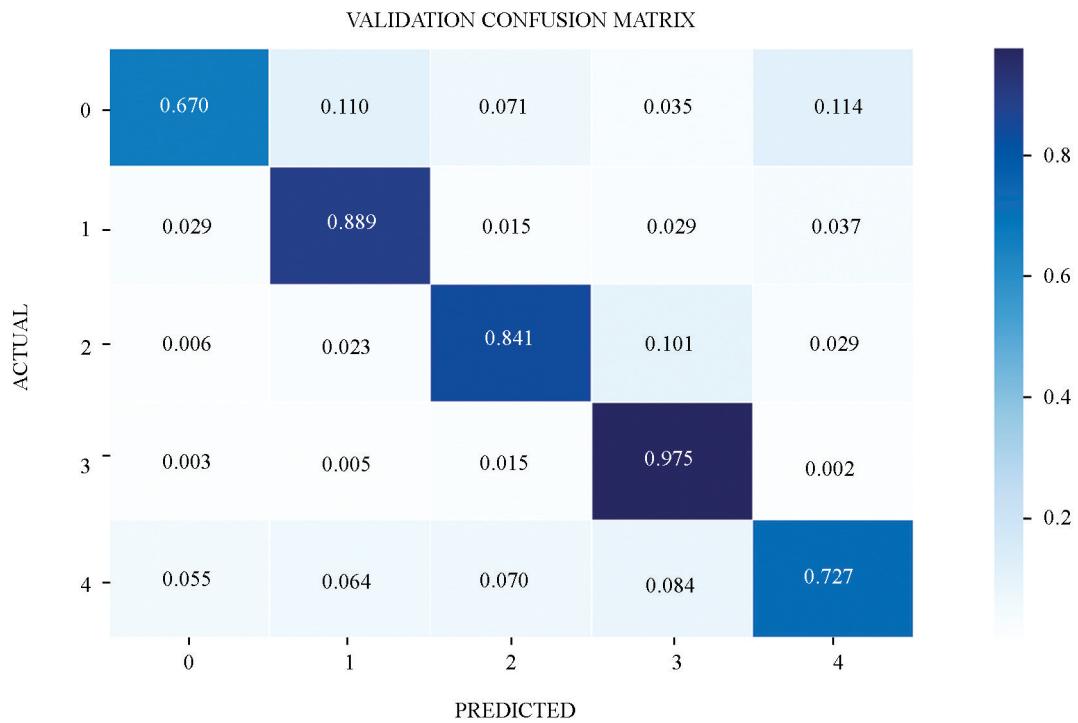


图 9 模型混淆矩阵

3.2 不同模型训练结果分析

为了检验本文提出的模型对木薯病害分类的性能, 选取 VGG16、ResNet-101 和 EfficientNet-B4 模型, 调用原始预训练模型结合交叉熵损失函数进行训练, 对比测试集和验证集的识别结果。由表 4 可知, EfficientNet-B4 模型不仅在测试集中表现最好, 在验证集上的准确率也最高, 其对噪声适应能力更强, 泛化能力更优越。对比 3 个模型的参数量可以发现, EfficientNet-B4 模型参数量相较于 VGG16 和 ResNet101 模型有更高的准确率。

表 4 不同模型分类准确率

模型	识别准确率/%		参数量
	测试集	验证集	
VGG16	79.64	82.33	1.38×10^8
ResNet101	83.77	84.44	4.45×10^7
EfficientNet-B4	89.51	90.70	1.9×10^7

4 结论

本文以木薯病害图像数据集为研究对象, 运用图像处理与深度学习技术提出一种基于 EfficientNet-B4 的木薯病害识别模型。为减小数据集中样本分布不均衡的影响, 增强模型的泛化能力, 本文选择 MixUp、CutMix 和 Gridmask 这 3 种数据增强技术, 并引入 WarmUp 结合余弦退火方法优化学习率, 防止模型出现训练初期陷入过拟合以及训练后期收敛慢的情况。与近年来流行的深度学习模型相比, 本文提出的模型具有参数量少、准确率高等优点。

参考文献:

- [1] 付海天, 郑华, 文峰, 等. 中国木薯研究及产业发展趋势 [J]. 农业研究与应用, 2022, 35(4): 9-22.

- [2] 刘琳玉, 赵平娟, 符艳等. 木薯花叶病毒 AC4 蛋白与 AtPARN 互作研究 [J/OL]. 热带作物学报, 2024, 45(1): 197-204.
- [3] PICON A, SEITZ M, ALVAREZ-GILA A, et al. Crop Conditional Convolutional Neural Networks for Massive Multi-Crop Plant Disease Classification over Cell Phone Acquired Images Taken on Real Field Conditions [J]. Computers and Electronics in Agriculture, 2019, 131(11): 356-366.
- [4] XIAO M H, MA Y, FENG Z X, et al. Rice Blast Recognition Based on Principal Component Analysis and Neural Network [J]. Computers and Electronics in Agriculture, 2018, 12(10): 482-490.
- [5] 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): 87-108.
- [6] 顾博, 邓蕾蕾, 李巍, 等. 基于 GrabCut 算法的玉米病害图像识别方法研究 [J]. 中国农机化学报, 2019, 40(11): 143-149.
- [7] 张善文, 谢泽奇, 张晴晴. 卷积神经网络在黄瓜叶部病害识别中的应用 [J]. 江苏农业学报, 2018, 34(1): 56-61.
- [8] 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(S1): 203-208.
- [9] 熊梦园, 詹炜, 桂连友, 等. 基于 ResNet 模型的玉米叶片病害检测与识别 [J]. 江苏农业科学, 2023, 51(8): 164-170.
- [10] 宋玲, 曹勉, 胡小春, 等. 基于 YOLOX 的复杂背景下木薯叶病害检测方法 [J]. 农业机械学报, 2023, 54(3): 301-307.
- [11] ZHANG H Y, CISSE M, DAUPHIN Y N, et al. Mixup: Beyond Empirical Risk Minimization [EB/OL]. (2018-04-27) [2023-04-01]. <https://arxiv.org/pdf/1710.09412v2.pdf>.
- [12] YUN S, HAN D, CHUN S, et al. CutMix: Regularization Strategy to Train Strong Classifiers with Localizable Features [C] // 2019 IEEE/CVF International Conference on Computer Vision (ICCV). Seoul: IEEE, 2019.
- [13] CHEN P G, LIU S, ZHAO H S, et al. GridMask Data Augmentation [EB/OL]. (2020-01-14) [2023-04-01]. <https://arxiv.org/pdf/2001.04086v2.pdf>.
- [14] XIE S N, GIRSHICK R, DOLLÁR P, et al. Aggregated Residual Transformations for Deep Neural Networks [C] // 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu: IEEE, 2017.
- [15] HE K M, ZHANG X Y, REN S Q, et al. Identity Mappings in Deep Residual Networks [M] // LEIBE B, MATAS J, SEBE N, et al. Computer Vision-ECCV 2016. Cham: Springer International Publishing, 2016.
- [16] ZAGOURUYKO S, KOMODAKIS N. Wide Residual Networks [EB/OL]. (2017-06-14) [2023-04-01]. <https://arxiv.org/pdf/1605.07146.pdf>.
- [17] HOWARD A G, ZHU M L, CHEN B, et al. MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications [EB/OL]. (2017-04-17) [2023-04-01]. <https://arxiv.org/pdf/1704.04861.pdf>.
- [18] Tan M, Le Q. Efficientnet: Rethinking Model Scaling for Convolutional Neural Networks [C]. Sacramento: International Conference on Machine Learning, 2019.
- [19] SARKAR C, GUPTA D, GUPTA U, et al. Leaf Disease Detection Using Machine Learning and Deep Learning: Review and Challenges [J]. Applied Soft Computing, 2023, 145: 110534.