

曾科, 王丽丽, 李玉辉, 等. 基于改进 MobileViT 的农作物病害识别研究[J]. 智能计算机与应用, 2025, 15(12): 137-143.
DOI: 10.20169/j.issn.2095-2163.24052203

基于改进 MobileViT 的农作物病害识别研究

曾科, 王丽丽, 李玉辉, 李程

(德州学院 计算机与信息学院, 山东 德州 253023)

摘要: 农业不仅关乎中国亿万百姓的生计,也是国家安全的基础产业。然而,农作物病害的频发已成为农业发展的一大障碍。为了快速且精准的识别农作物病害,减轻病害对农业生产的影响,本文以 MobileViT 为基础网络结构,提出了一种轻量级 HCA-MViT 网络模型。该模型将坐标注意力机制融入 MobileViT 网络中,以提升模型对作物病害特征的提取能力。同时,为了进一步优化网络性能并减少计算量,将 MobileViT 结构中的 ReLU6 激活函数替换为 Hardswish 激活函数,以确保信息的有效交互与管理。实验结果表明,提出的模型在性能上有了显著提升;该模型应用于 PlantVillage 数据集时达到了 99.6% 的准确率;应用在 PlantDoc 数据集时,识别准确率比 MobileViT 模型提升了 4.2%。此外,该模型权重文件体积仅为 4.1 MB,相较于其他现有网络模型,不仅识别准确率更高,而且更为轻量,占用的存储空间更少。

关键词: MobileViT; 农作物病害识别; 坐标注意力; Hardswish; 深度学习

中图分类号: TP391

文献标志码: A

文章编号: 2095-2163(2025)12-0137-07

Research on crop disease recognition based on improved mobilevit

ZENG Ke, WANG Lili, LI Yuhui, LI Cheng

(School of Computer and Information, Dezhou University, Dezhou 253023, Shandong, China)

Abstract: Agriculture is not only about the livelihood of hundreds of millions of Chinese people, it is also a basic industry for national security. However, the frequent occurrence of crop diseases has become a major obstacle to agricultural development. In order to quickly and accurately identify crop diseases and reduce the impact of diseases on agricultural production, this study proposes a lightweight HCA-MViT network model based on MobileViT network structure. This model integrates coordinate attention mechanism into MobileViT network to improve the model's ability to extract crop disease features. At the same time, in order to further optimize the network performance and reduce the amount of computation, the ReLU6 activation function in the MobileViT structure is replaced with the Hardswish activation function to ensure the effective interaction and management of information. The experimental results show that the performance of the proposed model has been significantly improved; the accuracy of the model reached 99.6% when applied to the PlantVillage dataset, and when tested on PlantDoc dataset, the recognition accuracy is higher than that of MobileViT model by 4.2%. In particular, the weight file size of the model is only 4.1 MB, which is not only higher recognition accuracy, but also lighter and occupies less storage space than other existing network models.

Key words: MobileViT; identification of crop diseases; coordinate attention; Hardswish; deep learning

0 引言

随着农业改革的不断深入,现代科技与农业的发展紧密相连。在农业发展进程中,生产过程中遭遇病害的情况越来越频繁,对农作物的生长造成了

非常严重的影响,进而加剧了农作物食品安全问题。因此,必须深入研究和防治农作物病害,减少不必要的损失。

在识别农作物病害时,传统的机器学习技术一般遵循 3 个核心阶段:首先是对图像进行初步处理

基金项目: 国家自然科学基金(11903008);德州学院省级大学生创新创业训练计划(S202310448017)。

作者简介: 曾科(2004—),男,本科生,主要研究方向:人工智能技术及应用;李玉辉(2004—),男,本科生,主要研究方向:深度学习;李程(2002—),男,本科生,主要研究方向:图像识别。

通信作者: 王丽丽(1981—),女,博士,教授,主要研究方向:大数据挖掘,机器学习。Email: jsjxwll@126.com。

收稿日期: 2024-05-22

以优化其质量;其次是通过一系列算法从图像中提炼关键信息;最后,利用这些信息对病害进行准确分类。在这一连续流程中,特征的提炼尤为关键,其直接影响到分类器的性能和识别的准确性。然而,农作物的病害图像呈现出背景复杂、病害特征不明显等问题。当面对具有复杂背景的病害图像数据时,传统的机器学习算法往往难以达到理想的效果,因为其在泛化方面的局限性较大。这种局限性导致了在从图像中抽取关键特征并执行精确分类时面临更多的挑战。

随着计算能力的提升和大数据时代的到来,深度学习快速发展起来,并应用于多个领域。深度学习技术可以自动从原始数据中提取特征,避免了手工特征提取的问题,极大地简化了模型设计的过程。此外,深度学习通过大量数据的训练提高模型的泛化能力,从而实现更准确的分类。在图像分类识别领域,卷积神经网络(CNN^[1])被认为是最有效的深度学习技术之一。在农作物病害检测的应用中,CNN同样展现了其出色的分类精确度。例如,肖天赐^[2]等通过对通道注意力机制进行改进,在 AI Challenger 2018 数据集上实现了 87.42% 的准确率;彭玉寒^[3]等通过对 MobileNetV2 模型的改良,在 PlantVillage 数据集上达到了 99.53% 的识别准确率;刘阳^[4]等通过调整 fire 模块参数改进 SqueezeNet,在 PlantVillage 数据集上获得了 98.13% 准确率;王泽钧^[5]等提出一种基于残差注意力机制和多尺度特征融合模块的轻量级模型,对 10 种农作物的叶片病害识别的准确率达 88.08%;李好^[6]等使用 ECA (Efficient Channel Attention) 注意力模块和深度可分离卷积对 ShuffleNetV2 模型进行改进,在 25 种叶片病害识别准确率达到 99.24%。虽然上述研究的准确率普遍较高,但数据集均为较简单的农作物病害图像,鲁棒性较差,并且应用的场景存在一定限制,难以完全适应农业生产环境的复杂性。对于较为复杂的农作物病害图像,SHILL^[7]等使用 YOLOv4 在对 13 种作物进行了检测,最优准确率达到 55.45%。王东方^[8]等提出了一种基于迁移学习的农作物病害分类模型 TL-SE-ResNeXt-101,准确率为 47.37%。这些研究遭遇的挑战主要包括模型的准确度尚未满足预期标准、处理数据所需的计算资源过于庞大、模型权重文件的存储需求过高,以及在移动设备上实施模型时遇到的困难。

MobileViT^[9]是苹果公司在 2021 年提出的一种 CNN 与 Vision Transformer 相结合的混合架构模型。

通过融合轻量级 CNN 和 Vision Transformer 的优势,本网络模型创新性地提出了一种既节省资源又通用、处理迅速,并且便于在移动设备上部署的网络解决方案。在现有研究成果的基础上,本文提出了一种名为 HCA-MViT (Hardswish and Coordinate Attention & MobileViT) 的新型网络模型。该模型采用 MobileViT 模型作为基础架构,通过将图像分割成块、将这些块展平成序列、注意力机制分配资源、应用 Transformer 编码、执行分类识别等一系列步骤,有效地处理了复杂背景中的噪声问题,从而提高了模型的准确性。

1 数据集构建及预处理

1.1 农作物病害数据集

本文用到的农作物病害图像来自于 PlantVillage^[10]和 PlantDoc^[11]两个公共数据集,在这些数据集上完成模型的训练和测试。

(1) 公共数据集 PlantVillage: 简单背景图像。该数据集包含 14 种农作物,共 38 个不同类别的 54 304 张叶片图像,这些图像均在实验室条件下采集,具有统一的、简单的背景。叶片摆放得整齐有序,全部为 RGB 彩色图像,涵盖了如西红柿、土豆、葡萄等多种农作物。所有图像的分辨率均为 256 像素×256 像素。图 1 展示了 PlantVillage 数据集的部分样本。

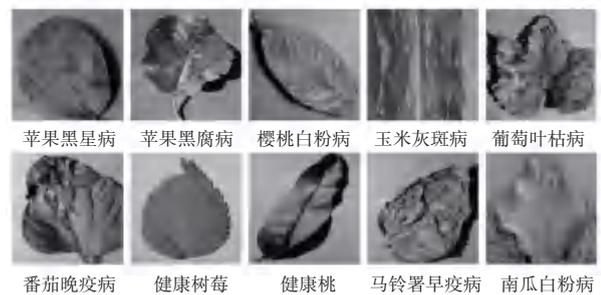


图 1 PlantVillage 数据集部分图像

Fig. 1 Partial images of PlantVillage dataset

(2) PlantDoc 数据集: 复杂背景图像。此数据集由印度理工学院的研究者们于 2019 年秋季开发,旨在支持图像分类与目标识别任务。经过 300 多个小时的辛勤工作,完成了数据的搜集与标注。数据集包含 13 个植物物种的图像,分为 27 个类别,具体为 17 种病理类别和 10 种表示健康状态的类别,图像总数达到 2 598 张。图片的大小、格式不完全相同。该数据集的图像背景复杂,更符合实际农业生产的真实情况。图 2 展示了该数据的部分样本。



图 2 PlantDoc 数据集部分图像

Fig. 2 Partial images of PlantDoc dataset

1.2 数据预处理

图像在输入模型之前需要的预处理操作主要包括调整图像分辨率和数据增强等。为了提高模型的学习效果和泛化能力,将每张图像的分辨率调整为 224 像素×224 像素。为了提升模型的泛化性能,对训练集进行数据增强是一种常用且有效的方法。在本文中,采用垂直翻转、水平翻转、高斯模糊和亮度调整等方法来扩充数据集。以一张番茄晚疫病图像为例,其数据增强后的效果示例如图 3 所示,具体的数据增强参数设置见表 1。

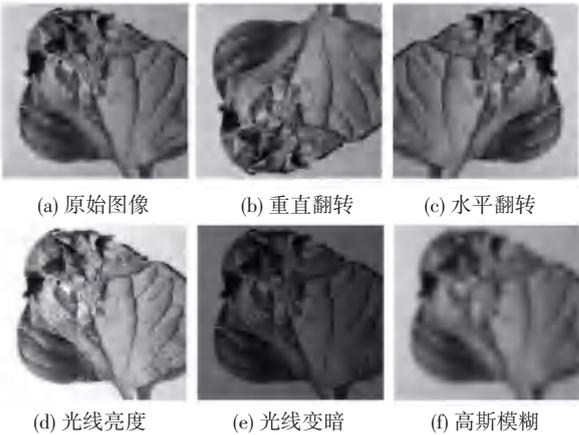


图 3 数据增强示例图

Fig. 3 Example of data augmentation

表 1 数据增强参数设置说明

Table 1 Explanation of data enhancement parameter setting

序号	增强方法	使用说明
1	垂直翻转	触发概率为 50%
2	水平翻转	触发概率为 50%
3	光线变暗	触发概率为 30%,亮度偏移度在 [0.5, 1] 区间随机产生
4	光线变亮	触发概率为 30%,亮度偏移度在 [1, 1.5] 区间随机产生
5	高斯模糊	触发概率为 30%,模糊程度在 [1, 5] 区间随机产生

2 农作物病害识别模型设计

卷积神经网络在提取图像的局部特征上具有强大的性能,但在处理全局特征时通常效果不佳。与此相对, Vision Transformer 模型在全局特征捕捉上更为出色,但其模型参数规模较大,这增加了对计算资源的需求,并影响其在移动设备上的部署效率。 MobileViT 将轻量级的卷积神经网络 MobileNet 和 Vision Transformer 结合,构建了一个轻量且高效的视觉网络模型,平衡了局部特征和全局特征的提取能力,既确保了精度,又具备出色的实时性能。在本文中,提出了一种名为 HCA-MViT 的分类模型,可轻松部署在移动设备上,以更快的速度准确分类作物病害图像。该模型以 MobileViT 为主干网络,为农作物病害的精准识别提供了可靠支持。通过充分利用 MobileViT 卓越的特征提取能力,本文能够准确地检测和分类各种农作物病害。 HCA-MViT 的结构如图 4 所示。

该网络结构的主要组成部分包括普通卷积层、 MobileViT 模块、改进的 MobileNetV2 模块、全局池化层以及全连接层。其中 MV2 表示改进的 MobileNetV2 模块, 2×表示执行该模块两次, ↓2 表示执行 stride=2 的下采样操作, $h = w = 2$ 表示将每个 batch 划分成大小为 2×2, 而 $L = 2$ 则表示 MobileViT 模块中的 Transformer 模块连续执行两次。

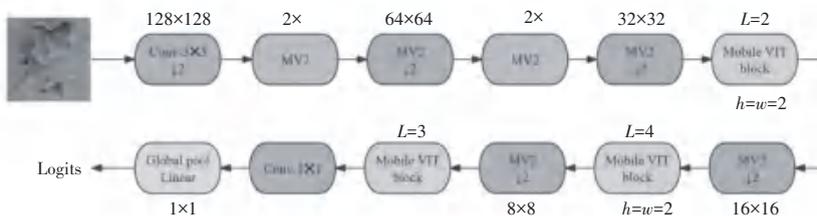


图 4 HCA-MViT 结构图

Fig. 4 HCA-MViT structure diagram

2.1 MobileViT Block

MobileViT Block 块包括 3 个子模块:局部特征编码、全局特征编码和特征融合模块。该模块的结构如图 5 所示。每个子模块均负责提取局部或全局

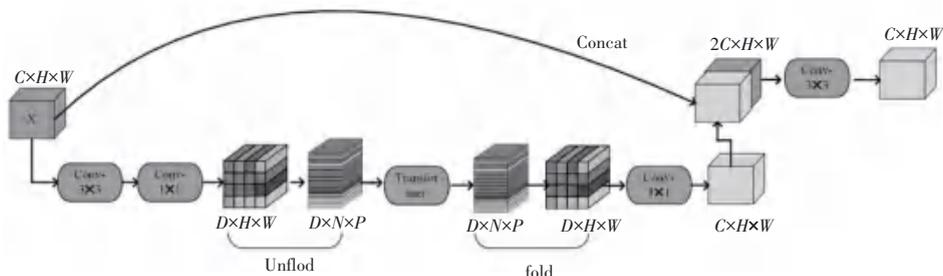


图 5 MobileViT Block 结构图

Fig. 5 MobileViT Block structure diagram

MobileViT Block 模块接收输入图像特征 $X \in R^H \times W \times C$, 其中 R 表示实数集合, X 是一个三维张量, 具有高度 H 、宽度 W 和通道数 C 。首先, 通过局部特征编码模块进行局部特征提取, 将特征张量投影到高维空间 $X_L \in R^{H \times W \times D}$ (其中 $D > C$)。接着, 该张量进入全局特征编码模块, 将局部特征信息通过 Unfold 操作, 再输入到 Transformer 中, 用于进行全局信息编码。在这个过程中, 注意力机制应用于计算列间的像素, 得到 $X_C \in R^{P \times N \times D}$, 如下式:

$$X_C(p) = \text{Transformer}(X_L(p)), 1 \leq p \leq P \quad (1)$$

经过 Transformer 模块后, 特征张量再通过 Fold 操作获得 $X_F \in R^{H \times W \times D}$, 与 X_L 具有相同的维度。这样, Transformer 输出中的每个像素都包含来自输入特征图中所有像素的信息, 从而将感受域扩展到 $H \times W$ 。接下来, 模型通过调整通道数, 将输入特征图的尺寸调整回与原始输入 X 相同的维度。引入捷径分支, 并将其与原始输入特征图 X 进行 Concat 操作, 得到融合后的特征图。随后, 利用卷积核大小为 $n \times n$ (这里 $n = 3$) 的卷积层对融合后的特征图进行进一步的融合, 得到输出特征图 Y 。这一系列操作有助于保留重要的特征信息, 并使网络更有效地学习和表达复杂的图像特征。

MobileViT Block 块通过增强模型对全局和局部信息的感知, 相较于传统卷积模块, 提高了特征提取的效率。这一改进是通过引入卷积式操作实现的, 使得 Transformer 能够获取位置信息, 从而减少变换器模块的需求, 更有效地学习更多信息, 实现轻量级模型。这种结构的优势在于它为处理图像特征提取任务提供了一种有效的方法, 同时保持了计算效率和性能。

特征信息, 并将这些特征进行融合。这种设计使得 HCA-MViT 模型能够高效地从图像中提取特征信息, 并且由于参数较少, 适用于计算资源有限的场景。

2.2 改进的 MobileNetV2 结构

MobileNetV2 模块采用了反向残差结构, 通过增减特征映射通道数量以及使用快捷连接来降低计算成本, 并同时保留足够的特征信息。这种设计使得 MobileNetV2 模块在维持高检测精度的同时减轻计算负担。尽管 MobileNetV2 架构通过整合输入特征与全局特征, 成功提高了图像分类的精确度, 但在将其直接应用于识别农作物病害图像时, 由于病害种类的多样性、高相似性以及复杂的背景等因素, 该架构仍存在一些局限性。为了进一步提升性能, 本文引入了 CA (Coordinate Attention) 注意力机制^[12] 到 MobileViT 网络的 MobileNetV2 模块中, 改进的 MobileNetV2 结构如图 6 所示, 在 MobileNetV2 模型中引入了坐标注意力机制。

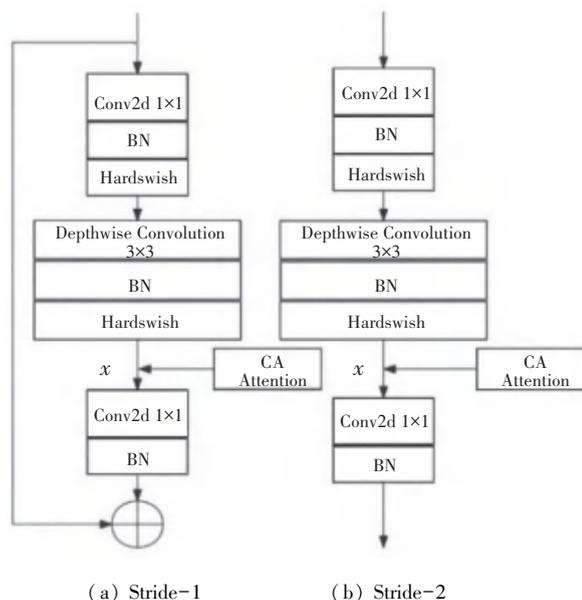


图 6 改进的 MobileNetV2 结构图

Fig. 6 Improved MobileNetV2 architecture diagram

坐标注意力是一种具有轻量级属性的注意力方法,其能有效地捕获位置信息和通道信息的关系,同

时避免显著的计算开销。本文的坐标注意力机制的结构图如图 7 所示。

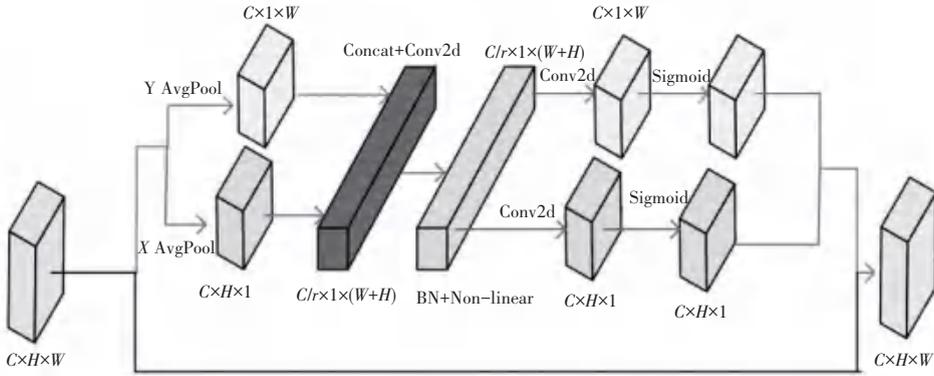


图 7 坐标注意力机制结构图

Fig. 7 Structure diagram of coordinate attention mechanism

坐标注意力模块可以取任意中间特征张量 $X = [x_1, x_2, \dots, x_c] \in R^{C \times H \times W}$ 作为输入进行变换和调整,得到与输入相同大小的具有增强表示的变换张量的输出张量 $Y = [y_1, y_2, \dots, y_c]$ 。该结构中,通过在垂直和水平方向上应用一维全局池化,对输入特征进行整合,生成了两个具备方向识别能力的独立特征图。

接下来,在全局感受野的宽度和高度两个方向上获得的特征图被连接在一起,通过一个共享的二维 1×1 卷积层和标准化进行变换,生成的 $f \in R^{C/r \times 1 \times (H+W)}$ 能够在水平方向和竖直的方向对空间信息进行编码。接着,将特征图 f 沿着空间维度分割成两个独立的张量: $f^b \in R^{C/r \times H}$ 和 $f^w \in R^{C/r \times W}$ 。再通过使用两个二维的 1×1 卷积层 F_b 和 F_w ,将 f^b 和 f^w 转换为与输入 X 具有相同数量的通道,得到以下输出:

$$g^b = \sigma(F_b(f^b)) \quad (2)$$

$$g^w = \sigma(F_w(f^w)) \quad (3)$$

其中, σ 表示 Sigmoid 激活函数。

最后,对得到的结果进行乘法加权来计算高度方向上的关注度权重 g^b 和宽度方向上的关注度权重 g^w 。然后,将这些权重应用于原始特征图 X 上,得到坐标注意力模块的最终输出。具体的计算过程如下:

$$y_c(i, j) = x_c(i, j) \times g_c^b(i) \times g_c^w(j) \quad (4)$$

坐标注意力模块插入在 DW 卷积和原始 MobileNetV2 块的卷积之间,从而创建了 MobileNetV2 的改进版本。既保证网络的轻量化,又使得资源更加合理的分配,CA 模块能够快速找到病害图像中感兴趣的区域,忽略背景和噪声信息。对

模型更准确地识别和定位感兴趣的对象非常有帮助。

在改进的 MobileNetV2 结果中,本文使用 Hardswish^[13] 作为激活函数,代替了原始版本中的 ReLU6,可以增强网络非线性能力,提高了神经网络的表达能力,有效解决了 ReLU6 激活函数神经元坏死的问题。而且还可以有效防止训练时梯度逐渐接近零时导致的饱和现象发生,进一步提升网络模型的表达能力,同时降低运算量,使得将来在移动设备上部署 HCA-MViT 模型更加便捷。Hardswish 函数如下式:

$$f(x) = \begin{cases} 0 & (x \leq -3) \\ x & (x \geq 3) \\ \frac{x(x+3)}{6} & (-3 < x < 3) \end{cases} \quad (5)$$

3 结果与分析

3.1 试验环境

本文采用 Python 编程语言, GPU 处理器为 NVIDIA GeForce GTX 2080Ti, 计算机内存为 16 GB, 操作系统为 ubuntu20.04, 深度学习框架为 Pytorch。

3.2 参数设置

为防止模型对固定顺序的数据出现依赖从而导致模型不具有泛化能力的现象,训练中将数据随机打乱。将数据增强后的数据按照 8:2 的比例划分为训练集和测试集,数据集的图片数量见表 2。经多次试验后,得到了本模型的最优参数组合,优化器为 RAdam^[14], 批量大小 Batch_size 为 32, 初始学习率为 0.0001, 损失函数设置为交叉熵损失函数。对于 PlantVillage 数据集, 迭代 epoch = 100 次; 对于

PlantDoc 数据集,迭代 epoch = 300 次。此外,由于 PlantDoc 数据集中的图像的格式不同,本文将该数据集全部转换为 RGB 图像。测试集不参与模型的训练,只进行模型评价指标的测试。

表 2 数据集分布情况

Table 2 Distribution of Datasets

数据集	训练集	测试集	总计
PlantVillage	43 443	10 861	54 304
PlantDoc	2 078	520	2 598

3.3 评价指标

本文通过准确率 (Accuracy)、F1 值、参数量、推理时间、模型大小等 5 个指标来评价模型的性能。这些评价指标综合考虑不同因素的评价方式,能够更全面地评估模型的性能表现。准确率和 F1 值的

定义如下式所示:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (6)$$

$$F1 = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \quad (7)$$

其中,TP 为真正例数;TN 为真负例数;FP 为假正例数;FN 为假负例数;Precision = TP/(FP+TP) 为精确率;Recall = (TP/TP+FN) 为召回率。

3.4 试验对比

3.4.1 HCA-MViT 模型与其它模型试验效果对比

为验证本文所构建的 HCA-MViT 模型在农作物病害识别任务中的有效性,将本文提出的 HCA-MViT 模型与 MobileViT、ResNet50、MobileNetV2 等六种神经网络模型进行对比,结果见表 3。

表 3 本文模型 HCA-MViT 与其他神经网络模型的对比结果

Table 3 Comparison results between the HCA-MViT model in this article and other neural network models

模型	准确率/%		F1		参数量	推理时间/ms	模型大小/MB
	Plant Village	Plant Doc	Plant Village	Plant Doc			
ResNet50 ^[15]	99.48	31.58	99.41	31.13	24.10	11.570	290.0
MobileViT	99.27	52.77	99.25	52.74	0.96	1.029	4.0
MobileNetV2 ^[16]	99.65	25.59	99.75	11.02	2.29	1.203	28.1
EfficientNetB0 ^[17]	99.38	31.78	99.38	23.55	4.08	1.999	49.8
VGG16 ^[18]	99.46	42.11	99.45	36.06	15.66	11.540	188.2
InceptionV3 ^[19]	99.38	40.29	99.38	37.55	21.88	12.568	263.7
HCA-MViT	99.68	56.96	99.68	56.80	0.97	1.044	4.1

由表 3 中可以得出,对于 PlantVillage 数据集,各个模型在该数据集的准确率均在 99% 以上,这是因为该数据集的背景较为单一、简单,训练识别起来较容易。其中准确率和 F1 分数最高的是 HCA-MViT 模型,准确率达到 99.68%,F1 分数为 99.68%,模型大小为 4.1 MB,其次是 MobileNetV2 模型,识别的准确率、F1 分数略低于 HCA-MViT,但是在参数量、推理时间、模型大小上均不及 HCA-MViT 模型。HCA-MViT 模型与 MobileViT 模型相比,在参数量、推理时间和模型大小基本一致的情况下,准确率提升了 0.41%,F1 分数提高了 0.43%。

对于 PlantDoc 数据集,模型的准确率均比较低,平均准确率在 60% 以下,这是因为 PlantDoc 数据集的背景复杂,干扰因素较多,该数据集更贴切实际生产的场景。准确率和 F1 分数最高的是 HCA-MViT 模型,该模型准确率为 56.96%,F1 分数为 56.80%,优于其它热门模型。相比于 MobileViT 模型准确率提升了 4.19%,F1 分数提高了 4.06%。

HCA-MViT 在 MobileViT 网络模型的基础上引入了坐标注意力和 Hardswish 激活函数,此改进使得模型能够保留更多的图像方向和位置感知信息,能够更好地捕捉图像中的细节和空间信息,通过这些改进,模型的识别准确率得到了提升,同时模型的收敛也更为迅速。这表明经过优化的模型在特征提取上更为强大,能够在一定程度上排除复杂背景等干扰信息的影响,增强了模型的鲁棒性。这使得 HCA-MViT 模型在图像识别领域,尤其是在面对复杂场景和任务时,变得更加高效。

3.4.2 消融试验

为了探究试验的改进方式对 MobileViT 模型带来的性能提升,在 PlantVillage 和 PlantDoc 两个数据集上进行消融试验,结果见表 4。其中“√”表示添加该结构,“×”表示未添加该结构。由表 4 可以得出,不论是在 PlantVillage 数据集还是在 PlantDoc 数据集上,引入坐标注意力机制和 Hardswish 激活函数后相较于原模型的识别准确率均为正提升,可以

有效的对农作物病害进行识别。

表 4 消融试验结果对比

Table 4 Comparison of ablation test results

数据集	Hardswish	CA	准确率/%	准确率提升/%
PlantVillage			99.27	
	✓		99.44	0.17 ↑
		✓	99.51	0.24 ↑
PlantDoc	✓	✓	99.68	0.41 ↑
			52.10	
	✓		53.83	1.73 ↑
		✓	54.91	2.81 ↑
	✓	✓	57.96	5.86 ↑

4 结束语

本文针对神经网络在农作物病害分类方面表现不佳、对计算资源的高消耗以及内存使用量大问题,提出了一种 HCA-MViT 分类识别算法。通过引入坐标注意力机制模块和 Hardswish 激活函数对 MobileViT 模型进行优化。本文在 PlantVillage 数据集和 PlantDoc 两个数据集上进行。试验结果表明,该模型与原始 MobileViT 模型和其他热门卷积神经网络算法相比,更准确地识别出农作物病害类型,该模型在参数量和准确率方面都展现出了优异的性能。这表明 HCA-MViT 算法在农作物病害分类任务上取得了有效的改进,能够有效地解决原有神经网络在该任务上的局限性,并且在资源消耗和模型性能方面取得了良好的平衡。此模型满足了在计算能力有限的设备上实现轻量化网络的快速推理和病害识别的部署需求。未来的工作将进一步针对自然背景下的应用场景,对网络模型进行改进,进一步提升模型的准确率。

参考文献

- [1] LECUN Y, BENGIO Y, HINTON G. Deep Learning [J]. Nature, 2015, 521(7553):436-444.
- [2] 肖天赐, 陈燕红, 李永可, 等. 基于改进通道注意力机制的农作物病害识别模型研究[J]. 江苏农业科学, 2023, 51(24):168-175.
- [3] 彭玉寒, 李书琴. 基于重参数化 MobileNetV2 的农作物叶片病害识别模型[J]. 农业工程学报, 2023, 39(17):132-140.
- [4] 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.

- [5] 王泽钧, 马凤英, 张瑜, 等. 基于注意力机制和多尺度轻量型网络的农作物病害识别[J]. 农业工程学报, 2022, 38(S1):176-183.
- [6] 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.
- [7] SHILL A, RAHMAN M A. Plant disease detection based on YOLOv3 and YOLOv4 [C]//Proceedings of 2021 International Conference on Automation, Control and Mechatronics for Industry 4.0 (ACMI). New York:ACM, 2021.
- [8] 王东方, 汪军. 基于迁移学习和残差网络的农作物病害分类[J]. 农业工程学报, 2021, 37(4):199-207.
- [9] MEHTA S, RASTEGARI M. MobileViT: Light-weight, general-purpose, and mobile-friendly vision transformer [J]. arXiv preprint arXiv, 2110.02178, 2021.
- [10] AbdallahAli. PlantVillage dataset [EB/OL]. (2019-09-06) [2025-01-21]. <https://www.kaggle.com/datasets/abdallahalidev/plantvillage-dataset>
- [11] SINGH D, JAIN N, JAIN P, et al. PlantDoc: A dataset for visual plant disease detection [C]//Proceedings of the 7th ACM IKDD CoDS and 25th COMAD. New York:ACM, 2020:249-253.
- [12] HOU Q B, ZHOU D Q, FENG J S. Coordinate attention for efficient mobile network design [C]//Proceedings of 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ:IEEE, 2021:13708-13717.
- [13] ALHASSAN A M, ZAINON W M N W. Brain tumor classification in magnetic resonance image using Hardswish-based ReLU activation function-convolutional neural network [J]. Neural Computing and Applications, 2021, 33(15):9075-9087.
- [14] LEONARDO S, KALLIL M C, ZIELINSKI L C, et al. RADAM: Texture recognition through randomized aggregated encoding of deep activation maps [J]. Pattern Recognition, 2023, 143:109802.
- [15] SHARMA A K, NANDAL A, DHAKA A, et al. Enhanced watershed segmentation algorithm-based modified ResNet50 model for brain tumor detection [J]. BioMed Research International, 2022, 2022(1):1-14.
- [16] SANDLER M, HOWARD A, ZHU M L, et al. Mobilenetv2: Inverted residuals and linear bottlenecks [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2018:4510-4520.
- [17] SUHARJIT O, NATANAEL G E, SEBASTIAN J P. Oil palm fresh fruit bunch ripeness classification on mobile devices using deep learning approaches [J]. Computers and Electronics in Agriculture, 2021, 188:106359.
- [18] FAHAD A, SAGHEER A, ATIFA A, et al. Identification of kidney stones in KUB X-ray images using VGG16 empowered with explainable artificial intelligence [J]. Scientific Reports, 2024, 14(1):6173-6173.
- [19] POORNIMA J J, ANITHA J, GNANA A H P, et al. Comparative analysis of melanoma classification using deep learning techniques on dermoscopy images [J]. Electronics, 2022, 11(18):2918.