

# 面向大规模多类别的病虫害识别模型

温长吉<sup>1,4</sup>, 王启锐<sup>1</sup>, 陈洪锐<sup>1</sup>, 吴建双<sup>1</sup>, 倪军<sup>2</sup>, 杨策<sup>3</sup>, 苏恒强<sup>1,4\*</sup>

(1. 吉林农业大学信息技术学院, 长春 130118; 2. 南京农业大学农学院, 南京 210095; 3. 明尼苏达大学食品、农业与自然资源科学学院, 圣保罗 55108; 4. 吉林农业大学智慧农业研究院, 长春 130118)

**摘要:** 早期病虫害精准识别是预警和防控的关键, 但是病虫害种类繁多数量巨大, 外部形态存在类间相似度高而类内差异性较大等性状特征, 导致病虫害识别仍然是一项极具挑战的工作。为实现病虫害识别分类任务中差异化特征的提取和表示, 该研究提出一种大规模多类别精细病虫害识别网络模型(a large-scale multi-category fine-grained pest and disease network, PD-Net)。首先通过在基准网络模型中引入卷积块注意力模型, 通过混合跨特征通道域和特征空间域实现模型在通道和空间两个维度上对关键特征提取和表示, 用以增强网络对差异化特征的提取和表示能力。其次引入跨层非局部模块, 提升模型在多个特征提取层之间对于多尺度特征的融合。在 61 类病害数据集和 102 类虫害数据集上的试验结果表明, 对比 AlexNet、VGG16、GoogleNet、Inception-v3、DenseNet121 和 ResNet50 模型, 该研究提出的面向大规模多类别病虫害识别模型, Top1 识别准确率在病害和虫害集上分别达到 88.617% 和 74.668%, 精确率分别达到了 0.875 和 0.745, 召回率分别达到 0.874 和 0.738, F1 值达到 0.874 和 0.732, 试验结果对比其他模型均有一定幅度的提升, 验证了 PD-Net 模型在大规模多类别病虫害识别上的有效性。

**关键词:** 模型; 深度学习; 病虫害; 精细分类; 卷积块注意力模块; 跨层非局部模块

doi: 10.11975/j.issn.1002-6819.2022.08.020

中图分类号: S24; TP391.41

文献标志码: A

文章编号: 1002-6819(2022)-08-0169-09

温长吉, 王启锐, 陈洪锐, 等. 面向大规模多类别的病虫害识别模型[J]. 农业工程学报, 2022, 38(8): 169-177.

doi: 10.11975/j.issn.1002-6819.2022.08.020 <http://www.tcsae.org>

Wen Changji, Wang Qirui, Chen Hongrui, et al. Model for the recognition of large-scale multi-class diseases and pests[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2022, 38(8): 169-177. (in Chinese with English abstract) doi: 10.11975/j.issn.1002-6819.2022.08.020 <http://www.tcsae.org>

## 0 引言

病虫害给农业造成巨大损失, 据世界粮农组织统计, 2020 年世界上因病虫害造成的粮食减损预计大于 10%, 局部地区甚至高达 30%<sup>[1]</sup>。因此, 早期实现病虫害精准识别, 对于后期防治具有重要意义。传统基于人工调查的病虫害识别方法, 存在专业技术人员短缺、费时费力、易错等问题, 因此开展半自动或全自动大规模病虫害精准识别研究与应用尤为重要<sup>[2-4]</sup>。

基于传统机器视觉技术实现病虫害半自动化识别, 在病虫害防治领域受到广泛关注。该项技术通过对可见光、高光谱、近红外光谱等病虫害影像提取视觉特征, 结合支持向量机、随机森林、朴素贝叶斯和 BP 神经网络等分类器, 实现病虫害识别, 对比人工识别方法实现效率提升<sup>[5-9]</sup>。传统机器视觉技术尽管取得较好的效果, 但

是仍然需要人工干预特征设计, 并且特征设计和目标识别优化是独立分阶段的工作, 缺少依据待识别目标自身多样性特征实现端到端的自适应调整<sup>[10-13]</sup>。

近年深度学习模型通过构建深层网络结构实现端到端的自动学习<sup>[14]</sup>。较为经典的深度学习模型包括 AlexNet、VGGNet、Inception 和 ResNet 等。其因较为优异的性能, 在病虫害识别等领域得到推广应用<sup>[15-19]</sup>。基本病虫害识别模型相关研究包括: 结合浅层迁移训练的 VGG16 模型应用于苹果黑腐病识别<sup>[20]</sup>, 利用 CaffeNet 模型对 13 类水稻害虫识别分类<sup>[21]</sup>, 使用 VGG16 模型对 5 种茄子病害进行识别<sup>[22]</sup>等。由于病虫害形态各异, 影像采集环境多样, 因此一些学者提出通过模型改进以适应病虫害识别任务的复杂性, 提升识别准确率。改进病虫害识别模型相关研究包括: 利用改进 AlexNet 网络对 6 种黄瓜叶片病害进行识别<sup>[23]</sup>, 结合余弦相似度改进 Inception 网络对荞麦病害进行识别分类<sup>[24]</sup>等。最近一些更为复杂的深度学习模型被构建用于更多类别的病虫害识别分类任务, 取得较好的识别效果<sup>[25-31]</sup>, 相关工作通过引入实现多层特征融合的金字塔模块, 优化网络学习性能的 RMSProp 算法和优化网络结构的轻量级残差模块 (Scale-Down Block), 实现改进基准模型, 应用于病虫害识别准确率达到 85% 以上<sup>[25-28]</sup>。为提升基准模型的识别性能, 将注意力机制, SE 模块 (Squeeze-and-Excitation) 和注意力池化策略, 引入到模型中用于包括番茄、叶霉

收稿日期: 2022-02-09 修订日期: 2022-04-10

基金项目: 吉林省自然科学基金 (20180101041JC); 吉林省发展与改革委员会产业技术与开发项目项目 (2021C044-8); 吉林省教育厅科学技术研究项目 (JKH20210335KJ, JKH20190924KJ); 国家自然科学基金重点项目 (U19A2061); 国家重点研发计划项目 (2017YFD0502001)

作者简介: 温长吉, 博士, 副教授, 硕士生导师, 研究方向为机器视觉、模式识别、动植物表型解析。Email: chagou2006@163.com

\*通信作者: 苏恒强, 副教授, 硕士生导师, 研究方向为机器视觉、农业物联网应用等。Email: suhengqiang@jlau.edu.cn

病等病虫害识别分类, 识别准确率均达到 90%以上<sup>[29-31]</sup>。尽管当前工作在各自的研究中均取得较好的结果, 但是仍然面临一定的问题。首先多数工作使用的数据集病虫害种类较为单一, 样本量较少, 且较多都是针对某一种或几种作物进行试验, 缺少在大规模多类别病虫害数据集上的识别分类应用。其次采集的病虫害影像多数在实验室环境下拍摄, 光照、背景和形态变化较为单一, 实际生产环境应用的适应性有待进一步提高。

综上所述, 大规模病虫害识别仍然是一项极具挑战性的工作, 其中如何实现病虫害类别间差异化判别特征提取和表示是重要的研究方向之一。针对该问题, 本研究在基准网络模型中引入一种性能较为优异的视觉注意力机制模块一卷积块注意力机制 (Convolutional Block Attention Module, CBAM)<sup>[32]</sup>, 通过混合跨特征通道域和空间域, 分别实现通道和空间两个维度上对关键特征的提取和表示, 即在通道域提取对病虫害分类任务贡献度较高的通道维度信息, 在空间域提取贡献度较高的位置维度信息<sup>[33]</sup>。其次, 多尺度的特征提取和表示尤为重要, 由于在病虫害识别分类任务中个体的尺寸和形状差异较大, 同层卷积核感受野尺度固定, 感受野与个体之间不匹配在一定程度上影响特征提取, 因此在基准模型中引入跨层非局部模块 (Cross-layer Non-local Module, CNL), 即在基准网络模型的多个特征提取层之间分别选取较深层和多个浅层建立空间响应关系, 从而学习更多多尺度特征, 用以提升基准模型的特征提取和表示能力<sup>[34]</sup>。在现有研究工作的基础上, 本研究的主要工作在于: 1) 构建一个面向大规模多类别的病虫害识别模型 PD-Net。以 ResNet50 作为基准网络, 分别引入卷积注意力模块和跨层非局部模块, 提升模型对差异化特征的提取和表示能力; 2) 开发基于 PD-Net 模型的手机端病虫害识别 APP, 通过手机拍摄实现对病虫害识别, 并且根据搜索的结果给出相应病害和虫害的性状特征以及防治措施等相关信息, 为作物病虫害识别提供了便利条件。

## 1 试验数据

### 1.1 数据集

病害数据集的构建由公开数据集和实施自侵染培养数据两部分构成。自侵染培养的病害包括樱桃白粉病、葡萄黑腐病、番茄斑枯病、柑橘黄龙病。将受侵染的病害叶片采摘后, 依据一般和严重程度进行划分, 在灰色的背景板下均匀光照使用相机进行拍摄, 每类图像 100 张, 共计 800 张图像。公开数据集选取 AI-challenger 2018 数据集中的农作物病害数据集, 该数据集共包含 10 种植物 (苹果、樱桃、葡萄、柑桔、桃、草莓、番茄、辣椒、玉米、马铃薯) 的 27 种病害, 其中的 24 种又分为“一般”和“严重”两种状态。病害数据集共计 61 类 37,058 张图像。图 1 为不同严重程度的病害数据集部分示例。

虫害数据集选用 IP102 作物病害数据集<sup>[35]</sup>。该数据集包含水稻、玉米、小麦、甜菜、苜蓿、葡萄、柑橘、芒果共 8 种作物的 102 类虫害, 共计 75 222 张图像。数据集构建分别包括在 Google、Flickr、Bing 等网站上输入

关键字进行识别收集, 从农业和昆虫科学等专业网站上识别获取, 以及在互联网上的虫害视频通过剪辑获取。最后对候选虫害影像进行标注过滤, 通过组织植保专家分别采用先进行独立标注, 然后进行协同标注, 保留专家一致标注的影像。部分虫害图像示例图片如图 2 所示。

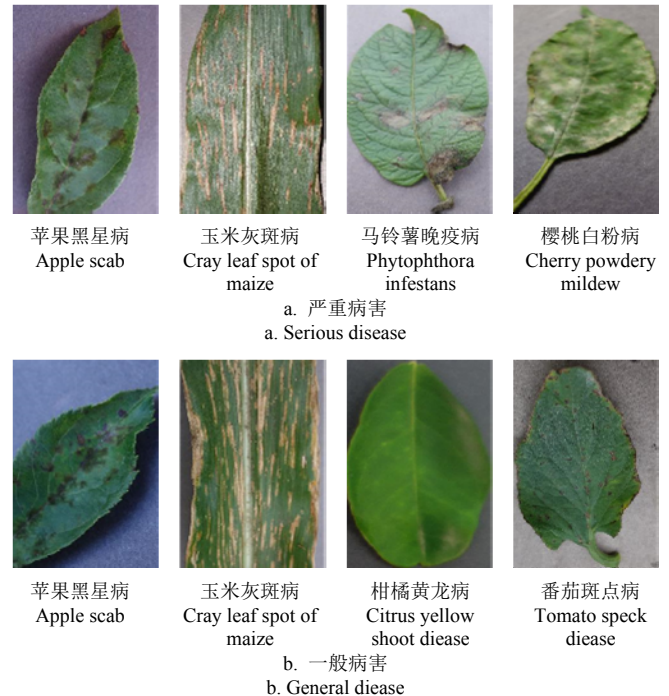


图 1 病害数据集部分示例

Fig.1 Partial examples of disease data sets



图 2 虫害数据集部分示例

Fig.2 Partial examples of pest data sets

### 1.2 数据预处理

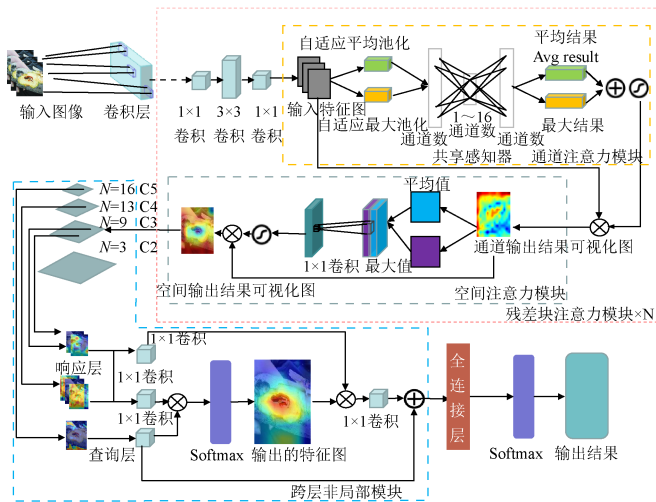
由于数据集中存在图像与类别标签不一致 (即样本丢失) 的问题, 因此本研究对两种数据集进行了数据清洗, 清洗后的病害和虫害数据集图像分别为 36 923 和 71 647 张。统计发现, 数据样本分布不平衡, 其中病害数据集清洗后共计 59 类, 最多的 3 种病害分别为番茄黄化曲叶病严重、柑橘黄龙病一般以及柑橘黄龙病严重,

分别有 2 825、2 197、2 161 张图像。最少的病害图像分别为番茄疮痂病严重、番茄疮痂病一般以及番茄斑点病严重，分别含有 1、2、27 张图像。虫害数据集清洗后共计 102 类，其中最多的 3 种虫害分别为叶蝉、斑衣蜡蝉和盲蝽，分别含有 5740、5310 和 5 081 张图像，最少的 3 种虫害分别为葡萄斑叶蝉、黑点蚱壳虫以及葡萄短须蚧，分别有 71、74、79 张图像。

针对数据样本分布不均衡问题，对病虫害数据集进行在线数据增强处理，在增加样本多样性的同时避免传统离线数据增强方法带来的存储负载问题。同时，数据集图像大部分尺寸 300 像素×400 像素。为了满足卷积神经网络输入图像尺寸要求，将图像尺寸统一为 448 像素×448 像素。在线数据增强主要包括水平翻转、垂直翻转、旋转、增加噪声和改变饱和度。

## 2 一种面向大规模多类病虫害精细识别模型

本研究构建了一种面向多类别病虫害识别模型 (a large-scale multi-category fine-grained pests and diseases network, PD-Net)，该模型在基准网络模型 ResNet50 中分别引入卷积块注意力模块 (Convolutional Block Attention Module, CBAM) 和跨层非局部模块 (Cross-layer Non-local Module, CNL)，分别用于增强通道和空间维度实现对于差异化特征信息的区分表示，以及融合多尺度特征从而实现差异化特征的有效提取和表示。具体模型框架结构如图 3 所示。



注：⊙表示 Sigmoid 函数，⊗表示矩阵的加权操作，⊕表示卷积特征与多注意力特征的合并数组计算。C2~C5 分别表示第 N 个残差块输出结果，N 为 3、9、13、16。

Note: “⊙” represents Sigmoid function, “⊗” represents weighting operation of the matrix, “⊕” represents combined array calculation of convolutional features and multi-attentional features. C2-C5 denote the results of the residual block output when N is 3, 9, 13, and 16, respectively.

图 3 病虫害识别模型框架图

Fig.3 Structure diagram of disease and pest recognition model

### 2.1 基于卷积块注意力改进卷积模块

视觉注意力 (Visual attention) 用于提升机器视觉模型对于感兴趣区域的关注，从而实现模型性能的提升<sup>[36-38]</sup>。卷积块注意力机制 CBAM<sup>[33]</sup>作为视觉注意力一种性能较为优异的模组，通过卷积操作在局部感受野内增强空间和通道对于特定目标识别任务的贡献度，即在基准网络

模型的特定卷积块中引入通道注意力机制模块和空间注意力机制模块，实现特征提取和表示提升模型性能。CBAM 模块具体实现机制如下，通道注意力模块 (Channel Attention Module, CAM)，首先将输入的特征图在宽和高两个维度上分别实现全局最大池化和全局平均池化，然后再经过多层感知器，输出的特征顺次通过加和和 sigmoid 运算最终输出通道注意力特征图。空间注意力模块 (Spatial Attention Module, SAM)，首先基于通道做最大池化和平均池化操作，将输出结果基于通道做连接操作，然后通过卷积操作实现通道降维，再经过 sigmoid 函数生成空间注意力特征图。最后构建注意力特征图，将通道注意力模块和空间注意力模块串行连接构建 CBAM 模块，并插入在基准网络特定层的卷积模块中。在本文提出的面向大规模多类别病虫害识别模型 PD-Net 中，在基准模型的第 C2~C5 层的指定卷积块中引入 CBAM 模块，构建改进的卷积模块。

### 2.2 跨层非局部模块

在图像识别分类任务中，跨层多尺度特征融合是提升模型特征提取和表示的重要方法<sup>[34,39]</sup>。跨层非局部模块 CNL<sup>[34]</sup>首先通过计算查询层特征与多层响应层特征之间的相关性，学习描述差异化信息的多尺度特征。然后依据相关性对响应特征进行加权，并添加到查询特征中实现加权机制。具体表达式如下：

$$CNL(X^q, X^{r1}, \dots, X^{rn}) = X^q + \sum Z^i \quad (1)$$

$$Z^i = Z(f(\theta(X^q), \phi_i(X^{ri}))g_i(X^{ri})) \quad (2)$$

式中 CNL 代表跨层非局部模块，给定 n 个响应层， $r1, r2, \dots, rn$  分别代表特定实现跨层操作由浅及深的响应层， $X^q, X^{r1}, \dots, X^{rn}$  分别代表对应响应层特征， $\theta(X^q), \phi_i(X^{ri}), g_i(X^{ri})$  分别代表实现可学习的特征变换函数，即特定的  $1 \times 1$  卷积操作， $X^q$  代表查询层局部特征， $f()$  用于相似度计算，实现过程为查询层和部分响应层分别经过  $1 \times 1$  卷积后进行矩阵相乘，然后通过 Softmax 生成特征图。Z 代表卷积， $Z^i$  代表 i 层的卷积。

### 2.3 构建模型

本文构建了面向大规模多类别病虫害的识别模型 PD-Net。模型算法实现步骤如下：1) 卷积特征提取：输入一张大小为  $448 \times 448$  像素的图像，第一层利用 64 个步长为 2 的  $7 \times 7$  的卷积核提取特征，再经过批标准化、激活函数以及步长为 2 的  $3 \times 3$  卷积进行最大池化操作；2) 通道注意力模块构建：将输出的特征输入残差块提取特征，将得到的特征分别经过最大池化和平均池化。将池化特征分别经过多层感知器 (Multilayer Perceptron, MLP)，利用  $1 \times 1$  卷积—ReLU 激活函数— $1 \times 1$  卷积，实现先降维后升维。将 MLP 输出的特征执行加和操作，再经过 sigmoid 激活操作，得到代表每个特征通道重要性的实数，将其和原始特征进行加权操作，得到带有特征侧重的特征集合从而输出通道注意力特征；3) 空间注意力模块构建：将通道注意力得到的特征分别经过全局最大池化和全局平均池化，然后将这 2 个结果做连接 (concat) 操作，经过一个  $7 \times 7$  卷积，将其降维为 1 个通道，再经

过 sigmoid 激活操作得到空间的权重向量和原始特征进行加权操作, 输出空间注意力特征; 4) 构建基于 CBAM 的注意力模块: 将 2)、3) 步骤结果与原始输入进行残差连接, 共同构建改进后的注意力模块。注意力模块需要重复操作  $N$  次,  $N$  为重复构建注意力块的执行次数 (按 ImageNet<sup>[14]</sup>将 ResNet50<sup>[15]</sup>分为 C1~C5 层), 随着残差块的不断叠加, 当  $N=3$  时, 经过降维操作 (步长为 1 的  $1 \times 1$  卷积) 使得图像尺寸缩小为原来的一半, 特征通道数为原来的一倍, 输出网络为 C2。当  $N=9$  时, 同样经过降维操作, 输出网络为 C3, 以此类推,  $N=13、16$  分别再输出 C4、C5; 5) 跨层非局部模块, 经过上述步骤得到 C1~C5 层的特征, 取 C3\_2、C3\_4、C4\_2、C4\_4、C4\_6 (C3\_2 表示 C3 层的第②个注意力模块的输出, 其余同理) 作为响应层, 将 C5\_2 作为查询层。将查询层和响应层作  $1 \times 1$  卷积进行降维操作, 得到 256 个通道, 然后将查询层和响应层作矩阵乘法, 经过 softmax 生成特征注意力权重和原始响应层特征进行加权操作, 经过  $1 \times 1$  卷积与查询层进行相加实现不同层特征之间的特征融合; 6) 将输出特征, 经过全连接层, 连接 softmax 得到代表害虫病害类别的向量  $K_{out}$ , 并在最终获得识别结果。本文项目开源代码地址 <https://github.com/wujians122/PD-Net>。

## 2.4 试验环境和参数设置

试验在 DELL Precision T7920 深度学习工作站进行, 该工作站配置为 2.1 GHz 的 Intel(R) Xeon(R) Gold 5218 CPU、62.5 GB DRAM、2 TB 硬盘和一个 GeForce RTX 2080 Ti/PCIe/SSE2 显卡组成。运行环境是 Ubuntu 18.0.4, Pytorch, Python 3.7。

病害和虫害数据集分别被划分为训练集和验证集, 其中训练集占比 70%, 验证集占比 30%。评价指标采用 TOP1 准确率、TOP2 准确率、损失值 (Loss)。由于病害和虫害数据集中类别数据存在严重不平衡问题, 因此召回率 (Recall)、精确率 (Precision)、 $F1$  值采用加权计算结果作为模型优劣的评判标准。

TOP1 准确率、TOP2 准确率 (Accuracy)、损失值 (Loss) 计算表达式如下:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FN} + \text{FP} + \text{TN}} \quad (3)$$

$$\text{Loss} = -\sum_x (p(x) \ln q(x)) \quad (4)$$

式中 TP 为真正例, TN 为真负例, FP 为假正例, FN 为假负例,  $p$  代表期望输出,  $q$  代表实际输出。

加权召回率、精确率、 $F1$  值计算式如下:

$$\left\{ \begin{array}{l} \text{Precision}_v = \frac{\text{TP}_v}{\text{TP}_v + \text{FP}_v} \\ \text{Precision}^{**} = \frac{\sum_{v=1}^L \text{Precision}_v \cdot \omega_v}{L} \end{array} \right. \quad (5)$$

$$\left\{ \begin{array}{l} \text{Recall}_v = \frac{\text{TP}_v}{\text{TP}_v + \text{FN}_v} \\ \text{Recall}^{**} = \frac{\sum_{v=1}^L \text{Recall}_v \cdot \omega_v}{L} \end{array} \right. \quad (6)$$

$$\left\{ \begin{array}{l} F1 = \frac{2(\text{Precision} \cdot \text{Recall})}{\text{Precision} + \text{Recall}} \\ F1^{**} = \frac{\sum_{v=1}^L F1_v \cdot \omega_v}{L} \end{array} \right. \quad (7)$$

式中  $v$  代表类别索引;  $L$  代表总类别数;  $\omega_v$  为权值, 以第  $v$  类样本占总体样本量的比例表示; Precision\*\*、Recall\*\*、 $F1^{**}$  分别代表加权精确率、召回率和  $F1$  值。

## 3 结果与分析

### 3.1 消融试验

试验以 ResNet50 为基准模型架构, 分别引入卷积注意力模块、跨层非局部模块, 在病害和虫害数据集上进行试验, 验证 PD-Net 模型的性能。消融试验结果如表 1 所示。

表 1 病害和虫害数据集消融试验结果  
Table 1 Ablation results of disease and pest data sets

模型 Model	数据集 Data set	准确率 Accuracy/%		损失值 Loss	精确率 Precision		召回率 Recall		F1 值 F1 value	
		Top1	Top2		未加 权	加权	未加 权	加权	未加 权	加权
		ResNet50	病害	85.502	95.734	0.431	0.846	0.864	0.832	0.865
	虫害	71.828	79.989	1.479	0.666	0.694	0.634	0.716	0.650	0.718
ResNet50+	病害	87.318	97.236	0.386	0.855	0.871	0.833	0.872	0.844	0.872
CNL	虫害	72.762	81.632	1.227	0.673	0.719	0.636	0.723	0.654	0.727
ResNet50+	病害	87.239	97.024	0.394	0.854	0.866	0.848	0.849	0.851	0.832
CBAM	虫害	72.748	81.426	1.242	0.671	0.709	0.645	0.716	0.658	0.723
PD-Net	病害	88.617	98.922	0.348	0.886	0.875	0.868	0.874	0.877	0.874
	虫害	74.668	83.298	1.154	0.686	0.745	0.702	0.738	0.694	0.732

依据表 1 试验结果, PD-Net 模型在病害和虫害集上的 Top1 识别准确率分别达到 88.617% 和 74.668%。在基准模型基础上引入 CNL 模块后, 在两个数据集上的试验结果对比基准网络模型 ResNet50 分别提升 1.816 个百分点和 0.934 个百分点, 在基准模型上加入 CBAM 注意力模块后, 试验结果分别提升 1.737 个百分点和 0.920 个百分点。PD-Net 模型, 在两个数据集上的试验结果对比 ResNet50 分别提升 3.115 个百分点、2.840 个百分点, Top2 的识别准确率分别提升了 3.188 个百分点和 3.309 个百分点。损失值方面, 改进的模型比其他 3 个模型在病害数据集上分别降低了 0.073~0.325, 在虫害数据集上分别降低了 0.038~0.083。在精确率、召回率以及  $F1$  值中, 分别给出了在病害和虫害数据集上的加权以及未加权的试验结果。未加权指标, 精确率最大提升了 0.04 和 0.02, 召回率最大提升了 0.036 和 0.068,  $F1$  值最大提升了 0.038 和 0.044。加权指标, 精确率最大提升了 0.011 和 0.051, 召回率最大提升了 0.009 和 0.022,  $F1$  值最大提升了 0.009 和 0.014。通过消融试验验证 PD-Net 模型的有效性。

为进一步验证本文工作提出的 PD-Net 模型的性能, 分别统计 PD-Net 在病害和虫害数据集上识别准确率区间范围内的类别数量。结果如表 2 所示。

表 2 病害和虫害混淆矩阵统计表  
Table 2 Disease and pest confusion matrix statistics

项目 Item	准确率 Accuracy			
	1.0	0.9~<1.0	0.8~<0.9	<0.8
病害种类数 Disease species number	4	18	16	21
虫害种类数 Pest species number	1	6	13	82

从表中可以看出，在病害数据集中，识别准确率准达到 1.0 的有 4 类，分别为苹果雪松锈病严重、樱桃白粉病一般、玉米健康类以及柑橘健康。在 0.9~<1.0 范围的共计 18 类。其中包含苹果健康、苹果灰斑病、樱桃健康、玉米花叶片病毒病、葡萄健康、辣椒健康、马铃薯健康、草莓健康、番茄健康、番茄花叶病毒病共计 10 类准确率达到了 0.97~0.99，其他 8 类达到了 0.90~0.96，玉米花叶病毒病等 16 类准确率达到 0.8~0.9，还有 21 类病害精度低于 0.87，其中葡萄褐斑病一般和番茄斑点病严重的分类准确率低至 0.33，所以导致总体精度有所下降，查看混淆矩阵识别结果发现葡萄褐斑病一般被误分为葡萄褐斑病严重以及番茄斑点病严重，番茄斑点病严重被误分为番茄斑点病一般和番茄斑枯病一般。在虫害数据集中，识别准确率达到 1.0 的为横线尾夜蛾。准确率 0.9~1.0 的共计 6 类，分别为蛆、蚜虫、苜蓿广肩小蜂、绿盲蝽、葡萄短须螨和斑衣蜡蝉。蜈蚣等 16 类准确率达到 0.8~<0.9，同时有 82 类准确率低于 0.8，其中部分类别识别精度低于 0.5 导致总体精度提升效果不明显，其中第 10 类星瓢萤叶甲类被全部误分为牛角花齿蓟马，通过 ID 索引误分图像可以看出两种虫害具有高度相似性。

### 3.2 对比试验

为进一步验证本文提出的 PD-Net 模型的有效性，对比其余较为经典的卷积网络模型 AlexNet、VGG16、GoogleNet、Inception-v3、DenseNet121 和 ResNet50，对比模型均参考模型原始框架以及参数设置进行。采用 Adam 作为优化器，样本批次 (batch size) 设置为 16，学习率为 0.0001。对比试验结果如表 3 所示。

从表 3 可知，与经典深度学习模型对比，PD-Net 模型在识别准确率上具有明显的提升。在病害数据集试验中，本文工作提出的 PD-Net 模型 TOP1 识别准确率达到 88.617%，对比 AlexNet、VGG16、GoogleNet、Inception-v3 以及 DenseNet121 分别高出了 4.331 个百分点、3.473 个百分点、2.117 个百分点、2.495 个百分点以及 1.748 个百分点。在虫害数据集试验中，模型 Top1 识别准确率达到 74.668%，对比其他模型分别提升 8.122 个百分点、5.662 个百分点、1.906 个百分点、2.159 个百分点以及 2.931 个百分点。PD-Net 模型在病害和虫害数据集上 Top2 识别准确率分别高达 98.922%和 83.298%，对比其他深度学习模型病害数据集上识别准确率提升 1.469 个百分点~6.076 个百分点，在虫害数据集上提高了 1.869 个百分点~6.644 个百分点。对比其他模型，病害和虫害数据集损失值方面，最多降低了 0.100 和 0.317。精确率方面分别提升 0.02~0.036 和 0.02~0.084。召回率以及 F1 值采用未加权和加权计算方式在病害和虫害数据集上均获得提升。

表 3 病害和虫害数据集对比试验结果  
Table 3 Comparative experimental results of disease and pest data sets

模型 Model	数据集 Data set	准确率 Accuracy/%		损失值 Loss	精确率 Precision		召回率 Recall		F1 值 F1 value	
		Top1	Top2		未加权	加权	未加权	加权	未加权	加权
		AlexNet	病害	84.286	92.846	0.448	0.801	0.839	0.787	0.841
	虫害	66.546	76.654	1.471	0.598	0.661	0.592	0.663	0.595	0.665
VGG16	病害	85.144	94.798	0.381	0.833	0.849	0.831	0.851	0.832	0.851
	虫害	69.006	78.347	1.383	0.613	0.684	0.576	0.687	0.594	0.69
GoogleNet	病害	86.500	96.706	0.348	0.853	0.873	0.845	0.874	0.849	0.875
	虫害	72.762	81.429	1.197	0.678	0.722	0.652	0.723	0.665	0.726
Inception-v3	病害	86.122	96.872	0.374	0.849	0.869	0.831	0.871	0.840	0.871
	虫害	72.509	81.032	1.178	0.670	0.725	0.627	0.728	0.648	0.731
DenseNet121	病害	86.869	97.453	0.355	0.854	0.866	0.834	0.868	0.844	0.868
	虫害	71.737	80.968	1.148	0.641	0.705	0.592	0.711	0.616	0.717
PD-Net	病害	88.617	98.922	0.348	0.886	0.875	0.877	0.874	0.868	0.874
	虫害	74.668	83.298	1.154	0.686	0.745	0.702	0.738	0.694	0.732

利用部分实测数据验证效果。利用训练后的 PD-Net 模型对自采数据，包括樱桃白粉病、葡萄黑腐病、番茄斑枯病、柑橘黄龙病划分为病害侵染一般和严重，8 类每类 100 张共计 800 张数据上进行试验，模型对比试验结果如表 4 所示。

表 4 自采数据集试验结果  
Table 4 Self-collected dataset test results

模型 Model	准确率 Accuracy/%		损失值 Loss	加权后 Weighted		
	Top1	Top2		精确率 Precision	召回率 Recall	F1 值 F1 value
	AlexNet	81.625	98.500	0.479 0	0.846	0.816
VGG16	91.250	99.875	0.222 1	0.917	0.913	0.913
GoogleNet	89.250	99.500	0.269 8	0.899	0.893	0.894
Inception-v3	91.375	99.625	0.219 8	0.941	0.936	0.937
DenseNet121	91.500	99.875	0.201 9	0.929	0.925	0.926
PD-Net	93.625	99.875	0.197 9	0.958	0.953	0.955

从表 4 可知，本文提出的模型 PD-Net，在所有评价指标上均取得了较好的结果，且明显优于病害和虫害数据集上的结果，分析原因在于自采构建数据样本均衡，标注准确，且部分数据可能由于作为训练集在模型训练过程中出现。因此 PD-Net 在公开数据集和自采数据上取得较好的结果。试验结果验证了本研究提出的 PD-Net 模型性能。

图 4 为 PD-Net 模型对比其他卷积网络模型分别在 Top1 识别准确率和 Loss 值在 100 轮的收敛情况。从图中可以更直观的看到本研究提出的 PD-Net 网络对比其他深度学习模型网络在 Top1 的精度的提高以及 Loss 值上的下降。

### 3.3 注意力机制可视化热图

为进一步解释 PD-Net 模型在基准模型引入注意机制对模型性能提升的影响，选取病害和虫害数据集的部分示例对注意力机制对模型抽取特征图进行可视化，结果如图 5 所示。

通过热图可视化结果可以看出，病害和虫害所在区

域为高亮偏暖色调,表明通过引入注意力机制后模型在特征提取和表示上对于分类识别任务共享度较高的目标

区域给予更高的关注,从而在一定程度上降低背景区域对病虫害识别的干扰,进而提升网络模型的性能。

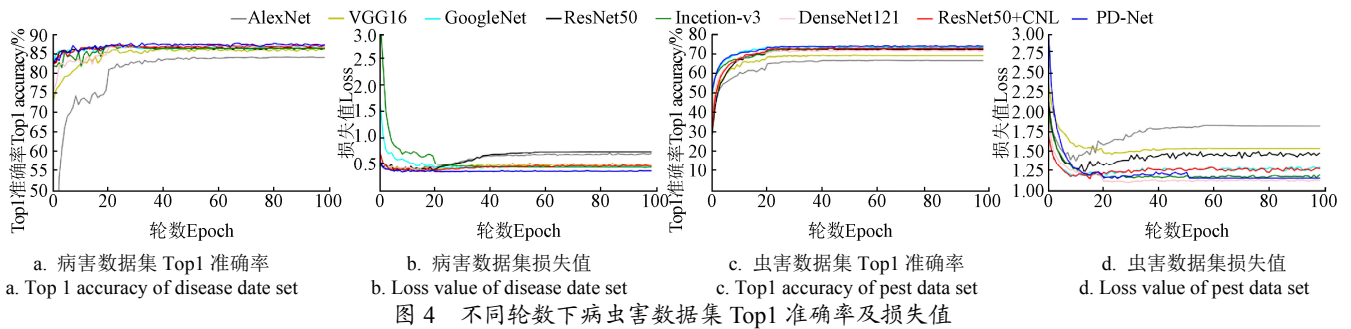
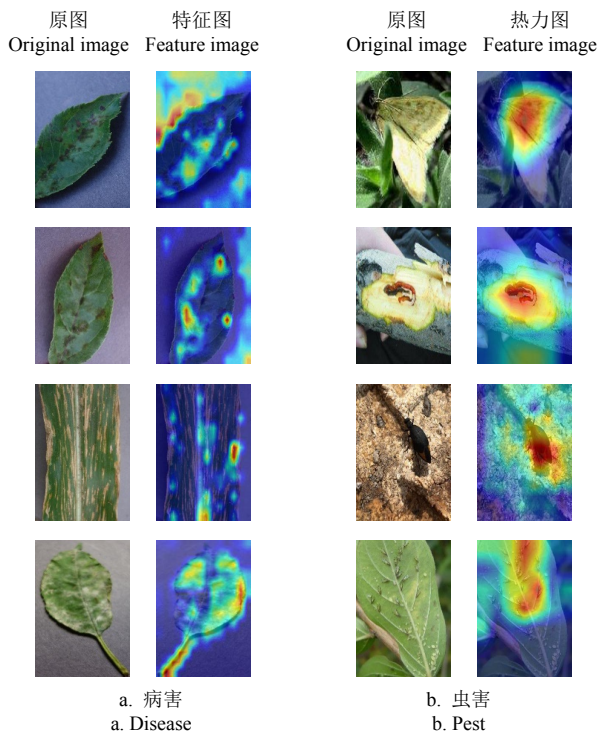


Fig.4 Top1 accuracy and loss of disease and pest data set with different epochs



注:不同颜色的热力特征代表了不同区域的关注度,颜色越红,关注度越高。  
Note: Different colors of heat features represent the attention of different regions, the red the color, the higher the attention.

图 5 注意力机制可视化图

Fig.5 Visual diagram of attention mechanism

### 4 病虫害识别手机端开发

本文工作在提出 PD-Net 模型基础上,将模型集成到基于 Android 系统的手机端,研发应用程序。病虫害识别应用程序开发步骤包括安装环境配置、模型导入和参数调试和软件测试。

病虫害识别程序是在 OPPO reno 4 pro 上进行测试,当前运行该 APP 需要运行内存不低于 512 MB,内存不低于 1 GB,Android 版本不低于 4.0。模型集成准备,模型准备主要是获取模型.pb 文件和标签文件.txt 文件,前期分别在服务器端训练和测试生成后导出。软件安装环境配置围绕开发工具和 SDK 的安装和环境配置。在 Android Studio 中打开 Project view,把.pb 文件和.txt 文件放在 assets 下,定义变量和接口初始化,配置文件对应的的版本,在 APP 页面添加 3 个标签按钮,分别是照片、相机、识别功能,之后配置按钮对 TF 模型的调用,执行“Build

APK”,即可生成可安装文件包(APK 文件)。然后,通过 3G/4G/5G/无线网络传输到 Android 手机,也可以通过有线方式把安装包移至 Android 手机,完成基于 Android 的植物病害叶片 APP 的移植。

在识别的过程中,初始登录 APP 界面如图 6a 所示,界面低端有 3 个功能按钮:拍照、图像、搜索。拍照获取权限后会自动打开手机摄像头拍摄待识别个体图像,点击图像按钮可以将拍摄照片保存,点击搜索按钮后将在应用软件端进行分类识别反馈相应结果。分别选取葡萄黑腐病和稻纵卷叶螟进行拍摄测试,病害和虫害测试识别结果界面截图如图 6b 和 6c 所示,分别给出识别病害和虫害的中英文名、发生区域以及病虫害性特征,防治措施等。



图 6 病虫害识别系统及识别结果图

Fig.6 Pest identification system and identification result chart

### 5 结论

病虫害给农业造成了巨大损失,由于类内差异性、

类间相似性、规模大等因素使得对病虫害进行精细识别产生较大的困难。为解决问题, 本研究提出一种面向大规模多类病虫害的精细识别模型 PD-Net, 分别在 61 类病害数据集和 102 类虫害数据集上进行试验。试验结果表明, 本文提出 PD-Net 模型在 Top1 识别准确率在病害和虫害集上分别达到 88.617% 和 74.668%, 损失值分别达到了 0.348 和 1.154, 精确率分别达到了 87.5% 和 74.5%, 对比其他经典深度学习模型, 在病害数据集上 Top1 准确率分别提升了 1.748 个百分点~4.331 个百分点, 损失值最多降低了 0.100, 精确率提升了 0.02~0.036, 虫害数据集上 Top1 准确率分别提升了 1.906 个百分点~8.122 个百分点, 在损失值上最多降低了 0.317, 精确率提升了 0.02~0.084。同时召回率以及 F1 值也有着一定的提升, 验证 PD-Net 模型的有效性。

#### [参 考 文 献]

- [1] Eskola M, Kos G, Elliott C T, et al. Worldwide contamination of food-crops with mycotoxins: Validity of the widely cited 'FAO estimate' of 25%[J]. *Critical reviews in Food Science and Nutrition*, 2020, 60(16): 2773-2789.
- [2] Cheng X, Zhang Y, Chen Y, et al. Pest identification via deep residual learning in complex background[J]. *Computers and Electronics in Agriculture*, 2017, 141: 351-356.
- [3] Rahman C R, Arko P S, Ali M E, et al. Identification and recognition of rice diseases and pests using convolutional neural networks[J]. *Biosystems Engineering*, 2020, 194: 112-120.
- [4] 贾少鹏, 高红菊, 杭潇. 基于深度学习的农作物病虫害图像识别技术研究进展[J]. *农业机械学报*, 2019, 50(S1): 313-317.  
Jia Shaopeng, Gao Hongju, Hang Xiao. Research progress on image recognition technology of crop pests and diseases based on deep learning[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2019, 50(S1): 313-317. (in Chinese with English abstract)
- [5] 李波, 刘占宇, 黄敬峰, 等. 基于 PCA 和 PNN 的水稻病虫害高光谱识别[J]. *农业工程学报*, 2009, 25(9): 143-147.  
Li Bo, Liu Zhanyu, Huang Jingfeng, et al. Hyperspectral identification of rice diseases and pests based on principal component analysis and probabilistic neural network[J]. *Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE)*, 2009, 25(9): 143-147. (in Chinese with English abstract)
- [6] 温芝元, 曹乐平. 椴柑果实病虫害的傅里叶频谱重分形图像识别[J]. *农业工程学报*, 2013, 29(23): 159-165.  
Wen Zhiyuan, Cao Leping. Citrus fruits diseases and insect pest recognition based on multifractal analysis of Fourier transform spectra[J]. *Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE)*, 2013, 29(23): 159-165. (in Chinese with English abstract)
- [7] Ropelewska E, Zapotoczny P. Classification of Fusarium-infected and healthy wheat kernels based on features from hyperspectral images and flatbed scanner images: A comparative analysis[J]. *European Food Research and Technology*, 2018, 244(8): 1453-1462.
- [8] Mishra M, Singh P K, Brahmachari A, et al. A Robust Pest identification system using Morphological analysis in neural networks[J]. *Periodicals of Engineering and Natural Sciences (PEN)*, 2019, 7(1): 483-495.
- [9] Bienkowski D, Aitkenhead M J, Lees A K, et al. Detection and differentiation between potato (*Solanum tuberosum*) diseases using calibration models trained with non-imaging spectrometry data[J]. *Computers and Electronics in Agriculture*, 2019, 167: 105056.
- [10] Zhang D, Chen G, Zhang H, et al. Integration of spectroscopy and image for identifying fusarium damage in wheat kernels using hyperspectral imaging[J]. *Spectrochimica Acta Part A Molecular and Biomolecular Spectroscopy*, 2020, 236: 118344.
- [11] Lins E A, Rodriguez J P M, Scoloski S I, et al. A method for counting and classifying aphids using computer vision[J]. *Computers and Electronics in Agriculture*, 2020, 169: 105200.
- [12] Khan S, Nadir S, Lihua G, et al. Identification and characterization of an insect toxin protein, Bb70p, from the entomopathogenic fungus, *Beauveria bassiana*, using *Galleria mellonella* as a model system[J]. *Journal of invertebrate pathology*, 2016, 133: 87-94.
- [13] Wen C, Guyer D E, Li W. Local feature-based identification and classification for orchard insects[J]. *Biosystems Engineering*, 2009, 104(3): 299-307.
- [14] 温长吉, 娄月, 张笑然, 等. 基于改进稠密胶囊网络模型的植物识别方法[J]. *农业工程学报*, 2020, 36(8): 143-155.  
Wen Changji, Lou Yue, Zhang Xiaoran, et al. Plant recognition method based on a improved dense CapsNet[J]. *Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE)*, 2020, 36(8): 143-155. (in Chinese with English abstract)
- [15] 张荣, 李伟平, 莫同. 深度学习研究综述[J]. *信息与控制*, 2018, 47(4): 385-397  
Zhang Rong, Li Weiping, Mo Tong. A review of deep learning research[J]. *Information and Control*, 2018, 47(4): 385-397. (in Chinese with English abstract)
- [16] Srdjan S, Marko A, Andras A, et al. Deep neural networks based recognition of plant diseases by leaf image classification[J]. *Computational Intelligence and Neuroence, USA 2016*, 2016: 1-11.
- [17] Amanda R, Kelsee B, Peter M C, et al. Deep learning for image-based cassava disease detection[J/OL]. *Frontiers in Plant Science*, 2017: 1852. <https://doi.org/10.3389/fpls.2017.01852>
- [18] Xie C, Wang R, Jie Z, et al. Multi-level learning features for

- automatic classification of field crop pests[J]. Computers and Electronics in Agriculture, 2018, 152: 233-241.
- [19] Aravind K, Raja P. Disease classification in eggplant using pre-trained VGG16 and MSVM[J]. Scientific Reports, 2020, 10(1): 1-11.
- [20] Wang G, Sun Y, Wang J. Automatic image-based plant disease severity estimation using deep learning[J]. Computational Intelligence and Neuroscience, 2017, 2017:2917536.
- [21] Alfarisy A A, Chen Q, Guo M. Deep learning based classification for paddy pests & diseases recognition[C]//Proceedings of 2018 International Conference on Mathematics and Artificial Intelligence. SiChuan ChengDu 2018: 21-25.
- [22] Rangarajan A K, Purushothaman R. Disease Classification in Eggplant Using Pre-trained VGG16 and MSVM[J]. Scientific Reports, 2020, 10(1): 2322.
- [23] Zhang S, Zhang S, Zhang C, et al. Cucumber leaf disease identification with global pooling dilated convolutional neural network[J]. Computers and Electronics in Agriculture, 2019, 162: 422-430.
- [24] 陈善雄, 伍胜, 于显平. 基于卷积神经网络结合图像处理技术的荞麦病害识别[J]. 农业工程学报, 2021, 37(3): 155-163.  
Chen Shanxiong, Wu Sheng, Yu Xianping. Buckwheat disease recognition using convolution neural network combined with image processing[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2021, 37(3): 155-163. (in Chinese with English abstract)
- [25] 宋余庆, 谢熹, 刘哲. 基于多层 EESP 深度学习模型的农作物病虫害识别方法[J]. 农业机械学报, 2020, 51(8): 196-202.  
Song Yuqing, Xie Xi, Liu Zhe. Crop pests and diseases recognition method based on multi-level EESP model[J]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(8): 196-202. (in Chinese with English abstract)
- [26] Wang J, Li Y, Feng H, et al. Common pests image recognition based on deep convolutional neural network[J]. Computers and Electronics in Agriculture, 2020, 179: 105834.
- [27] 李书琴, 陈聪, 朱彤. 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. 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 with English abstract)
- [28] 甘雨, 郭庆文, 王春桃. 基于改进 EfficientNet 模型的作物害虫识别[J]. 农业工程学报, 2022, 38(1): 203-211.  
Gan Yu, Guo Qingwen, Wang Chuntao. Recognizing crop pests using an improved EfficientNet model[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2022, 38(1): 203-211. (in Chinese with English abstract)
- [29] Zhao X, Li K, Li Y, et al. Identification method of vegetable diseases based on transfer learning and attention mechanism[J]. Computers and Electronics in Agriculture, 2022, 193: 106703.
- [30] Qi J, Liu X, Liu K, et al. An improved YOLOv5 model based on visual attention mechanism: Application to recognition of tomato virus disease[J]. Computers and Electronics in Agriculture, 2022, 194: 106780.
- [31] 唐詹, 柏召, 刁磊, 等. 基于注意力池化和堆叠式结构的病虫害文献识别模型[J]. 农业机械学报, 2021, 52(S1): 178-184.  
Tang Zhan, Bai Zhao, Diao Lei, et al. Diseases and pests articles identification model based on attention pooling and stacked structure[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(S1): 178-184. (in Chinese with English abstract)
- [32] Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need[J]. Advances in Neural Information Processing Systems, 2017, 30: 5998-6008.
- [33] Woo S, Park J, Lee J Y, et al. CBAM: Convolutional block attention module[C]//Proceedings of the European Conference on Computer Vision (ECCV). Zürich: Springer 2018: 3-19.
- [34] Ye Z, Hu F, Liu Y, et al. Associating multi-scale receptive fields for fine-grained recognition[C]//2020 IEEE International Conference on Image Processing (ICIP). Abu Dhabi National Exhibition Center. IEEE, 2020: 1851-1855.
- [35] Wu X, Zhan C, Lai Y K, et al. Ip102: A large-scale benchmark dataset for insect pest recognition[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach, CA, USA 2019: 8787-8796.
- [36] Itti L, Koch C, Niebur E. A model of saliency-based visual attention for rapid scene analysis[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1998, 20(11): 1254-1259.
- [37] Mnih V, Heess N, Graves A. Recurrent models of visual attention[C]. Advances in Neural Information Processing Systems, Montreal, Canada 2014, 2204-2212.
- [38] Hu J, Shen L, Sun G. Squeeze-and-excitation networks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Conference Location: New York, NY, USA 2018: 7132-7141.
- [39] Wang X, Girshick R, Gupta A, et al. Non-local neural networks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Conference Location: New York, NY, USA 2018: 7794-7803.

## Model for the recognition of large-scale multi-class diseases and pests

Wen Changji<sup>1,4</sup>, Wang Qirui<sup>1</sup>, Chen Hongrui<sup>1</sup>, Wu Jianshuang<sup>1</sup>, Ni Jun<sup>2</sup>, Yang Ce<sup>3</sup>, Su Hengqiang<sup>1,4\*</sup>

(1. College of Information and Technology, Jilin Agricultural University, Changchun 130118, China; 2. College of Agriculture, Nanjing Agricultural University, Nanjing 210095, China; 3. University of Minnesota College of Food, Agricultural and Natural Resource Sciences, São Paulo 55108, USA; 4. Institute for the Smart Agriculture, Jilin Agricultural University, Changchun 130118, China)

**Abstract:** Diseases and pests have posed a huge loss to agricultural production in recent years. Food losses that resulted from pests and diseases can be greater than 10% in the world, and even up to 30% in local areas, according to the latest statistics from the World Food and Agriculture Organization (FAO). It is a high demand to early identify the pests and diseases for the early warning, prevention, and control. However, an accurate identification of large-scale pests and diseases still remains a great challenge, due to the wide variety of pests and diseases and trait characteristics, such as high inter-class similarity and high intra-class variability of external morphology. This study aims to effectively extract and characterize the subtle features between categories for the large-scale multi-category pest classification and recognition task. A typical fine-grained classification was first established using the convolutional block attention mechanism, which was a visual attention mechanism module with better performance in the benchmark network. The key features were extracted to represent in both channel and spatial dimensions via blending across the feature channel and spatial domain. The feature dimension information with a high contribution rate to the pest classification task was extracted in the channel domain, while, the location dimension information with a high contribution rate was extracted in the spatial domain, where the benchmark network was achieved a fine-grained differential enhancement and representation. Secondly, the multi-scale feature extraction and representation were also particularly important for the fine-grained classification tasks. Since there were the different sizes and shapes of individuals to be recognized in fine-grained classification tasks, the sensory field scales were fixed for the same layer of convolutional kernels, indicating the influencing feature extraction when the sensory fields did not match the individuals. Therefore, a cross-layer nonlocal module was introduced into the benchmark model, in order to select a deeper layer and multiple shallow layers between multiple feature extraction layers of the benchmark network model. As such, the spatial response relationships were established to learn more multi-scale features for the improved feature extraction and representation capability of the benchmark model. Thirdly, a probability of default (PD)-Net model was also built in a large-scale multi-category fine-grained pest and disease network. A total of 800 images were collected, including the cherry powdery mildew general and severe, grape black rot general and severe, tomato spotted wilt general and severe, as well as citrus yellow dragon disease general and severe. Finally, a pest identification APP was developed using the PD-Net model. A pest and disease identification APP was developed on 61 categories of disease datasets and 102 categories of pests. After that, the 88.617% and 98.922% accuracies were achieved for the Top1 and Top2 recognition on the disease dataset, respectively, the 74.668% and 83.298% accuracy for the Top1 and Top2 recognition on the pest dataset. Respectively, compared with the AlexNet, VGG16, GoogleNet, Inception-v3, and DenseNet121 deep learning models. More importantly, the Top1 and Top2 recognition accuracies were improved by 1.748 percentage points to 4.331 percentage points, and 1.469 percentage points to 6.076 percentage points on the disease dataset, respectively. By contrast, the Top1 and Top2 recognition accuracies were improved by 1.906 percentage points to 8.122 percentage points, and 1.869 percentage points to 6.644 percentage points on the pest dataset, respectively. Meanwhile, the loss value, accuracy, precision, recall rate, and *F1* were all improved, compared with the others. The experimental data was selected to verify the effectiveness of the PD-Net model. Consequently, a fine-grained recognition model can be widely expected for large-scale multi-category pests and diseases.

**Keywords:** models; deep learning; diseases and pests; fine-grained classification; convolutional block attention module; cross-layer non-local module