

胡志伟, 杨华, 黄济民, 等. 基于注意力残差机制的细粒度番茄病害识别 [J]. 华南农业大学学报, 2019, 40(6): 124-132.

HU Zhiwei, YANG Hua, HUANG Jiming, et al. Fine-grained tomato disease recognition based on attention residual mechanism[J]. Journal of South China Agricultural University, 2019, 40(6): 124-132.

基于注意力残差机制的细粒度番茄病害识别

胡志伟¹, 杨 华¹, 黄济民², 谢倩倩²

(1 山西农业大学 信息科学与工程学院, 山西 太谷 030801;

2 武汉大学 计算机学院, 湖北 武汉 430072)

摘要:【目的】解决温室环境下细粒度番茄病害识别方法不足问题。【方法】以早、晚期 5 种番茄病害叶片为研究对象, 提出一种基于注意力与残差思想相结合的新型卷积神经网络模型 ARNet。通过引入多层注意力模块, 层次化抽取病害分类信息, 解决早期病害部位分散、特征难以提取难题; 为避免网络训练出现退化现象, 构建残差模块有效融合高低阶特征, 同时引入数据扩充技术以防止模型过拟合。【结果】对 44 295 张早、晚期病害叶片数据集进行模型训练与测试的结果表明, 与 VGG16 等现有模型相比, ARNet 具有更好的分类表现, 其平均识别准确率达到 88.2%, 显著高于其他模型。ARNet 对早期病害识别准确率明显优于晚期病害, 验证了注意力机制在提取细微区域特征上的有效性, 且在训练过程中未发生过度抖动的状况。【结论】本文提出的模型具有较强鲁棒性和较高稳定性, 在实际应用中可为细粒度番茄病害智能诊断提供参考。

关键词: 细粒度; 注意力机制; 残差网络; 卷积网络; 番茄叶片; 病害识别

中图分类号: S436.36; TP391.4

文献标志码: A

文章编号: 1001-411X(2019)06-0124-09

Fine-grained tomato disease recognition based on attention residual mechanism

HU Zhiwei¹, YANG Hua¹, HUANG Jiming², XIE Qianqian²

(1 College of Information Science and Engineering, Shanxi Agricultural University, Taigu 030801, China;

2 School of Computer Science, Wuhan University, Wuhan 430072, China)

Abstract: 【Objective】To solve the insufficient identification of fine-grained tomato diseases in greenhouse. 【Method】Taking tomato leaves with five early or late diseases as research objects, we proposed a new convolutional neural network model ARNet based on the combination of attention and residual thought. A multi-layered attention module was introduced to solve the problem of early disease location dispersion and the difficulty of feature extraction by extracting hierarchically disease classification information. In order to avoid the degradation of network training, we constructed a residual module to effectively integrate high- and low-order features. Meantime, we introduced the data expansion technology to prevent model over-fitting. 【Result】Model training and testing results of early and late disease leaf datasets with 44 295 pictures showed that ARNet has better classification performance with an average recognition accuracy of 88.2%, which was significantly higher than those of other existing models. In addition, the identification accuracy of ARNet for early disease was significantly better than that for late disease, which verified the effectiveness of attention mechanism in extracting fine region features, and there was no excessive jitter during training process.

收稿日期: 2018-12-25 网络首发时间: 2019-10-28 09:21:00

网络首发地址: <http://kns.cnki.net/kcms/detail/44.1110.s.20191025.0844.034.html>

作者简介: 胡志伟 (1992—), 男, 硕士, E-mail: zhiwei@whu.edu.cn; 通信作者: 杨 华 (1973—), 男, 副教授, 博士, E-mail: yanghua@sxau.edu.cn

基金项目: 国家自然科学基金 (31671571); 国家自然科学基金青年基金 (21803037)

本文模型代码已开源: https://github.com/zhiwei0329/identify_tomato_disease

【Conclusion】 This model proposed in this paper has strong robustness and high stability, and can provide a reference for intelligent diagnosis of fine-grained tomato diseases in practical application.

Key words: fine-grained; attention mechanism; residual network; convolutional network; tomato leaf; disease recognition

病害问题是造成蔬菜品质下降,从而导致农民经济损失的主要原因之一,与日常经济活动关系密切^[1-2]。番茄作为温室栽培主要作物之一,同样存在着病害问题。如何快速准确发现并识别番茄病害,从而采取相应防治措施,对保障作物生产、提高农民收入具有重要意义。不同病害在可见光波段虽呈现出某些症状,但病状信息复杂多变,往往需要经过专业训练的植保专家才能识别诊断^[3]。随着机器视觉和模式识别等技术逐渐成熟,研究人员开始关注采用先进信息技术辅助识别番茄病害,基于机器视觉的番茄病害识别已取得较大进展。Mokhtar 等^[4]利用 Gabor 小波变换技术,结合具有交替核函数的支持向量机,提取番茄叶片图像相关特征,检测和识别其病害类型。Xie 等^[5]结合高光谱图像波长及图像自身纹理特征,对番茄早疫病、晚疫病进行识别。柴阿丽等^[6]提取番茄叶片颜色、纹理及形状参数构建特征提取与判别模型,对其早疫病、晚疫病、叶霜病和棒孢叶斑病进行自动识别。基于计算机视觉的识别方法需针对特定番茄病害人工构建对应病害特征,而实际应用中的病害图像受光照条件及复杂背景等噪声扰动影响,其病害症状特征提取十分复杂,限制了上述方法的应用效果。随着深度学习的发展,卷积神经网络(Convolutional neural network, CNN)通过引入局部连接、权重共享、池化等操作,自动地训练网络抽取数据特征,并已在香蕉^[7]、黄瓜^[8]等作物病害识别上取得较大进展。在番茄病害识别领域, Durmus 等^[9]利用 AlexNet 与 SqueezeNet 卷积网络对包括健康图像在内的 10 类番茄叶片进行训练测试。Fuentes 等^[10]结合 VGG 等深度特征提取器,构建基于区域的卷积神经网络、基于区域的全卷积网络及单镜头多盒检测器训练端到端系统,对 9 种番茄病害进行识别。Brahimi 等^[11]验证深层卷积网络性能优于浅层网络,并得出使用预训练权重初始化模型参数能提升模型性能的结论,在此基础上对 9 种番茄病害区域进行鉴定。上述研究工作仅停留在将各个变种深度卷积网络用于粗粒度番茄病害分类,并未关注病害的发生时期。实际生产中,同一病害的早晚期图像呈现不同特点,若在病害发生的早期对其进行准确识别并采取相应防治措施能极大地降低病害带来的损失。

然而,早期病害发生部位较为隐蔽且感染区域面积也偏小,其占据整幅图像像素比重较低。因此,利用深度卷积网络有效抽取早、晚期番茄病害特征,进行细粒度病害识别问题仍有待解决。

相较于以往方法对全部信息进行处理,注意力机制在处理信息时只会将注意力集中在部分显著或者感兴趣的信息上,并针对不同信息分配不同权重,从而滤除不重要的信息、提升处理效率与模型效果,注意力机制已在自然语言处理^[12-13]与图像分割领域^[14]取得较大进展,但在图像分类领域仍处于探索阶段。Mnih 等提出基于 RNN(Recurrent neural network)的循环注意力模型(Recurrent attention model, RAM),将注意力机制运用于图像分类任务中(<https://arxiv.org/pdf/1406.6247.pdf>)。Wang 等^[15]提出将基于注意力的残差学习引入图像分类领域,解决了前向过程中无法提取注意力的问题,并简化了模型结构,加速了模型训练。Fu 等^[16]将注意力机制用于细粒度图像分类任务中,使用每一层的上一层输出作为当前层的输入,从而形成尺度间循环,能够充分学习细粒度特征信息。注意力机制在图像分类领域已被证明能够有效提取细粒度的特征信息,但在细粒度病害分类问题中的应用仍有待研究。本文创新性地研究细粒度番茄病害识别问题,提出基于注意力与残差机制相结合的新型模型识别方法,对不同严重级别的 5 类病害进行训练及测试,以期为细粒度番茄病害智能诊断系统提供技术支持。

1 材料与方法

1.1 数据采集及预处理

本文模型采用的数据集来自于 2018 全球 AI 挑战赛农作物病害数据集(<https://challenger.ai/dataset/pdd2018>)。首先,筛选其中与番茄相关图片以构建番茄病害数据集,筛选后数据集包含 8 861 幅图片,数据集中所有图片均是在自然光照条件下拍摄得到,且分辨率大小不等。每张图片中仅包含 1 片番茄叶片,且叶片占图片的主要位置。数据集包括番茄白粉病、早疫病、叶霉病、斑枯病和黄化曲叶病共 5 类病害。对于每种病害的样例进一步区分为早期和晚期 2 个级别,区分依据如下。

1) 白粉病早、晚期症状: 早期叶面出现褪绿色小点或白色霉点, 散布于叶片不同区域; 晚期褪绿色小点互相黏合为大小不等的不规则白色粉斑同时表面伴有白色絮状物。

2) 早疫病早、晚期症状: 早期叶部感染出现小黑点并慢慢扩大为同心轮纹斑, 轮斑周边出现黄色或浅绿色晕轮; 晚期茎部分枝处出现褐色至深褐色不规则圆形或椭圆形病斑, 同时伴有灰黑色霉状物。

3) 叶霜病早、晚期症状: 早期叶片表面出现椭圆形或不规则形淡黄色病斑; 晚期病斑发展为灰色或黑紫色的不规则形霉层, 严重时叶片常出现干枯卷缩。

4) 斑枯病早、晚期症状: 早期叶片表面产生水渍状小斑, 随着病情加重出现呈鱼眼状病斑且其周围散生许多黑色小点; 晚期褐色病斑布满全叶, 叶片逐渐枯黄。

5) 黄化曲叶病早、晚期症状: 早期叶片变小呈现褶皱簇状、叶部稍发黄且边缘上卷, 叶厚脆硬; 晚期症状加重, 叶片明显发黄且边缘上卷严重, 叶片明显增厚。

为获取模型有效输入, 对数据集进行以下预处理: 1) 为降低运算量, 提高模型运算速度, 本文将图像统一调整为 256×256 (宽 256 像素, 高 256 像素), 并对不足该像素的部分进行填充、对超出该像素尺寸的区域进行像素压缩; 2) 为丰富数据集, 避免模型训练过拟合, 提高模型泛化能力, 对 8 861 张图片分别进行随机旋转角度 (旋转的最大角度为 30°)、随机水平或垂直平移 (平移的最大距离为 $0.1 \times$ 图像宽或高)、随机缩放、随机水平翻转等操作, 通过引入轻微扰动来实现数据扩充, 并对数据增强操作后出现的像素缺失部位采用邻近像素填充的方式进行补全。

1.2 病害识别模型

本文基于注意力机制与 ResNet^[17-19] 残差网络提出注意力残差网络 ARNet (Attention residual network), 用于番茄叶片的细粒度病害识别。该模型结构如图 1a 所示, 其由注意力卷积块 (Attention convolutional block, ACB) 和残差卷积块 (Residual convolutional block, RCB) 构成。模型网络共包含 5 层, 其中 2~5 层通过上述 2 类模块层叠构成, 并使用上一层的输出作为下一层输入。其中, 层内通过 ACB 块间相互连接构建深度网络结构以抽取多维图像特征信息, 并采用最后一个 ACB 模块的输出作为 RCB 模块的输入; 层间则通过 RCB 块的输出拼接 ACB 块的输出, 以降低图像特征图大小。最终, 模型采用 2 层全连接网络实现对输入病害图像

的类别划分。ARNet 中第 2 层卷积通道数为 256, 并随着网络层数的递增, 卷积通道数翻倍, 直至第 5 层达到 2 048。

模型中的 ACB 注意力卷积模块结合注意力机制捕获丰富的语境关联, 以解决图像分类问题, 其具体操作如图 1b 所示。深层网络中有着丰富的语义信息, 可用于引导浅层网络信息的选择, 因而可得到更为精准的分辨率信息, ACB 块可利用全局池化提供的信息作为指引引导选择浅层特征。具体经历以下 3 个阶段: 1) 浅层网络输入通过 3 次 3×3 卷积、批归一化以及激活函数操作获得中间层输出, 经操作①将浅层输入与中间层输出叠加作为步骤 2) 输入; 2) 对 1) 结果经全局池化与 2 步 1×1 卷积、批归一化以及 Relu 与 Sigmoid 激活函数操作, 得到中间层输出权重信息, 经操作②将权重信息与浅层输入相乘得到加权后的浅层输入信息; 3) 最后经操作③将中间层输出与步骤 2) 的结果叠加, 得到 ACB 模块输出。

此外, 本文模型基于 ResNet 残差网络构建了 RCB 残差卷积模块。随着卷积神经网络深度的增加, 模型的学习存在退化问题, 即准确率在训练集上出现饱和甚至下降的现象。为解决上述问题, 本文结合文献 [17-19] 提出残差网络, 以保证在增加网络层数的同时提升网络性能, 其具体结构如图 1c 部分所示, 该模块在传统网络结构基础上增加跳跃信息。图 2 展示了传统深层网络与加入残差单元深层网络间的操作差别。其中, 图 2a 表示简单深层网络的操作, 其仅仅将卷积、批归一化以及激活函数操作进行简单堆叠; 图 2b 的残差卷积模块对输入信息进行了 2 路卷积操作, 左路卷积操作等同于传统深层网络, 而右路卷积块为另一浅层网络, 并最终叠加 2 路卷积输出信息后经过 Relu 激活函数生成当前残差卷积模块的输出。

本文残差卷积模块可由式 (1) 和式 (2) 表示:

$$l_i = G(L_i) + F(L_i, W_i), \quad (1)$$

$$L_{i+1} = f(l_i). \quad (2)$$

式中, L_i 与 L_{i+1} 表示残差模块的输入与输出, $G(L_i)$ 表示对输入 L_i 的恒等映射, 即图 2b 中的右路卷积操作。通常情况下 $G(L_i)$ 等于 L_i , 本文 RCB 模块中将其设为对输入 L_i 进行 1 次 3×3 卷积操作的结果, $F(\bullet)$ 表示常规卷积操作, W_i 表示卷积操作中的卷积核权重参数, l_i 表示残差操作后未经激活函数处理的输出内容, $f(\bullet)$ 表示激活函数, 本文选取 Relu 函数。

一般状态下, RCB 块同时更新 W_i 与 $G(L_i)$ 参数。

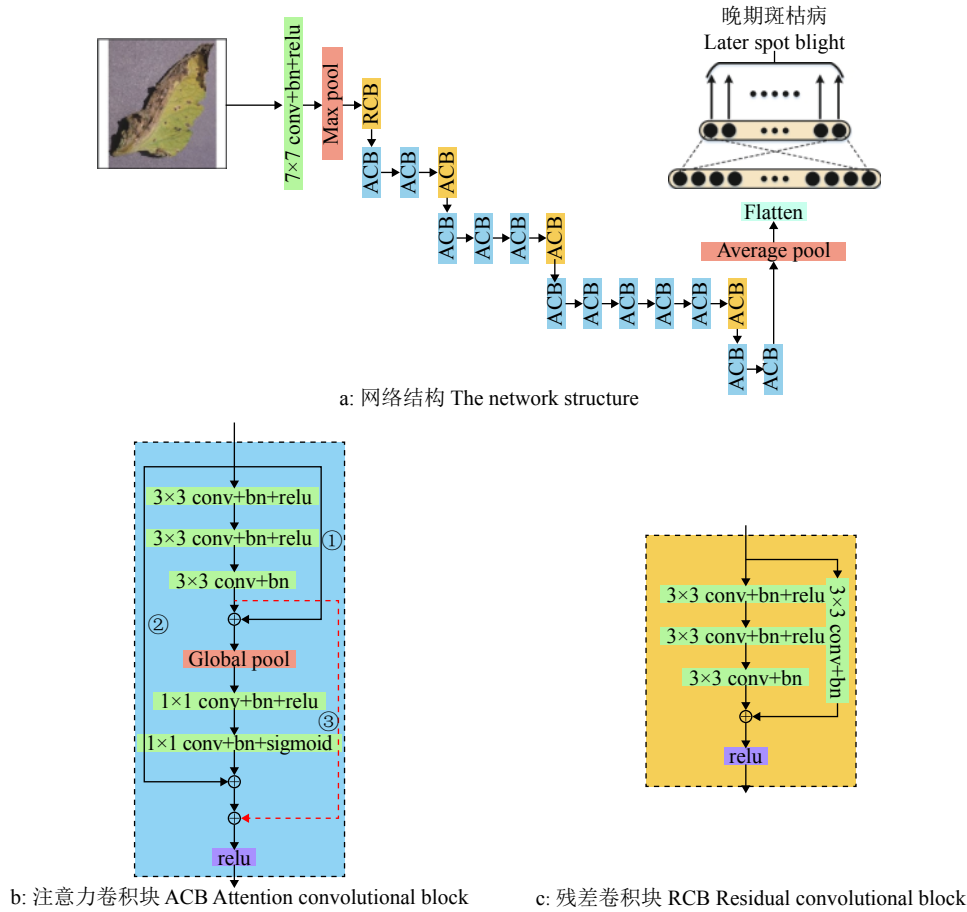


图 1 ARNet 模型结构图

Fig. 1 The structure of ARNet model

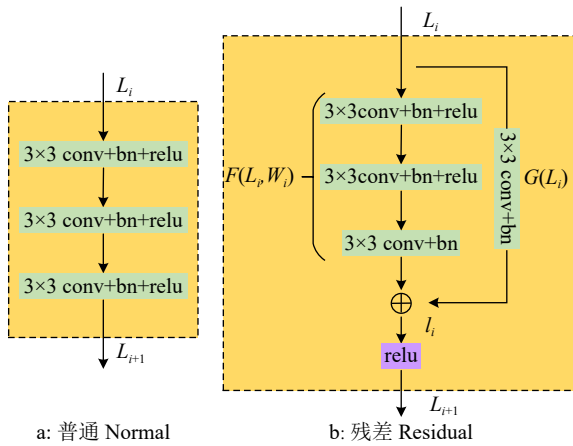


图 2 RCB 模块引入前后对照操作

Fig. 2 The control operations before and after introduction of RCB module

而当学习已经达到较饱和状态时, 学习目的转化为仅对 $G(L_i)$ 恒等映射权重的学习, 使输入 L_i 近似转化为输出 L_{i+1} , 以保证后续层不会因为当前层的存在而造成精度的下降。该残差式跳跃结构使得某一层的输出能够直接跨过中间几层作为后面某一层的输入, 解决了叠加多层网络的学习退化问题, 同时降低了模型的分错误率。

1.3 模型评价指标

本文采用图像分类领域公认的平均准确率 (\bar{A}) 评价指标评估模型分类结果, 其定义如式 (3) 所示^[20]:

$$\bar{A} = \frac{1}{n_c} \sum_{i=1}^{n_c} \frac{n_{ii}}{n_i}, \quad (3)$$

式中, n_c 表示样本类别总数, 本文取 10; i 为类别标签, 本文取 1~10; n_i 表示类别为 i 的样本总数; n_{ii} 表示类别 i 正确预测为第 i 类的样本总数。

1.4 试验参数设置

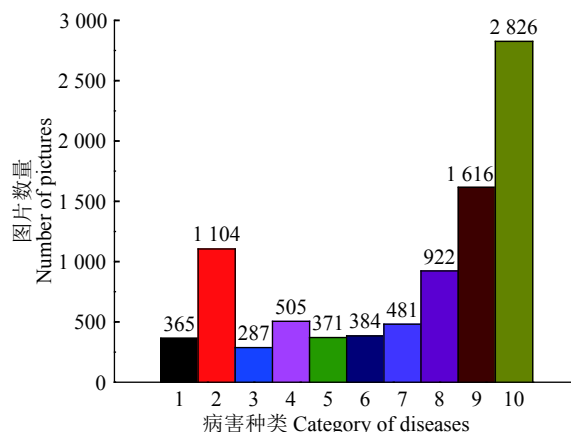
本文将数据集分为训练集、验证集和测试集 3 个部分, 为避免内存溢出, 采取批训练方式对 ARNet 模型与 VGG16(<https://arxiv.org/pdf/1409.1556.pdf>)、InceptionV3^[21]、Xception^[22]、MobileNetV2^[23]、ResNet34 在训练集和验证集进行对比试验。每个批次分别训练 16 张图片, 而验证集批大小与训练集批大小一致。本文设置迭代轮数为 50, 并采用 Keras 中多分类对数损失函数 (Categorical_crossentropy) 作为代价函数。为解决反向传播过程中梯度消失和爆炸问题, 引入批规范化 (Batch

normalization), 对网络隐藏层输入进行标准化。为加快模型收敛速度以获得最优模型参数, 本文同时迁移学习已训练好的部分模型参数作为对应模型初始权重值。此外, 为防止过拟合, 引入 Keras 中的早停 (Early stopping) 机制, 本文设置当验证集的损失在最近 10 迭代过程中不再减小时停止训练。为提高调参效率, 本文采用自适应矩估计算法 (Adaptive moment estimation, Adam)(<https://arxiv.org/pdf/1412.6980.pdf>) 优化模型, 其初始学习率设置为 0.001。为自动修改学习率, 引入 Keras 的学习率调度器 (Reduce LR on plateau), 本文设置当验证集的损失函数值在 5 轮迭代过程中均未降低时, 将学习率降低到当前学习率的 0.9 倍。为保存最优模型参数, 每轮迭代完成之后, 通过观察验证集损失函数值是否降低以决定是否保存当前模型。最终将保存的模型结构及参数用于测试集病害图片类别预测。

2 结果与分析

2.1 番茄病害数据集及预处理结果

筛选后得到的番茄病害数据集包含 8 861 幅图片, 番茄叶片 5 种病害早、晚期的病害分类具体分布如图 3 所示。由图 3 可见, 各类别的样本数目分布不均衡, 这可能导致训练模型时对于不同类别存



1: 早期白粉病; 2: 晚期白粉病; 3: 早期早疫病; 4: 晚期早疫病; 5: 早期叶霉病; 6: 晚期叶霉病; 7: 早期斑枯病; 8: 晚期斑枯病; 9: 早期黄化曲叶病; 10: 晚期黄化曲叶病

1: Early powdery mildew; 2: Later powdery mildew; 3: Early early blight; 4: Later early blight; 5: Early leaf frost disease; 6: Later leaf frost disease; 7: Early spot blight; 8: Later spot blight; 9: Early yellow flower curly leaf disease; 10: Later yellow flower curly leaf disease

图 3 番茄病害初始数据集分布

Fig. 3 The distribution of initial dataset for tomato disease

在偏差。本文构建的数据集能够进一步检测模型的训练稳定性。数据集经“1.1”中 2 步预处理后, 最终扩充为 44 295 幅图片, 图 4 为部分数据预处理结果。按照 8:1:1 的比例随机将其分为训练集、验证集与测试集, 其中, 训练集大小为 35 436 幅、验证集大小为 4 429 幅、测试集大小为 4 430 幅。

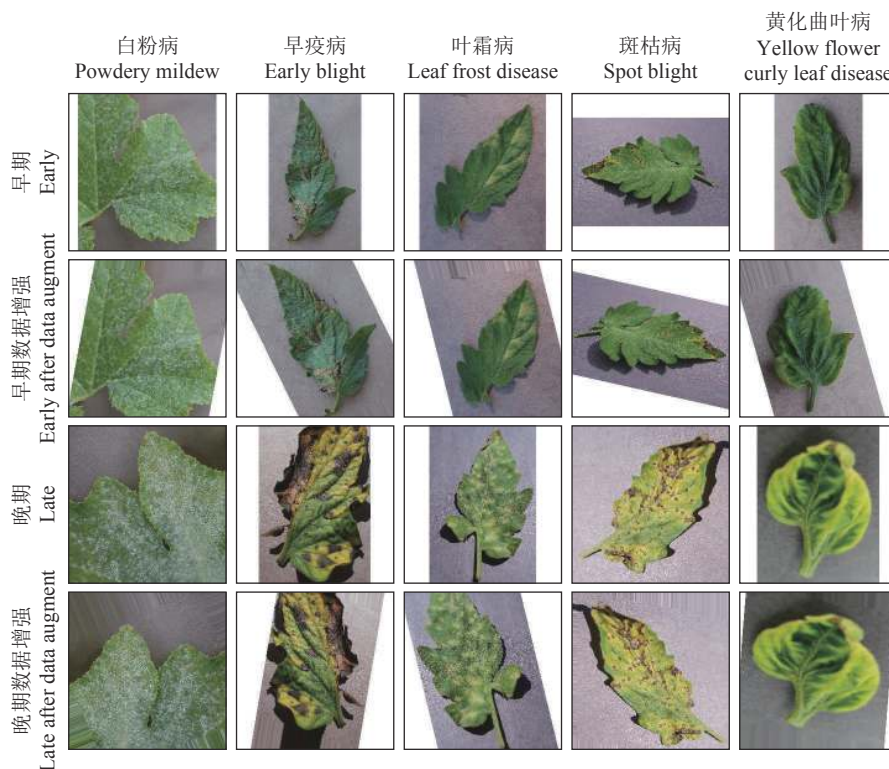


图 4 部分数据预处理前后

Fig. 4 The partial data before and after preprocessing

2.2 模型不同层 ACB 模块可视化

可视化 ARNet 模型不同层 ACB 模块所获取的注意力信息如图 5 所示。图 5a 和 5b 分别表示晚期早疫病、晚期叶霜病 2 类病害图片在 ARNet 模型不同层级 ACB 模块的输出热度图。其中, 每类图片的第 1 行表示输出热度、第 2 行展示热度在原图上的叠加, 从左至右依次为第 2 层、第 3 层、第 4 层和第 5 层

最后一个 ACB 模块的输出。从图 5 可看出, 在引入注意力机制后, ACB 模块均能较为准确地提取每类病害的关键特征; 虽然浅层 ACB 模块提取到的特征较为分散, 不足以作为类别划分依据, 但深层能够提取更为集中的特征, 其颜色也更近于红色, 表示对应部位对最终分类决策贡献更大。

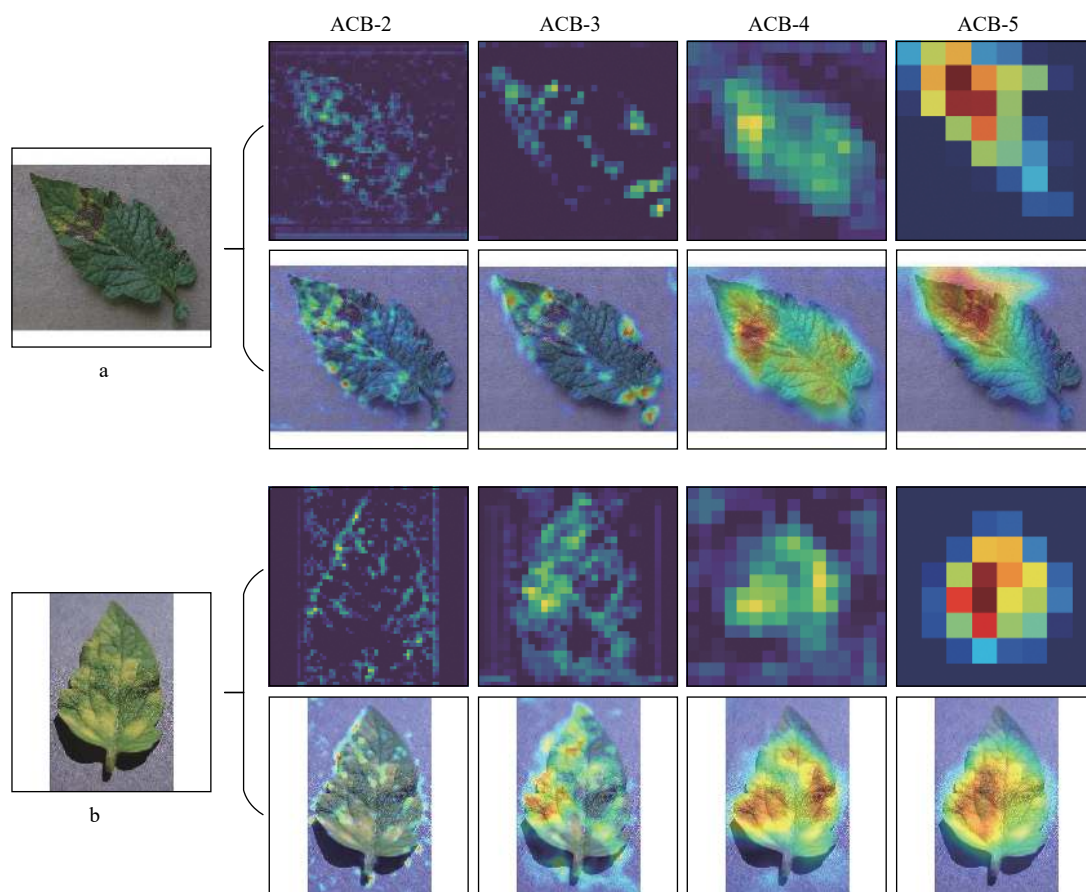


图 5 各层最后 ACB 模块热力图输出

Fig. 5 The thermodynamic diagram output of final ACB module at each layer

2.3 模型分类的准确率

不同模型采用相同的试验参数; 为加快模型收敛速度, VGG16、InceptionV3 和 Xception 这 3 种模型分别迁移已经在 ImageNet^[24] 上训练好的参数, 作为对应模型的初始化权重值, 并在训练过程中固定除最后全连接层外的其余层; 而 MobileNetV2、ResNet34 和 ARNet 这 3 种模型均采用 Glorot 均匀分布方法初始化权重参数。

表 1 展示了测试集中每种病害类别在不同模型上的预测准确率, 其中, 最后一行表示在整个测试集上的准确率。表 1 的结果表明: 1) 本文提出的 ARNet 模型对早、晚期 5 种番茄病害均能取得最好的分类准确率。与其他 5 种模型相比, ARNet 模型的分类准确率提高了 3.2%~12.9%; 具体到单个类别

上, ARNet 模型在早期白粉病类别上性能优势最大。2) ARNet 模型对早期病害类别识别准确率优势明显高于晚期病害。其在除早期黄化曲叶病外的其余所有早期病害类别上均能取得最优分类性能, 这是因为即使早期病害部位隐蔽、辨识度低, 但 ARNet 模型中的 ACB 模块仍然能够在训练中分层自学习注意力信息, 对有效病害特征区域赋予较高权重, RCB 模块则能够将低层获取到的病害特征通过跳跃机制有效传入高层, 引导高层进行特征选择, 进而充分提取出细微关键病害特征, 极大地提升模型在病害特征不明显时的识别能力。此外, 在早期黄化曲叶病上虽然未能取得最佳效果, 但也能得到有竞争力的分类准确率, 仅仅比最优模型低 0.4 个百分点。

表 1 不同模型对番茄病害的分类准确率¹⁾

Table 1 Classification accuracy of tomato diseases by different models

病害 Disease	时期 Stage	VGG16	InceptionV3	Xception	MobileNetV2	ResNet34	ARNet
白粉病 Powdery mildew	早期 Early	0.176	0.191	0.197	0.521	0.521	0.638
	晚期 Late	0.996	0.978	0.972	0.939	0.967	0.967
早疫病 Early blight	早期 Early	0.647	0.571	0.541	0.797	0.872	0.940
	晚期 Late	0.909	0.683	0.706	0.794	0.833	0.865
叶霜病 Leaf frost disease	早期 Early	0.734	0.814	0.824	0.830	0.920	0.920
	晚期 Late	0.773	0.432	0.492	0.589	0.627	0.843
斑枯病 Spot blight	早期 Early	0.573	0.300	0.496	0.712	0.838	0.888
	晚期 Late	0.938	0.913	0.825	0.877	0.931	0.948
黄化曲叶病	早期 Early	0.715	0.693	0.713	0.830	0.830	0.826
Yellow flower Curly leaf disease	晚期 Late	0.911	0.868	0.855	0.804	0.856	0.887
平均值 Average		0.817	0.753	0.758	0.807	0.850	0.882

1)加框数据表示对应列模型在对应行病害中的最佳准确率

1)Framed data indicated the best accuracy of corresponding column model in the corresponding row disease

2.4 模型稳定性

图 6 展示了各模型在 50 轮迭代过程中验证集上准确率与损失函数变化情况。图 6 的结果表明,不同模型在迭代过程中验证准确率变化呈现不同特点。除 ARNet 外的 5 种模型在训练过程中均触发了 EarlyStopping 机制,即出现连续 10 轮迭代中准确率虽仍在提升但损失函数值未下降现象。具体来讲, VGG16、InceptionV3 以及 Xception 均在第 12 轮迭代时触发该机制,而 MobileNetV2 以及 ResNet34 分别在第 16 和 18 轮触发。且这 5 种模型在中途部分迭代过程中其准确率均出现抖动现象,而对于 ARNet 模型,虽在初始迭代时其验证准确率较低,但随着训练的深入,准确率一直呈现上升趋势,且在第 37 轮超越了其余 5 种模型的最大准确率,说明 ARNet 模型相较于其他对比模型具有较高稳定性,能够稳步提升模型准确率。这得益于其网络结构中 ACB 模块的存在,通过对整幅图片的不同位置引入注意力,可以使得模型尽量减少不

必要区域特征的学习,增大有效病害部位对分类决策的贡献,滤除噪声区域,避免过多学习训练集上无用特征而出现验证集准确率下降现象。综合对比准确率与损失函数变化图可得,训练过程中即使验证准确率仍处于上升态势,但验证损失函数却可能出现不减反增的现象,此时模型发生过拟合。这种现象容易出现在迁移大数据集上已训练好的权重参数的模型上,如图 6 中前 3 种模型,说明迁移学习权重值仅仅能加速对应模型的收敛速度,但对模型准确率提升贡献度不大。这主要是因为深度卷积网络通过训练可获得分层特征表达,模型低层能够获取诸如边缘、颜色等低级语义特征,而这些特征在不同任务中是不变的,因而可通过将在大数据集上训练好的模型参数迁移至新模型,进而优化模型学习效率,不至于使模型从零开始学习,从而加速收敛。但这也只是将大量共性特征有效迁移过来,而未能获取特定任务数据集的独特特征,因而对其预测准确率的贡献度并不会很大。

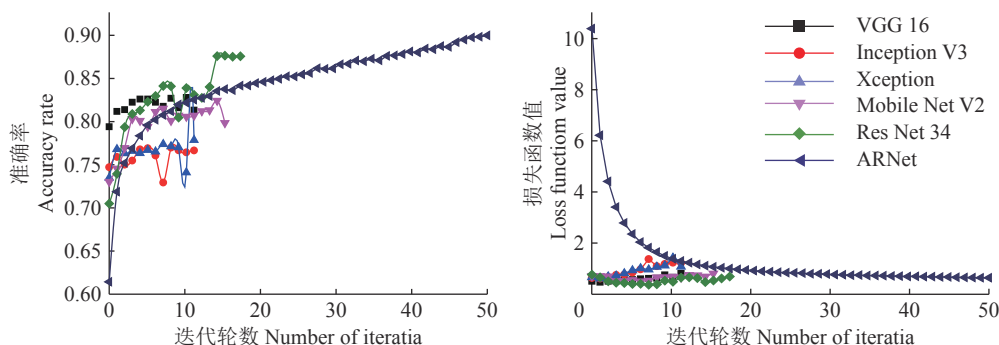


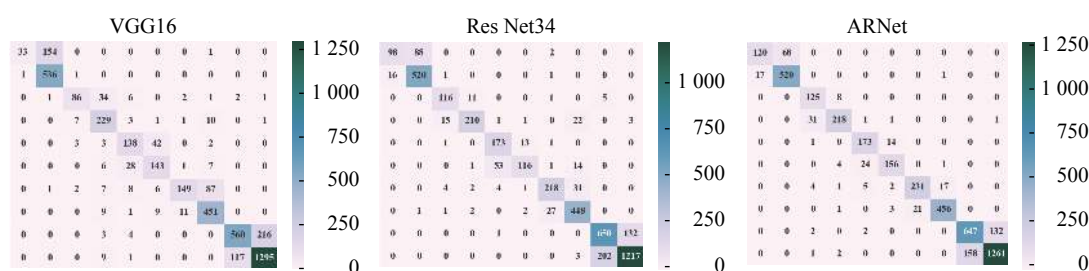
图 6 不同模型迭代过程中验证集准确率与损失函数变化

Fig. 6 The changes of verification set accuracy and loss function in the iterative process of different models

2.5 模型预测的结果

为进一步观察各模型在预测集上的表现,本文可视化展示了 VGG16、ResNet34 和 ARNet 这 3 种模型在测试集上的混淆矩阵。如图 7 所示,图中每列代表预测类别,其总数表示该类别的预测数据;每行表示数据的真实类别,其总数表示属于该类别的真实数据;行列交叉处的数值表示数据被预测为对应行类别的数目,而对角线所处位置处的值为预测正确的结果总和。图 7 的结果表明:3 种模型在不同类别病害分类上表现差异较大。VGG16 模型能够在晚期黄化曲叶病类别上取得最佳准确率,但对早期白粉病类别预测效果最差,甚

至出现预测正确的数目远远小于错误的情况。而 ARNet 的预测效果最理想,在同类别预测中效果均不逊于其他 2 种模型。对于早、晚期相同病害预测时,容易出现预测类别正确,但病害时期预测错误的情况,即早期病害被预测成晚期或者相反。这是因为不同时期的同一病害,其在图像上存在共性特征且病害的严重程度不存在统一的标准,无法明确早期与晚期间的界限,若现实中仅仅只是识别某种病害类别而不区分其病害时期,3 种模型分类准确率均较高,且 ARNet 模型最高,准确率达到 99.26%,说明 ARNet 模型同样可为粗粒度番茄病害识别提供参考。



行号 1~10 分别表示早期白粉病、晚期白粉病、早期早疫病、晚期早疫病、早期叶霜病、晚期叶霜病、早期斑枯病、晚期斑枯病、早期黄化曲叶病以及晚期黄化曲叶病,列号 1~10 类别与其对应编号的行号类别一致

Line numbers 1 to 10 indicate early powdery mildew, later powdery mildew, early early blight, later early blight, early leaf frost disease, later leaf frost disease, early spot blight, later spot blight, early yellow flower curl leaf disease, and in the case of later yellow flower curl leaf disease, the column numbers 1 to 10 are consistent with the row number category corresponding to the corresponding number

图 7 3 种模型测试集混淆矩阵

Fig. 7 The confusion matrixes of test sets of three models

3 结论

本文基于注意力与残差机制构建 ARNet 模型对番茄病害进行细粒度识别,并将其与 VGG16、InceptionV3、Xception、MobileNetV2、ResNet34 等 5 种模型进行对比,就其在分类准确率以及模型预测等方面进行详细分析,得出以下结论:

1) ARNet 可以较好地自动提取病害的关键特征,具有较高的分类性能,平均识别准确率达到 88.2%,比其他模型提高了 3.2~12.9 个百分点,且在早期白粉病类别上性能优势最大,预测准确率达 63% 以上;

2) ARNet 模型对早期病害类别识别准确率优势明显高于晚期病害,说明 ARNet 模型可用于对早期病害进行高精度识别,便于及早发现番茄病害以采取防治措施;

3) 虽然通过数据扩充在一定程度上破坏了原始数据集的样本分布,但 ARNet 在迭代过程中能在提高验证准确率的同时避免过拟合现象的发生,且未发生过度抖动的状况,说明 ARNet 模型具有较强

鲁棒性和较高稳定性。

4) 对于不同时期同种病害预测时,容易出现病害类别预测正确但病害时期预测错误的情况;但如果仅关注病害类别而不考虑其发生时期,则 ARNet 准确率更高,可为粗粒度番茄病害识别提供参考。

目前模型在训练中通过数据扩充方式引入大量模拟数据,且局限于背景简单、病害单一的番茄病害叶片图像,不能充分识别同一叶片上的多种病害。今后应进一步丰富番茄病害图像数据集,以充分提取病害症状的多尺度、多维度特征,并建立端到端识别模型,进一步提升模型的识别准确率与实用性。

参考文献:

- [1] BAI X, LI X, FU Z, et al. A fuzzy clustering segmentation method based on neighborhood grayscale information for defining cucumber leaf spot disease images[J]. *Comput Electron Agric*, 2017, 136: 157-165.
- [2] 王翔宇,温皓杰,李鑫星,等. 农业主要病害检测与预警技术研究进展分析[J]. *农业机械学报*, 2016, 47(9): 266-277.
- [3] 龙满生,欧阳春娟,刘欢,等. 基于卷积神经网络与迁移学习的油茶病害图像识别[J]. *农业工程学报*, 2018,

- 34(18): 194-201.
- [4] MOKHTAR U, ALI M A S, HASSENIAN A E, et al. Tomato leaves diseases detection approach based on support vector machines[C]//IEEE. Computer Engineering Conference (ICENCO). Egypt: IEEE, 2015: 246-250.
- [5] XIE C, SHAO Y, LI X, et al. Detection of early blight and late blight diseases on tomato leaves using hyperspectral imaging[J]. *Sci Rep*, 2015, 5: 16564.
- [6] 柴阿丽, 李宝聚, 石延霞, 等. 基于计算机视觉技术的番茄叶部病害识别[J]. *园艺学报*, 2010, 37(9): 1423-1430.
- [7] AMARA J, BOUAZIZ B, ALGERGAWY A. A deep learning based approach for banana leaf diseases classification[M]// MITSCHANG B. Lecture Notes in Informatics (LNI). Bonn : Gesellschaft Für Informatik, 2017: 79-88.
- [8] 马浚诚, 杜克明, 郑飞翔, 等. 基于卷积神经网络的温室黄瓜病害识别系统[J]. *农业工程学报*, 2018, 34(12): 186-192.
- [9] DURMUS H, GUNES E O, KIRCI M. Disease detection on the leaves of the tomato plants by using deep learning[C]//IEEE. 2017 6th International Conference on Agro-Geoinformatics. Fairfax VA: IEEE, 2017: 1-5.
- [10] 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): 2022.
- [11] BRAHIMI M, BOUKHALFA K, MOUSSAOUI A. Deep learning for tomato diseases: Classification and symptoms visualization[J]. *Appl Artif Intell*, 2017, 31(4): 299-315.
- [12] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[C]//Curran Associates. Advances in Neural Information Processing Systems(NIPS). New York: Curran Associates, 2017: 5998-6008.
- [13] LETARTE G, PARADIS F, GIGUERE P, et al. Importance of self-attention for sentiment analysis[C]// Proceedings of the 2018 EMNLP Workshop BlackboxNLP: Analyzing and Interpreting Neural Networks for NLP. Brussels: Association for Computational Linguistics. 2018: 267-275.
- [14] YU C, WANG J, PENG C, et al. BiSeNet: Bilateral segmentation network for real-time semantic segmentation [C]//Springer. Proceedings of the European Conference on Computer Vision (ECCV). Munich: Springer, 2018: 325-341.
- [15] WANG F, JIANG M, QIAN C, et al. Residual attention network for image classification[C]//IEEE. IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu: IEEE, 2017: 3156-3164.
- [16] FU J, ZHENG H, MEI T. Look closer to see better: Recurrent attention convolutional neural network for fine-grained image recognition[C]//IEEE. IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu: IEEE, 2017: 4476-4484.
- [17] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]//IEEE. IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas: IEEE, 2016: 770-778.
- [18] HE K, ZHANG X, REN S, et al. Identity mappings in deep residual networks[C]//Springer. Proceedings of the European Conference on Computer Vision (ECCV). Amsterdam : Springer, 2016: 630-645.
- [19] ZHANG K, SUN M, HAN T X, et al. Residual networks of residual networks: Multilevel residual networks[J]. *IEEE T Circ Syst Vid*, 2018, 28(6): 1303-1314.
- [20] 孙俊, 谭文军, 毛罕平, 等. 基于改进卷积神经网络的多种植物叶片病害识别[J]. *农业工程学报*, 2017, 33(19): 209-215.
- [21] SZEGEDY C, VANHOUCKE V, IOFFE S, et al. Rethinking the inception architecture for computer vision[C]//IEEE. IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas: IEEE, 2016: 2818-2826.
- [22] CHOLLET F. Xception: Deep Learning with Depthwise Separable Convolutions[C]//IEEE. IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu: IEEE, 2017: 1800-1807.
- [23] SANDLER M, HOWARD A, ZHU M, et al. MobileNetV2: Inverted residuals and linear bottlenecks[C]//IEEE. IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Salt Lake City: IEEE, 2018: 4510-4520.
- [24] RUSSAKOVSKY O, DENG J, SU H, et al. Imagenet large scale visual recognition challenge[J]. *Int J Comput Vision*, 2015, 115(3): 211-252.
- 【责任编辑 周志红】