

李晓振, 徐 岩, 吴作宏, 等. 基于注意力神经网络的番茄叶部病害识别系统[J]. 江苏农业学报, 2020, 36(3): 561-568.
doi: 10.3969/j.issn.1000-4440.2020.03.005

基于注意力神经网络的番茄叶部病害识别系统

李晓振, 徐 岩, 吴作宏, 高 照, 刘 林

(山东科技大学电子信息工程学院, 山东 青岛 266590)

摘要: 基于注意力机制的卷积神经网络构建了番茄叶部病害识别系统。依据注意力机制构建并行注意力模块以提升特征提取能力, 并与残差结构相结合构建 PARNet 模型。以分别患有早疫病、晚疫病、叶霉病、斑枯病和花叶病毒病这 5 类病害的叶片和健康叶片的叶部图像为研究对象, 将 PARNet 模型与 VGG16、ResNet50、SeNet 等模型相对比, 结果显示 PARNet 模型的识别率为 96.91%, 高出其他模型 2.25%~11.58%。各类预测结果的精确率平均为 96.84%。最后使用 Flask 完成 WEB 应用程序的开发, 实现了跨平台的番茄叶部病害识别。

关键词: 番茄; 叶部病害; 注意力机制; 并行池化; WEB 界面

中图分类号: TP391 **文献标识码:** A **文章编号:** 1000-4440(2020)03-0561-08

Recognition system of tomato leaf disease based on attentional neural network

LI Xiao-zhen, XU Yan, WU Zuo-hong, GAO Zhao, LIU Lin

(School of Electronic Information Engineering, Shandong University of Science & Technology, Qingdao 266590, China)

Abstract: The convolutional neural network based on attention mechanism was proposed to construct tomato leaf disease recognition system. According to the attention mechanism, a parallel attention module was constructed to improve the feature extraction ability, and combined with the residual structure to construct a PARNet model. The images of healthy leaves and leaves suffering from early blight, late blight, leaf mold, leaf blight and mosaic virus were used as research objects. Compared with other models such as VGG16, ResNet50 and SeNet the recognition rate of the PARNet model was increased by 2.25%~11.58%. The accuracy rate of various prediction results was 96.84% on average. Flask was used to complete the development of WEB application program, and the cross-platform tomato leaf disease recognition was realized.

Key words: tomato; leaf disease; attention mechanism; parallel pooling; WEB interface

番茄果实营养丰富, 含有丰富的胡萝卜素、维生素(主要为维生素 C 和 B 族维生素)、番茄红素和矿物质。除直接食用外, 番茄还被广泛用来加工成番

茄酱、番茄汁等。然而在番茄的种植过程中, 各种病害严重制约着番茄的生产^[1], 常见的有晚疫病、早疫病、叶霉病、花叶病毒病和斑枯病等。

近年来, 机器学习技术发展迅速, 人们开始尝试机器视觉技术在作物病害识别中的各种应用。传统机器视觉算法需要根据目标任务和先验知识选择合适的特征, 这些特征通常包括图像的颜色、形状、纹理等, 用来提取这些特征的特征提取器需要人为构建, 较为繁琐和困难, 而且由此设计出的特征提取器还缺乏泛化能力^[2-4]。而卷积神经网络方法可以根据数据及其类别自行调整权值参数以构建合适的特

收稿日期: 2020-03-23

基金项目: 国家自然科学基金项目(11547037、11604181); 山东省研究生教育创新计划项目(01040105305); 海信(山东)冰箱有限公司研发中心资助课题

作者简介: 李晓振(1994-), 男, 山东临沂人, 硕士研究生, 研究方向为计算机视觉、深度学习和人工智能等。(E-mail) lxz201019@126.com

通讯作者: 徐 岩, (E-mail) xuyan@sdust.edu.cn

特征提取器,相对高效便捷,构建的特征提取器还具备更优异的泛化能力,可以有效克服传统方法的缺陷。LeNet^[5]的出现奠定了现代卷积神经网络的基础。而 AlexNet 的优异性能引发了深度学习的研究热潮,各种卷积神经网络模型不断提出,如 VGG 模型、GoogleNet 模型和 ResNet 模型等^[6-8]。

注意力机制在自然语言处理^[9]领域进展迅速,但在图像分类领域注意力机制仍处于探索阶段。Fu 等^[10]将 VGG 模型与注意力机制相结合应用到图像分类任务中。Hu 等将注意力机制与捷径连接相结合构建残差注意力模型 SeNet^[11],但该模型仅采用了单一池化操作,故仅能根据纹理特征区分特征图重要度,而且注意力是作用于输出特征图与输入特征图的残差上,不利于区分特征通道的重要度。

为实现对番茄健康叶片以及早疫病、晚疫病、叶霉病、斑枯病或花叶病毒病叶片的分类识别,本研究设计一种并行注意力模块并与捷径连接相结合构建 PARNet 模型,再采用 Flask 开发 WEB 界面,实现多平台下番茄叶部病害识别。

1 数据与方法

1.1 数据简介

番茄叶部病害图像主要来自 Plant Village Dataset^[12]数据集和互联网。在去除异常图像后,番茄健康叶片和患有早疫病、晚疫病、叶霉病、斑枯病和花叶病毒病的叶片图像数量分别为 395、397、397、389、342、364 张,共 2 284 张图像。

番茄叶部图像经裁剪、旋转等方式处理后,数量扩增到 9 136 张,按照 7:2:1 的比例随机划分训练集、验证集和测试集。各类番茄叶部病害图像如图 1 所示。

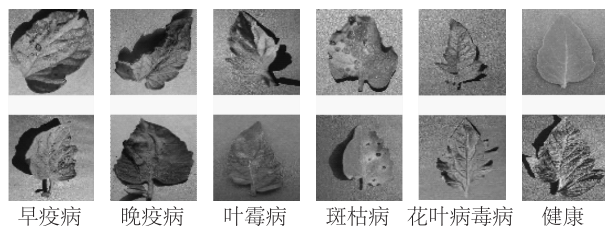


图 1 番茄叶部病害图像示例

Fig.1 Image of tomato leaf diseases

1.2 并行注意力模块的构建

图像信息经过卷积和池化操作得到的特征图通

常缺乏对各个通道重要性的区分,但实际上这些特征图中各个通道的重要程度是不一致的,因此引入权值来表明各个特征通道的重要性,以提升和目标相关的信息的表达,抑制其余无用信息的表达。

构建并行池化的注意力卷积模块 (Parallel pooled attention convolution module, PPA) (图 2)。PPA 模块包括由 2 种全局池化层、1 个多层感知器、1 个激活函数和 1 个捷径连接组成的注意力模块。

采用捷径连接后,设 PPA 模块的输入为 x ,其卷积输出为 $F(x)$,最终输出为结果为 $H(x) = F(x) + x$ 。根据反向传播的链式法则可以得到:

$$\frac{\partial \varepsilon}{\partial x_l} = \frac{\partial \varepsilon}{\partial x_L} \cdot \frac{\partial X_L}{\partial x_l} = \frac{\partial \varepsilon}{\partial x_L} \left[1 + \frac{\partial}{\partial x_l} \sum_{i=1}^{L-1} F(x_i, W_i) \right] \quad (1)$$

梯度分为 2 个部分:不通过权重层传递的 $\frac{\partial \varepsilon}{\partial x_L}$ 和

通过权重层传递的 $\frac{\partial \varepsilon}{\partial x_L} \left(\frac{\partial}{\partial x_l} \sum_{i=1}^{L-1} F \right)$ 。其中 $\frac{\partial \varepsilon}{\partial x_L}$ 保证了梯度信号能够直接传回到任意浅层 x_L ,避免了梯度消失带来的退化问题。

若在残差结构后执行注意力机制,则先从输入中提取特征,再判断特征的重要性,且引入的激活函数会作用于不通过权重层传递的 $\frac{\partial \varepsilon}{\partial x_L}$,干扰捷径连接通路反向传播,导致梯度不断被缩小进而发生退化问题。

PPA 模块先根据注意力机制划分输入信息的重要度再提取特征,不但更符合视觉注意力机制的逻辑,还有利于减少参数量(卷积层输出的特征图数量比输入的更多)。其次,PPA 模块同时采用 2 种池化结构,可分别提升注意力模块对特征图的背景特征和目标(病害)特征的识别,进而更好地区分特征图的重要度。相较于将 2 种池化操作的输出通过同一个 MLP 后再求和(每个神经元要先后根据 2 种特征值调整,容易顾此失彼且易倾向针对于后输入的特征值进行调整),PPA 模块先在通道维度上拼接 2 组池化值再经过 MLP 得到通道重要度,可以有效提升 MLP 的训练效果。最后,简单堆叠注意力模块,会在 (0,1) 范围内对特征值进行重复加权,造成特征值的消减,而捷径连接将注意力模块的输入作为输出的初始值,可避免此问题。而且捷径连接可使注意力模块学习为恒等映射,使得网络在训练过程中,可自行“开关”注意力模块,保障注意力模

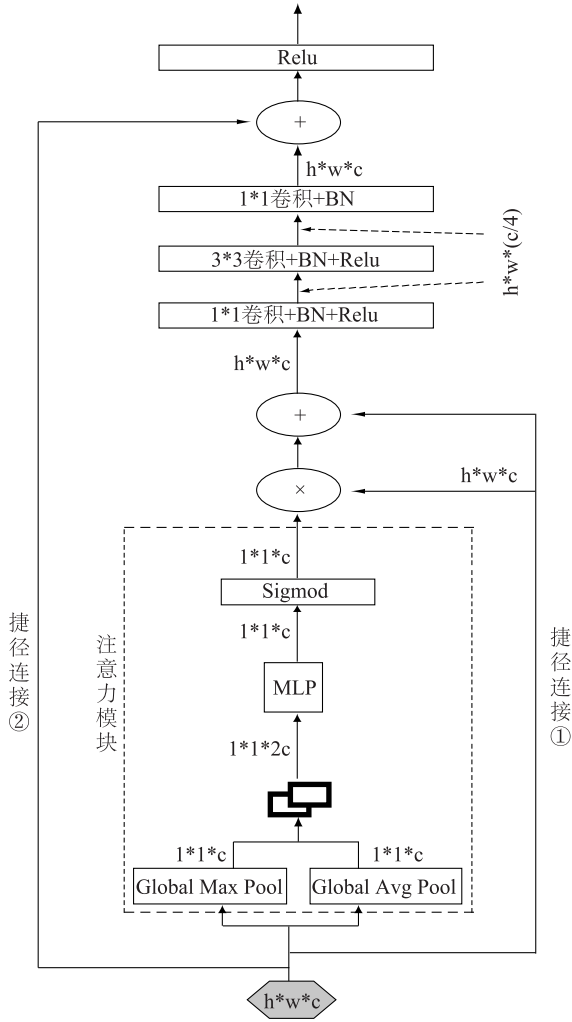


图2 并行池化的注意力卷积模块

Fig.2 Parallel pooled attention convolution module

块不会造成网络精度的降低。

1.3 模块整体构建

图3中下采样残差卷积模块(Downsampling residual convolution module, Down RCM, 未采用注意力机制)用来将浅层特征组合为高层特征并降低特征图大小。

采用 PyTorch 深度学习框架构建 PPA 和 RCM 模块并完成网络模型 PARNet 的搭建(图4),该模型由1个输入层、1个7×7大小卷积层、4个RCM层、13个PPA层、1个全连接层和1个输出层组合而成。

模型在读取图像数据后,会将图像大小调整为224×224像素,然后进行50%概率的翻转,最后经归一化后输入到卷积层进行特征提取。采用 Stochastic

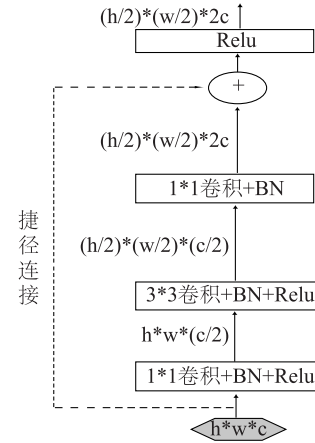


图3 下采样残差卷积模块

Fig.3 Downsampling residual convolution module

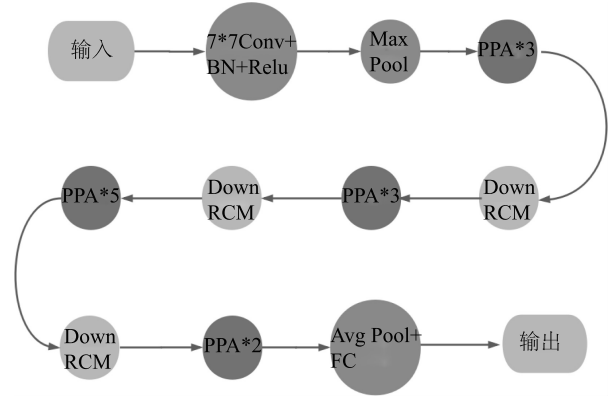


图4 本研究模型的网络结构

Fig.4 Network structure of the model constructed in this study

tic gradient descent (SGD) 训练网络参数并引入 Momentum 项抑制 SGD 的震荡。在普通的 SGD 中,参数的迭代更新公式为:

$$W = W - \alpha dW \quad (2)$$

$$b = b - \alpha db \quad (3)$$

而引入 Momentum 项后,参数的迭代更新公式为:

$$v_{dw} = \beta v_{dw} + (1 - \beta) dW \quad (4)$$

$$v_{db} = \beta v_{db} + (1 - \beta) db \quad (5)$$

$$W = W - \alpha v_{dw}, b = b - \alpha v_{db} \quad (6)$$

式中, α (学习率) 和 β 为自行设置的超参数, β 值决定上次的更新值对本次更新影响。引入 Momentum 项后,网络参数的每次更新都会考虑上次的更新值,加强梯度方向与上次梯度方向相同的参数更新,削减当前梯度方向与上一次梯度方向不同的参数更新。进而增加稳定性和提升学习速度,并且

能在一定程度上摆脱局部最优解^[13]。

1.4 迁移学习

迁移学习可以从以前的任务之中学习知识或经验并将其应用于新的任务之中^[14-16]。深度学习是一个数据驱动的学科^[17],普通的机器学习工作者通常难以获得足够多的数据和强大的算力,很难从零开始训练出足够优秀的 CNN 网络,但已经完成训练的模型通常能够迁移到其他不同的数据集上使用^[18]。因为尽管各个数据集里的图像不尽相同,但其基本特征大多是常见的,因此可以使用近似的提取方式。

因此,残差结构的参数使用 ResNet50 的预训练参数初始化,注意力模块参数采用随机初始化,对比试验的模型使用预训练参数初始化。以上预训练参数由深度学习框架 Pytorch 提供。

1.5 试验超参数设置

根据 GPU 的显存大小,在每次迭代中,分多个批次进行训练,每个批次使用 32 张图像(即 Batch size 为 32)。训练的 Epoch 设为 60(训练数据从头到尾过一遍为 1 个 Epoch)。初始学习率为 0.01,采用指数衰减进行调整,第 30 个 Epoch 时学习率衰减到 0.001,第 60 个 Epoch 时学习率衰减到 0.000 1。引入权重衰减,使得权重能够衰减到更小的值,减轻模型的过拟合问题,权重衰减的参数值为 1×10^{-4} 。Momentum 项的参数为 0.9。

2 结果与分析

2.1 试验过程

试验在具有 Tesla P4 GPU 的深度学习服务器上进行,其 Linux 系统的版本为 4.19.104,Python 的版本为 3.6.9,Pytorch 的版本为 1.5.0。完成数据划分和模型搭建后,通过使用同一套训练代码调用不同模型完成相应训练,以控制变量。模型 PARNet 使用 Pytorch 搭建实现,对比网络模型由深度学习框架 Pytorch 提供。使用训练集和验证集训练和验证网络模型,每个 Epoch 中,首先使用训练集训练模型的权值,然后使用验证集测试模型的识别率(此时不进行反向传播,即不改变模型的权值),最后保存模型的权值,并开始下一个 Epoch 训练,直至完成 60 个 Epoch 训练。完成训练后,根据训练过程中模型在验证集上的表现,选出最优权值(识别率最高的权值),并保存为模型文

件。测试时,加载保存的模型文件,经前向传播后,给出识别结果。

2.2 PARNet 模型的识别率

统计训练得到的各个模型在验证集上的识别率(表 1),计算方法为:识别率=预测正确的样本数/总样本数。由表 1 可知,相比于其他模型,PARNet 模型的识别率提高了 2.25%~11.58%,达到了 96.91%,这主要得益于 PPA 模块。PARNet 模型中 PPA 模块能在训练中自学习特征图的注意力信息,抑制与病害无关的特征图的表达,加强与病害相关的特征图的表达,而 RCM 模块负责将低层特征图组合为高层特征图,通过两者的组合提升病害相关信息在高层的表达,抑制背景等信息的表达,进而提升模型识别效果。而将注意力模块置于卷积之前,为先通过注意力机制对特征进行约束增强再提取特征,更符合视觉注意力机制,还能有效控制参数量的增加。

使用 Python 中的 THOP(PyTorch-OpCounter)库计算参数量。除 VGG16 的参数量为 138.36 M 外,其余模型的参数量均在 25 M 至 29 M 之间,这主要得益于对 1×1 卷积的使用。而 PARNet 的注意力模块通过并行的池化操作压缩输入特征图,减少了注意力模块的参数量,以实现在提高识别率的同时,尽可能减少增加的参数量。最终其参数量比 ResNet50 高出 1.26 M,为 26.82 M,低于其余模型。

表 1 不同模型的识别率与参数量

Table 1 Recognition rate and parameter amount of different models

模型	识别率(%)	参数量(M)
VGG16 ^[6]	85.33	138.36
Inception V3 ^[19]	91.75	27.16
ResNet50 ^[8]	93.25	25.56
SeNet ^[11]	94.66	28.07
PARNet	96.91	26.82

2.3 PARNet 模型的预测结果

为进一步观察各模型在测试集(预测集)上的表现,通过混淆矩阵(表 2)展示识别效果前三的 ResNet50、SeNet 和 PARNet 模型在各类病害识别上的表现。混淆矩阵是机器学习中总结分类模型预测结果的情形分析表,以矩阵形式将真实的类别(每一行为一类真实类别)与分类模型预测的类别(每一列为一类预测类别)进行汇总。

表 2 ResNet50、SeNet 和 PARNet 模型预测结果的混淆矩阵

Table 2 Confusion matrix of prediction results of ResNet50, SeNet and PARNet models

实际类别	ResNet50 预测类别						SeNet 预测类别						PARNet 预测类别					
	早疫病	晚疫病	叶霉病	斑枯病	花叶病毒病	健康叶	早疫病	晚疫病	叶霉病	斑枯病	花叶病毒病	健康叶	早疫病	晚疫病	叶霉病	斑枯病	花叶病毒病	健康叶
早疫病	89	7	1	2	1	0	91	4	1	3	1	0	94	3	1	2	0	0
晚疫病	6	91	2	1	0	0	6	92	1	1	0	0	4	95	1	0	0	0
叶霉病	2	3	95	0	0	0	2	1	96	0	1	0	0	1	99	0	0	0
斑枯病	3	0	3	92	1	1	3	2	1	92	1	1	3	0	1	94	1	1
花叶病毒病	0	1	0	2	96	1	0	1	0	1	98	0	0	0	0	1	99	0
健康叶	0	0	0	0	0	100	0	0	0	0	0	100	0	0	0	0	0	100

由表 2 中的混淆矩阵可以得知 ResNet50 模型对番茄花叶病毒病的识别效果最好,对早疫病的识别效果最差,两者的识别效果相差 7 个百分点左右,每类病害的预测结果中包含的错误类别较多,共 16 类识别错误。相比于 ResNet50 模型,SeNet 模型识别效果得到一定提升,且每类病害预测出的错误类别有所减少。而 PARNet 模型的识别效果最好,各类病害的识别效果均优于 Resnet50 模型和 SeNet 模型,且每类病害预测出的错误类别进一步减少,只有 11 类识别错误。

以混淆矩阵中的数据为基础,计算精确率(表 3)以进一步分析识别效果。精确率针对的是预测结果,表示预测结果中被正确预测的比率。精确率计算公式为:精确率=(A 类预测为 A 类)/(A 类预测为 A 类+非 A 类预测为 A 类)。PARNet 模型在 5 类病害上的精确率比 ResNet50 模型和 SeNet 模型高出 0.09~6.74 个百分点,这得益于依据注意力思想和残差思想构建的 PPA 模块。网络模型对病害的识别,依赖于对病害特征的提取,病害特征对分类决策的贡献越大,越能取得更好的识别效果。而借助注意力思想,PPA 模块可通过对特征图的约束增强,给予病害区域更多的注意力。即减少对非重要区域的学习,增强对病害区域的学习,进而提高病害特征对分类决策的贡献,提升识别效果。

网络提取出的一些病害特征是不同病害共有的,病害之间相似特征越多,越容易出现识别错误。更重要的是,病害是一个动态的发展过程,图像中病害的严重度不一致,若图像中病害轻微,则不利于提取病害特征,容易导致网络给出错误的识别结果。

表 3 ResNet50、SeNet 和 PARNet 模型的精确率

Table 3 Accuracy of ResNet50, SeNet and PARNet models

病害类别	预测精确率(%)		
	ResNet50 模型	SeNet 模型	PARNet 模型
早疫病	89.00	89.22	93.07
晚疫病	89.22	92.00	95.96
叶霉病	94.06	96.97	97.06
斑枯病	94.85	94.85	96.91
花叶病毒病	97.96	97.03	99.00
健康叶	98.04	99.01	99.01

PARNet 模型对早疫病的识别精确率最低,为 93.07%;对晚疫病的识别精确率为 95.96%,对叶霉病和斑枯病的识别精确率在 97.00%左右;对花叶病毒病的识别精确率最高,为 99.00%。PARNet 模型预测结果的精确率平均为 96.84%,精确率的最大差值为 5.94%,故 PARNet 模型可为番茄叶部病害的识别提供参考。

2.4 PARNet 模型的稳定性

图 5 和图 6 展示了 ResNet50、SeNet 和 PARNet 模型在 60 轮迭代中验证集上的损失函数值和准确率的变化。可以看出在训练过程中损失值和准确率虽然存在一定波动,但总体上趋于稳定。PARNet 模型的损失值和准确率在训练过程中无明显波动,SeNet 模型的损失值和准确率存在少数大幅度波动,ResNet50 模型的损失值和准确率存在较多大幅波动。在第 7 轮迭代之后,PARNet 模型的损失值小于 ResNet50 和 SeNet 模型,准确率高出 ResNet50 和 SeNet 模型。尽管三者训练过程中的变化均相对稳定,但 PPA 模块使 PARNet 模型在训练过程中尽

可能地减少了对无关的特征学习,提升了病害特征对预测的贡献度,避免了学习无用特征导致的损失值和准确率的波动,进而取得了更高的稳定性和更好的识别效果。综上所述,PARNet 模型可以为番茄早疫病、晚疫病、叶霉病、斑枯病和花叶病毒病的识别提供参考。

2.5 番茄叶部病害可视化识别系统设计

将模型代码中与预测无关的部分(如反向传播、损失计算等)去除,仅保留必要的前向传播部分。经测试在主频为 2.20 GHz 的单核 CPU(支持超线程)上,如果每次识别 1 张图像,连续识别 100 图像需 27.6 s,平均每张图像 0.276 s。而实际应用中 CPU 通常在 4 核 8 线程以上,主频也高于 2.2 GHz,可通过并发或并行计算等技术降低识别所需时间。如果使用 Tesla P4(Compute capability: 6.1)进行识别,同样每次仅识别 1 张图像,连续识别 100 图像仅需 7.06 s,平均识别每张图像仅需 0.070 6 s。但是在 CPU 计算模式下加载模型需 3.23 s,而在 GPU 计算模式下需 12.87 s。因此模型的参数加载和调用接口必须位于全局部分,即在 WEB 应用启动后便完成加载,进行识别时可直接计算,不需要先加载参数再计算。

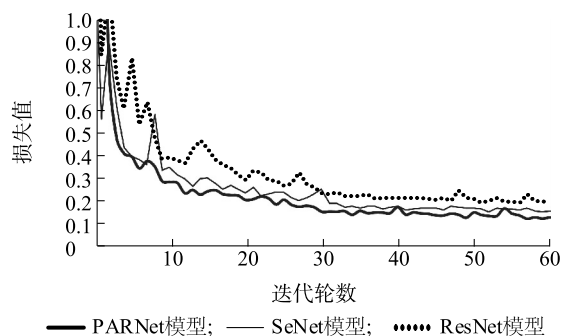


图 5 ResNet50、SeNet 和 PARNet 模型的损失值变化趋势

Fig.5 Change trend of loss value of ResNet50, SeNet and PARNet models

Flask 是 Python 环境下的轻量级 Web 应用框架,本研究采用 Flask 开发 WEB 应用,通过网页将番茄叶部图像上传到识别服务器,服务器完成识别后,通过网页将识别结果反馈给用户,因而将识别系统分为用户端(网页界面)与服务器端。借助于网页的便利性,实现在 Linux、Windows、Android 等系统下的番茄叶部病害识别。其交互过程如图 7 所示,图形界面的效果如图 8 至图 11 所示。

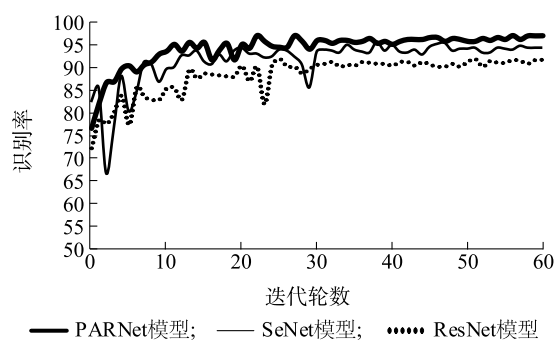


图 6 ResNet50、SeNet 和 PARNet 模型的准确率变化趋势

Fig.6 Change trend of accuracy rate of ResNet50, SeNet and PARNet models

注册时,用户密码经 MD5 算法计算后和用户名等信息储存于数据库。登录时,将输入密码经同样的 MD5 算法计算后,与数据库中对应密文比对,实现密码的验证。

完成登录后方可进入图像上传界面(图 10)。选择图像后,服务端先对图像进行预处理,然后通过模型得到识别结果,根据识别结果给出病害类别及各类病害的概率(图 11),点击参考意见,会根据识别结果,跳转到相应网站。

3 结论

本研究基于注意力机制构建了 PARNet 模型对番茄叶部早疫病、晚疫病、叶霉病、斑枯病、花叶病毒病进行识别,并与 VGG16、InceptionV3、ResNet50、SeNet 模型进行了初步对比,然后与其中表现较好的 ResNet50、SeNet 模型进行了深入对比,得出以下结论:1) PARNet 模型的并行注意力机制能有效区分特征图中的背景信息和目标信息,将更多的网络资源投入到与目标相关特征图中,更有效地提取病害相关特征,进而具备较高的分类性能,其平均识别率达到 96.91%,与其他模型相比高出 2.25~11.58 个百分点。2) PARNet 模型注意力机制于卷积之前实现即先划分输入信息的重要度再进行卷积,能更有效地利用卷积提取目标信息,而且不会干扰捷径连接通路反向传播,可以有效保证前层网络的训练效果。3) PARNet 模型的预测精确率平均为 96.84%,精确率的最大差值为 5.94%,能为番茄病害识别提供有效建议。4) 训练过程中未发生过度抖动,具备较高的稳定性。

通过开发 WEB 界面,实现通过网页完成识别

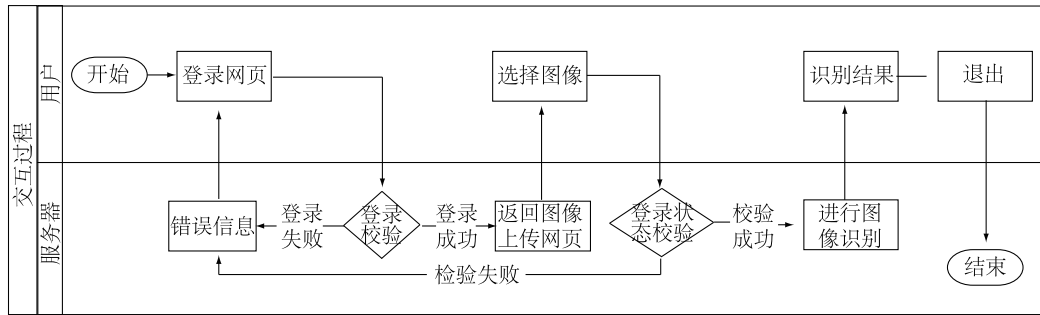


图7 界面交互的基本流程图

Fig.7 Basic flow chart of interface interaction



图8 系统登录界面

Fig.8 System login interface



图10 系统上传界面

Fig.10 System upload interface



图9 系统注册界面

Fig.9 System registration interface



图11 识别结果

Fig.11 Identification results

并获取识别结果,在使用 GPU (Compute capability: 6.1)的条件下,平均识别每张图像仅需0.070 6 s,能快速给出预测结果。但是模型对病害的识别较为单一,不能在同一叶片上对多种病害进行识别,还依赖单一背景。不同类别病害图像间病害严重度的分布不一致,对数据的均衡存在影响,影响网络的训练。故番茄病害图像数据需进一步扩充,在保证数据量的同时,保证数据的均衡,并进一步实现多病害、多环境下的识别,提升实用性。

参考文献:

- [1] 刘子记,杜公福,牛 玉,等.番茄主要病害的发生与防治技术[J].长江蔬菜,2019(19):59-62.
- [2] KEERTHI S S, SHECADE S K. SMO algorithm for least-squares SVM formulations[J]. Neural Computation, 2014, 15(2):487-507.
- [3] 许伟栋,赵忠盖. 基于卷积神经网络和支持向量机算法的马铃薯表面缺陷检测[J].江苏农业学报,2018,34(6):1378-1385.
- [4] 徐 岩,米 强,刘 斌,等. 基于字典学习的煤与矸石图像特

- 征识别方法[J].山东科技大学学报(自然科学版),2018,37(3):66-72.
- [5] LECUN Y, BENGIO Y, HINTON G. Deep learning [J]. Nature, 2015, 521(7553):436-444.
- [6] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[J]. Computer Science, 2014, 14(9):1409-1556.
- [7] SZEGEDY C, LIU W, JIA Y, et al. Going deeper with convolutions[C]// IEEE. 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Boston: IEEE, 2015:1-9.
- [8] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]// IEEE. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Boston: IEEE, 2016:770-778.
- [9] LETARTE G, PARADIS F, GIGUERE P, et al. Importance of self-attention for sentiment analysis[C]// ACL. Proceedings of the 2018 EMNLP Workshop BlackboxNLP: Analyzing and Interpreting Neural Networks for NLP. Brussels: Association for Computational Linguistics. Brussels: ACL, 2018: 267-275.
- [10] 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.
- [11] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]// IEEE. The IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Salt Lake City: IEEE, 2018:7132-7141.
- [12] 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.
- [13] RUDER S. An overview of gradient descent optimization algorithms [DB/OL]. (2016-09-15) [2017-6-15]. <https://arxiv.org/abs/1609.04747>.
- [14] 刘嘉政. 基于深度迁移学习模型的花卉种类识别[J].江苏农业科学, 2019, 47(20):231-236.
- [15] 庄福振. 迁移学习研究进展[J]. 软件学报, 2015, 26(1):26-39.
- [16] LONG J, SHEKHANER E, DARRELL T. Fully Convolutional networks for semantic segmentation[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2014, 39(4):640-651.
- [17] 王秀美, 牟少敏, 邹宗峰, 等. 基于深度学习的小麦蚜虫预测预警[J].江苏农业科学, 2018, 46(5):183-187.
- [18] HONG-WEI N, DUNG N, VASSILIOS V, et al. Deep learning for emotion recognition on small datasets using transfer learning[C]// ACM. Proceedings of the 2015 ACM on International Conference on Multimodal Interaction. Seattle: ACM, 2015:443-449.
- [19] SZEGEDY C, VANHOUCHE V, IOFFE S, et al. Rethinking the inception architecture for computer vision[C]// IEEE. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Boston: IEEE, 2016:2818-2826.

(责任编辑:张震林)