

# 多卷积神经网络模型融合的农作物病害图像识别

龚 安, 井晓萌

(中国石油大学(华东) 计算机科学与技术学院, 山东 青岛 266580)

**摘 要:**农作物病害是粮食安全的主要威胁,病害的诊断对于农业生产来说至关重要。针对单一卷积神经网络在农作物病害识别上的局限性,分类准确率不高的问题,采用多个卷积神经网络模型融合的方式,对10种农作物的27种病害及其3种病害程度的农作物叶子图片进行病害及病害程度的细粒度识别。首先选用Resnet101、RestNext50、SE-ResNet50、SE-ResNext50这4种网络模型运用迁移学习的方式,固定底层模型参数,修改顶层的全连接层进行训练,然后采用Stacking方法将模型预测结果输入第二层元学习器XGBoost,最后对单模型预测结果和Stacking融合后的结果进行对比。实验结果表明,经过模型融合后的准确率能达到87.19%,具有较高的识别准确率及较强的鲁棒性,可以作为农作物病害的早期诊断方式,并可以进一步研究将该方法应用到真实的农业生产中。

**关键词:**农作物病害识别;模型融合;卷积神经网络;元学习器;迁移学习

中图分类号:TP391

文献标识码:A

文章编号:1673-629X(2020)08-0134-06

doi:10.3969/j.issn.1673-629X.2020.08.023

## Image Recognition of Crop Diseases Based on Multi-convolution Neural Network Model Ensemble

GONG An, JING Xiao-meng

(School of Computer and Communication Technology, China University of Petroleum (East China),  
Qingdao 266580, China)

**Abstract:** Crop disease is a major threat to food security and its diagnosis is of significant importance for agricultural production. For the limitation of a single convolutional neural network in the identification of crop diseases and the problem of low classification accuracy, we adopt the ensemble method of multiple convolution neural network models to identify 27 diseases of 10 crops and their 3 degrees of disease by fine particle size. Firstly, the four network models Resnet101, RestNext50, SE-ResNet50 and SE-ResNext50 are used to transfer learning, fix the underlying model parameters, and modify the top-level fully connected layer for training. Then, the model prediction result is input into the second layer element learner XGBoost by Stacking, and finally the single model prediction result and the stacking result are compared. The experiments show that the accuracy of model fusion can reach 87.19%, which has high recognition accuracy and strong robustness. It can be used as an early diagnosis method of crop diseases, and can be further studied and applied to real agricultural production.

**Key words:** crop disease identification; model ensemble; convolutional neural network; meta learner; transfer learning

## 0 引言

中国作为农业大国,农业生产在中国经济发展中发挥着重要的作用,农作物是保障人民生活顺利运转的重要资源,然而农作物病害却会严重影响农作物的产量。农作物病害的识别通常通过直接观察植物叶片进行症状判断,但是这种方法对于没有专业知识的普通农民来说具有一定的困难,因而会导致病害误诊或者错过最佳治理时间。因此,运用深度学习<sup>[1]</sup>的方法

帮助农民识别诊断病害图像,尽早检测农作物的健康状况并及时发现病害症状至关重要。

近年来,卷积神经网络(convolutional neural network, CNN)在图像分类和识别中取得了显著的研究成果<sup>[2-3]</sup>。因此,国内外专家学者也将卷积神经网络应用于农作物的病害诊断中,并取得了不错的研究成果。在国外,将图像处理技术应用在农作物病害识别方面的研究相对较早,且有着完善的图像数据库。

收稿日期:2019-09-05

修回日期:2020-01-10

基金项目:国家科技重大专项(2017ZX05013-001)

作者简介:龚 安(1971-),男,副教授,硕导,CCF会员(62929M),研究方向为图像识别、大数据智能处理;井晓萌(1994-),女,CCF会员(A2977G),硕士研究生,研究方向为图像识别、大数据智能处理。

Ramcharan 等人<sup>[4]</sup>运用迁移学习训练深度卷积神经网络,对木薯的三种疾病和两种病虫害进行识别,利用 Inception v3 模型训练,通过多次对单个模型进行调参取得了不错的结果,并将其利用的卷积神经网络模型在移动端使用,这对帮助农民对农作物病害的进行识别有着极大的现实意义。Too 等<sup>[5]</sup>对 14 种植物 38 种不同的类别进行植物疾病识别,其采用目前比较先进的深度卷积神经网络进行微调,通过不断加深网络模型层数来提高网络性能,网络模型包括 VGG16<sup>[6]</sup>、InceptionV4<sup>[7]</sup>、ResNet50<sup>[8]</sup>、ResNet101、ResNet152、DenseNet121<sup>[9]</sup>网络结构。Khamparia 等人<sup>[10]</sup>利用卷积神经网络和自动编码器组合设计了一种检测农作物叶片病害的混合方法,在小样本上对 3 种作物和 5 种作物病害进行训练,在训练时采用  $2 \times 2$  和  $3 \times 3$  不同的卷积滤波器,最后在小样本数据集上取得了非常不错的效果。相对传统的图像识别方法,深度学习方法准确率更高,检测速度快,国内的研究人员也将目光投向了利用深度学习来解决农作物病害识别的问题。杨晋丹等人<sup>[11]</sup>利用卷积神经网络对草莓叶部白粉病害进行识别,设计了一种混合池化的 CNN-9 模型,其首先设计由 3 种不同网络深度与 3 种不同尺寸的卷积核交叉组合的网络结构,然后选择 4 种采样层构建方法,经多次训练测试后,最终对草莓白粉病病害识别达到了 98.61% 的正确识别率。刘永波等人<sup>[12]</sup>将卷积神经网络应用到自然环境下玉米 10 种常见病害的识别上,应用 Triplet loss 双卷积神经网络学习图像特征,使用 sift 算法提取图像纹理细节,最后通过 SoftMax 对图像进

行分类,最后取得了 90% 的准确率。蒋丰千等人<sup>[13]</sup>采用卷积神经网络对生姜的炭疽病、姜瘟病、根结线虫病和白星病进行研究分析达到了 96% 的识别率,并设计了人机交互界面,通过可视化界面提高了系统使用的便捷性。

结合上述研究发现,目前国内外的农作物病害识别还存在一些问题:(1)通过不断增加单一网络模型的深度来提高准确率,往往会导致出现训练难以收敛的情况;(2)对单一模型不断地进行调参,从而增加了时间成本;(3)多数针对单一农作物的病害进行识别。针对这些问题,文中在现有国内外农作物病害识别研究的基础上,对原有卷积神经网络模型进行优化,采用 Stacking<sup>[14]</sup>融合算法对苹果、樱桃、玉米、柑橘、马铃薯等十种农作物的物种-病害-程度进行细粒度分类识别。采用多模型融合的方法,提高识别的准确率,优化识别效率,从而得到高质量的检测效果,弥补农民专业知识的不足。

## 1 农作物病害识别网络模型

文中所使用的基础网络模型都是深度卷积神经网络,卷积神经网络是一种多层神经网络,包括输入层、卷积层、池化层、全连接层、输出层。每层是由多个二维平面(即特征映射)组成,而每个平面由多个独立的神经元组成,通过不同的卷积核提取图像的局部特征,将该局部特征组合成更高层的全局特征,从而完成图像从局部到整体的映射过程。图 1 为常见卷积网络结构。

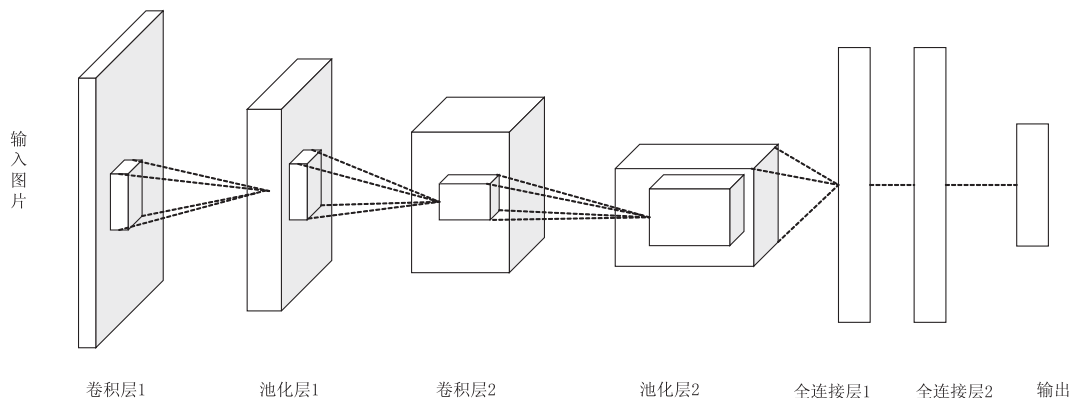


图1 卷积网络结构

### 1.1 网络特征提取

卷积神经网络通常是采用卷积层对输入的数据进行特征提取,卷积层内部包含很多个卷积核,组成卷积核的每个元素都对应一个权重系数和一个偏差量(bias vector),类似于一个前馈神经网络的神经元(neuron)。卷积层内每个神经元都与前一层中位置接近的区域的多个神经元相连,区域的大小取决于卷积核的大小。卷积核在工作时,会有规律地扫过输入特

征,在感受野内对输入特征做矩阵元素乘法求和并叠加偏差量。只有对输入图像进行准确、全面的特征提取才能确保分类识别的正确性。但是随着 CNN 网络层数的不断增加,训练难度逐渐增大,为了降低训练成本,文中使用 ImageNet 预训练好的模型,固定底层模型参数,修改顶层的全连接层,分别使用 4 个模型进行特征提取。

为提取能更好地表达语义信息的特征,文中在深

度卷积网络的基础上,将残差学习的思想应用到农作物病害特征的提取上。

图2是残差学习模块的示意图,残差学习模块可以作为神经网络的一部分也可作为多个部分。假设该部分卷积神经网络的输入为 $x$ ,要拟合的函数映射(即

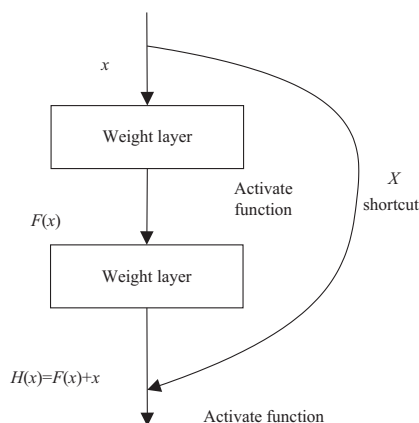


图2 残差学习示意图

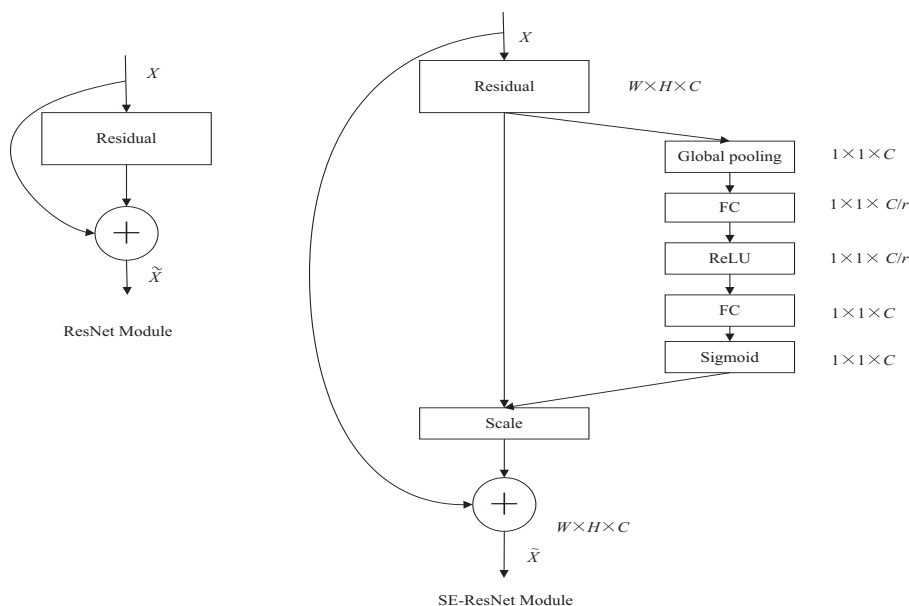


图3 SE-ResNet 结构

## 1.2 损失函数

常见的损失函数有很多,在处理多分类任务时通常采用交叉熵损失函数,其表达式为:

$$\text{loss} = - \sum_{i=1}^n y_i \log(p_{ii})$$

其中, $y_i$ 表示期望输出, $p_{ii}$ 表示模型对样本的实际输出。

然而在文中使用的数据集各类别间图片数量存在不均衡的问题,类别的不均衡问题会影响各类别结果的输出,故使用改进的交叉熵损失函数—focal loss<sup>[16]</sup>。为了解决类别不均衡问题,何凯明团队在2016年提出focal loss损失函数,给交叉熵损失函数添加了一个预测概率 $P_i$ 和超参数 $\gamma$ ,公式可以表达为:

输出)为 $H(x)$ ,可以定义另外一个残差映射为 $F(x) = H(x) - x$ ,则原始的函数映射 $H(x)$ 可以表示为 $F(x) + x$ 。He的实验表明:优化残差映射 $F(x)$ 比优化原始映射 $H(x)$ 要比较容易。

$F(x) + x$ 在前馈神经网络中可以理解为捷径(short cut) $x$ 与主径 $F(x)$ 的加和。然而捷径并没有因此而引入额外的参数,并不影响原始神经网络的复杂度,整体神经网络依然可使用现有的深度学习进行反馈训练求解。

为提高网络的表示能力,文中引入SE<sup>[15]</sup>(squeeze-and-excitation)架构单元嵌入到现有的卷积神经网络结构中。SE的思想是在2017年由Jie Hu等人提出,网络结构根据损失去学习特征权重,自适应地校准通道的特征响应,通过给予有效的特征更大的权重而抑制无效或者效果小的特征,从而得到更好的效果。SE模块是一个子结构,图3是将SE模块嵌入到ResNet网络的结构图。

$$\text{FL}(P_i) = -(1 - P_i)^\gamma \log(P_i)$$

该损失函数通过减少易分样本的权重,使模型在训练时更专注难分类的样本,从而解决农作物病害分类的不均衡问题。

## 1.3 自适应学习率优化算法

传统的优化算法不是将学习率设置为常数就是根据训练次数调节学习率,因此极大忽视了学习率其他变化的可能性。然而学习率对网络模型的性能有着非常大的影响,因此需要采取一系列策略来更新学习率,从而提高网络训练速度。Adam<sup>[17]</sup>,AdaGrad,RMSProp等算法都是常见的自适应学习率优化算法。文中选用Adam算法,Adam是一种将Momentum算法和RMSProp算法结合起来使用的算法,通过

Momentum 利用类似于移动指数加权平均的方法来对网络的参数进行平滑处理,让梯度的摆动幅度变得更小,网络收敛速度更快,同时利用 RMSProp 算法对梯度计算微分平方加权平均数,消除摆动幅度大的方向,用来修正摆动幅度,使得各个维度的摆动都较小。

## 2 基于 Stacking 的模型融合策略

### 2.1 XGBoost 元学习器

XGBoost<sup>[18]</sup> (extreme gradient boosting) 是由华盛顿大学的陈天奇博士提出的分布式梯度提升树算法,由于其较高的准确率和出色的性能,越来越多的人将其应用在工业界和 Kaggle 等数据竞赛中。XGBoost 算法是在传统的 GBDT<sup>[19]</sup> 算法基础上进行改进的,GBDT 算法采用一阶导数信息,而 XGBoost 算法同时用到了一阶导数和二阶导数,对损失函数做二阶泰勒展开,有利于梯度下降的更快,XGBoost 算法还在目标函数之外加入了正则项,控制模型的复杂性,降低了出现过拟合的可能性。XGBoost 可以处理高维度特征,并有着处理速度快的优点,所以文中在 Stacking 算法的第二层选用 XGBoost 元学习器。该算法的思想就是通过不断进行特征分裂来生成树,每次添加树的过程就是学习一个新的函数来拟合上次预测的残差的过程。如果要预测一个种类的得分,就是根据这个种类的特征,在训练的每棵树中落到对应的一个叶子节点上,这样每个叶子节点就会对应一个得分,最后将每棵树对应的得分加起来就得到这个种类的预测值。

### 2.2 模型融合步骤

模型融合也被称为集成学习 (ensemble learning),主要思路是先通过一定的规则生成多个学习器,再采用集成策略进行模型组合,最后综合判断从而输出最终结果。融合多个模型可以弥补单个模型的不足,从而达到比训练单个模型更好的分类结果,并且比单模型具有更高的稳定性。常见的集成方法有投票法、平均法、Boosting、Bagging、Stacking,文中采用基于 Stacking 的集成学习方式,以下是 Stacking 算法的伪代码。

输入:训练数据集  $D = \{(x_1, y_1), (x_2, y_2), \dots, (x_m, y_m)\}$ ;  
次级分类器  $L_1, \dots, L_T$ ; 元学习器  $L$

过程 1:学习次级分类器

For  $t = 1, 2, \dots, T$ :

$h_t = L_t(D)$

End;

过程 2:构建新的数据集用于预测

For  $t = 1, 2, \dots, m$ :

$X'_i = \{h_1(x_i), \dots, h_T(x_i)\}$

$D_h = \{x'_i, y_i\}$

End;

过程 3:学习元分类器

学习  $H(D_h)$

输出:集成学习分类器  $H$

首先,为了防止发生过拟合,文中采用 5-折交叉验证法,将农作物病害原始训练数据集随机划分成 5 个大小基本相等的子集。

其次,采用 Resnet101、RestNext50、SE-ResNet50、SE-ResNext50 这 4 种网络模型作为 Stacking 算法的基础模型,对每一份训练数据,都用剩余 4 份数据来训练模型,预测出该份数据的结果,这样经过 5 次交叉检验,将得到 5 个预测结果,然后分别对原始测试集进行预测,将几个结果加权平均。

最后,在这四个基础模型完成交叉验证过程后,根据它们输出的样本预测结果来构建新的预测结果,作为第二层元学习器 XGBoost 输入,由 XGBoost 进行训练后得到最终预测结果。

### 2.3 算法流程

文中采用了 Resnet101、RestNext50<sup>[20]</sup>、SE-ResNet50、SE-ResNext50 这 4 种网络来提取图像特征,然后将其预测结果作为 Stacking 算法第二层元学习器 XgBoost 的输入,应用于农作物病害识别领域。文中流程如图 4 所示,输入为农作物病害图像,这几种网络分别从输入的农作物病害图像中检测提取出病害特征,然后应用 SoftMax 分类器对图像进行分类,最后将预测结果作为 XgBoost 的输入,得到最终的分类结果。

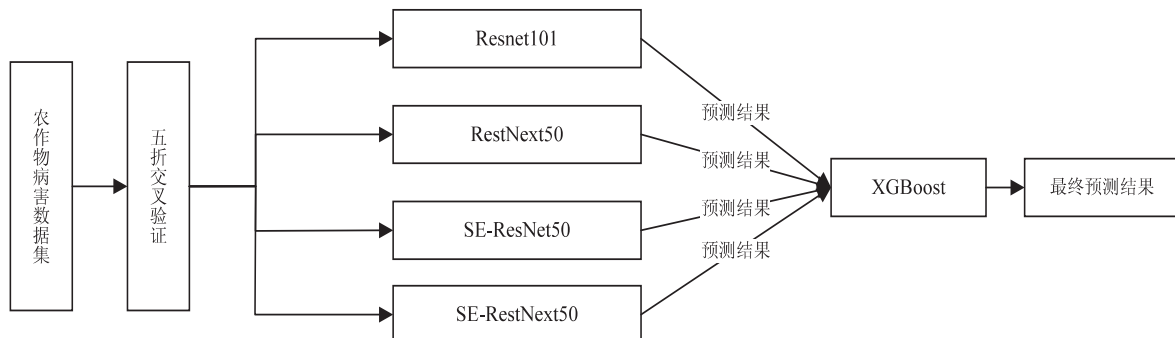


图4 算法流程



### 3 实验及结果分析

实验环境如表 1 所示。

表 1 实验环境

名称	版本
CPU 处理器	Intel(R) Core(TM) i7-6700K
GPU 处理器	NVIDIA GeForce GTX TITAN X
内存 (RAM)	16.00 GB
操作系统	64bit Windows 10 专业版
程序设计语言	Python 3.6.5
框架	Pytorch 0.4

#### 3.1 数据集

文中所用数据集来自 AI Challenger (<https://challenger.ai/competition/pdr2018>), 共有 47 393 张图片, 数据集分为训练集 70%, 验证集 10%, 测试集 20%, 为方便结果对比, 使用原验证集作为测试集。该数据集按照物种-病害-程度分为 61 种不同的类别, 10 个物种 (包括苹果、樱桃、玉米、葡萄、柑桔、桃、辣椒、马铃薯、草莓、番茄叶子图片), 27 种病害 (包括苹果黑星病、苹果灰斑病、苹果雪松锈病、樱桃白粉病、玉米灰斑病、玉米锈病、玉米叶斑病、玉米花叶病毒病、葡萄黑腐病、葡萄轮斑病、葡萄褐斑病、柑桔黄龙病、桃疮痂病、辣椒疮痂病、马铃薯早疫病、马铃薯晚疫病、草莓叶枯病、番茄白粉病、番茄疮痂病、番茄早疫病、番茄晚疫病、番茄叶霉病、番茄斑点病、番茄斑枯病、番茄红蜘蛛损伤、番茄黄化曲叶病毒病、番茄花叶病毒病, 其中 24 个病害分一般和严重两种程度), 3 种病害程度 (分为健康、一般、严重)。其中番茄疮痂病一般 (第 44 类) 和番茄疮痂病严重 (第 45 类) 训练集中分别只有 1 张图片, 验证集分别有 1、0 张图片, 所以为了提高准确率, 在图像识别的过程中将 44 类、45 类删除。

#### 3.2 数据预处理

深度学习是一种基于大数据的方法, 数据的规模越大、质量越高就越好, 模型也能够拥有更好的泛化能力, 然而在实际采集数据的时候, 往往很难覆盖全部的场景, 数据增强是扩充数据样本规模的一种有效的方法, 为了避免网络过拟合现象, 提升整体性能, 文中对训练集中所有图片进行了随机旋转、上下翻转、随机裁剪的数据增强操作, 测试集不做任何处理。

#### 3.3 评价指标

所用评价指标公式如下:

$$ACC = \frac{N(c)}{N(t)} * 100\%$$

其中,  $N(c)$  表示判断正确的个数,  $N(t)$  表示总个数。

#### 3.4 实验结果和分析

先使用前面所提到的 4 种深度卷积神经网络对农

作物叶子病害及病害程度图像进行识别, 表 2 是实验所用参数。首先对训练集进行 5-折交叉验证, 将数据随机划分为 5 份, 其中 4 份作为训练集, 1 份作为测试集, 重复五次实验, 再用测试结果对原测试集进行预测。在训练时, 为了节省时间和提高分类性能, 实验运用迁移学习, 将低层次模型参数固定, 在训练时不改变这些层的参数, 修改顶层的全连接层的参数为 59 类进行训练。表 3 展示的就是在测试集上不同网络模型的农作物病害种类及其程度的识别结果。为方便结果记录, 将 Resnet101 记为 M1, RestNext50 记为 M2, SE-ResNet50 记为 M3, SE-ResNext50 记为 M4。

表 2 实验所用参数

参数	大小
初始 Imagesize	224×224×3
Batch size	32
Epoch	40
Momentum	0.9
Learning rate	0.01 ~ 0.000 1

表 3 不同网络模型病害种类及程度实验结果

模型名称	准确率%
M1	84.60
M2	85.71
M3	84.61
M4	86.83

从表 3 中可以看出, SE-ResNext50 在单独使用时比其他网络提供了更好的性能, 带来了更高的准确率。那么在对模型进行融合时, 可以优先考虑使用 SE-ResNext50 网络与其他网络组合。实验的最后, 采用 Stacking 方法将模型预测结果输入第二层元学习器 XgBoost, 从而得到最终的预测结果。对比单个卷积神经网络模型, 为了实现更高的识别准确率, 文中将网络采用排列组合的方式对单个网络模型进行融合。模型融合实验结果如表 4 所示。

表 4 模型融合实验结果

模型组合	准确率%
M1+M4	86.92
M2+M4	87.09
M3+M4	87.14
M1+M2+M3+M4	87.19

表 4 说明基于模型融合的方法相比单模型有一定的优势, 弥补了单网络模型的不足, 提高了识别准确率。尤其是 M1+M2+M3+M4 组合, 即将 Resnet101, RestNext50, SE-ResNet50, SE-ResNext50 这四种网络组合比性能最好的单模型在准确率上提高了 0.26%,

并且从一定程度上来说,模型融合相比单网络模型结构来说具有较好的稳定性和泛化能力。

#### 4 结束语

针对现有农作物病害识别时单一模型准确率不高、且多是对单一物种的不同病害的识别,提出了一种将多种主流神经网络融合的方法,还借鉴使用了 focal loss 损失函数解决了数据分类不均衡的问题。通过融合后的模型对农作物病害及程度识别最高准确率达到 87.19%,基于这种高水平的性能,很明显卷积神经网络通过分析简单的叶子图像非常适合于农作物病害的自动检测和诊断。但是文中所用数据多为简单背景并且数据量小,所以扩大现有数据库以纳入更多种类的植物物种和疾病应成为下一个紧迫的任务;另一个关键问题是未来在该领域验证该方法的工作应旨在开发基于移动设备的应用程序,使农民能够快速监测疾病情况。

#### 参考文献:

- [1] DENG L, YU D. Deep learning: methods and applications [J]. Foundations & Trends in Signal Processing, 2014, 7 (3): 197–387.
- [2] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks [J]. Communications of the ACM, 2017, 60(6): 84–90.
- [3] ZEILER M D, FERGUS R. Visualizing and understanding convolutional networks [C]//Proceedings of the 13th European conference on computer vision. Zurich: ECCV, 2014: 818–833.
- [4] RAMCHARAN A, BARANOWSKI K, MCCLOSKEY P, et al. Using deep learning for image-based cassava disease detection [J]. Frontiers in Plant Science, 2017, 8(1): 1852–1859.
- [5] TOO E C, YUJIAN L, NJUKI S, et al. A comparative study of fine-tuning deep learning models for plant disease identification [J]. Computers and Electronics in Agriculture, 2018, 161: 272–279.
- [6] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition [C]//Proc of CVPR. [s. l.]: [s. n.], 2014: 1049–1062.
- [7] SZEGEDY C, IOFFE S, VANHOUCKE V, et al. Inception-v4, Inception-ResNet and the impact of residual connections on learning [C]//International conference on learning representations. [s. l.]: [s. n.], 2016: 536–548.
- [8] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition [C]//2016 IEEE conference on computer vision and pattern recognition (CVPR). Las Vegas: IEEE, 2016: 770–778.
- [9] HUANG G, LIU Z, MAATEN L V D, et al. Densely connected convolutional networks [C]//2017 IEEE conference on computer vision and pattern recognition (CVPR). Las Vegas: IEEE, 2017: 4700–4708.
- [10] KHAMPARIA A, SAINI G, GUPTA D, et al. Seasonal crops disease prediction and classification using deep convolutional encoder network [J]. Circuits, Systems, and Signal Processing, 2019, 34(1): 818–836.
- [11] 杨晋丹, 杨涛, 苗腾, 等. 基于卷积神经网络的草莓叶部白粉病病害识别 [J]. 江苏农业学报, 2018, 34(3): 527–532.
- [12] 刘永波, 雷波, 曹艳, 等. 基于深度卷积神经网络的玉米病害识别 [J]. 中国农学通报, 2018, 34(36): 159–164.
- [13] 蒋丰千, 李畅, 余大为, 等. 基于 Caffe 的生姜病害识别系统研究与设计 [J]. 中国农机化学报, 2019, 40(1): 126–131.
- [14] JUREK A, BI Yaxin, WU Shengli, et al. Clustering-based ensembles as an alternative to stacking [J]. IEEE Transactions on Knowledge & Data Engineering, 2014, 26(9): 2120–2137.
- [15] HU J, SHEN L, SUN G. Squeeze-and-excitation networks [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 10(1): 7132–7141.
- [16] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection [C]//IEEE international conference on computer vision (ICCV). [s. l.]: [s. n.], 2017.
- [17] KINGMA D P, BA J. Adam: a method for stochastic optimization [C]//3rd international conference for learning representations. San Diego: [s. n.], 2015.
- [18] CHEN T, GUESTRIN C. XGBoost: a scalable tree boosting system [C]//Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining. Ithaca: ACM, 2016: 785–794.
- [19] FRIEDMAN J H. Greedy function approximation: a gradient boosting machine [J]. The Annals of Statistics, 2001, 29(5): 1189–1232.
- [20] XIE S, GIRSHICK R, DOLLAR P, et al. Aggregated residual transformations for deep neural networks [C]//2017 IEEE conference on computer vision and pattern recognition (CVPR). Las Vegas: IEEE, 2017: 1492–1500.