

基于快照集成卷积神经网络的苹果叶部病害程度识别

刘 斌^{1,2} 徐皓玮¹ 李承泽¹ 宋鸿利¹ 何东健^{3,4} 张海曦¹

(1. 西北农林科技大学信息工程学院, 陕西杨凌 712100; 2. 农业农村部农业物联网重点实验室, 陕西杨凌 712100;
3. 西北农林科技大学机械与电子工程学院, 陕西杨凌 712100;
4. 陕西省农业信息感知与智能服务重点实验室, 陕西杨凌 712100)

摘要: 针对苹果叶部病害程度识别准确率低的问题,构建了一种基于快照集成方法的苹果叶部病害程度识别模型。首先,通过多种数字图像处理技术对原始苹果叶部病害图像进行数据增强;然后,选取 Inception-ResNet V2 作为基模型,引入 CBAM 模块提升网络的特征提取能力,使用焦点损失函数缓解苹果叶部病害数据集类别不平衡问题;最后,通过快照集成方法进行模型集成,得到苹果叶部病害程度识别模型。利用苹果黑星病和锈病的早期和晚期病害数据集进行了模型验证,准确率高达 90.82%,比单一 Inception-ResNet V2 模型的准确率提高了 2.50 个百分点。实验结果表明,基于快照集成的识别模型准确率较高,为苹果叶部病害程度识别研究提供了参考。

关键词: 苹果叶部; 病害识别; 卷积神经网络; CBAM 模块; 焦点损失函数; 快照集成

中图分类号: TP391.41 **文献标识码:** A **文章编号:** 1000-1298(2022)06-0286-09 **OSID:** 

Apple Leaf Disease Identification Method Based on Snapshot Ensemble CNN

LIU Bin^{1,2} XU Haowei¹ LI Chengze¹ SONG Hongli¹ HE Dongjian^{3,4} ZHANG Haixi¹

(1. College of Information Engineering, Northwest A&F University, Yangling, Shaanxi 712100, China
2. Key Laboratory of Agricultural Internet of Things, Ministry of Agriculture and Rural Affairs, Yangling, Shaanxi 712100, China
3. College of Mechanical and Electronic Engineering, Northwest A&F University, Yangling, Shaanxi 712100, China
4. Shaanxi Key Laboratory of Agricultural Information Perception and Intelligent Service, Yangling, Shaanxi 712100, China)

Abstract: To address the problem of low recognition accuracy for identifying different apple leaf diseases, an apple leaf disease identification model was proposed based on snapshot ensemble. Firstly, the original dataset was augmented by various digital image processing methods. Then, an Inception-ResNet V2 was chosen as base model. The convolutional block attention module (CBAM) was introduced to enhance the feature extraction capability for apple leaf diseases. And focal loss was used to alleviate the imbalance of samples in each category. Finally, the model was integrated through snapshot ensemble to obtain the final identification model for different degrees of diseases on apple leaves. The image was input to the final model for identification. Compared with the original single Inception-ResNet V2, the recognition accuracy of the improved model was increased from 88.32% to 90.82%. Experimental results showed that the ensemble model had a high accuracy rate, which provided an idea and explored a approach for diseases of different degrees on apple leaves.

Key words: apple leaf; disease identification; CNN; CBAM module; focal loss; snapshot ensemble

收稿日期: 2021-07-12 修回日期: 2021-08-08
基金项目: 陕西省重点研发计划项目(2021NY-138)、CCF-百度松果基金项目(2021PP15002000)、国家重点研发计划项目(2020YFD1100601-02-13)、陕西省重点研发计划项目(2019ZDLNY07-06-01)、宁夏智慧农业产业技术协同创新中心项目(2017DC53)和国家级大学生创新创业训练计划项目(S202010712083)
作者简介: 刘斌(1981—),男,副教授,主要从事并行计算、深度学习与农作物病害诊断等研究,E-mail: liubin0929@nwsuaf.edu.cn
通信作者: 张海曦(1989—),男,讲师,主要从事深度学习及计算机视觉在农业领域应用研究,E-mail: zh.haixi@nwfufu.edu.cn

0 引言

苹果病害对其生长发育、产量、品质影响甚大，非常容易造成无法挽回的经济损失。因此，精确的早期苹果病害识别对降低苹果叶部的发病率和保证苹果产业的健康发展具有十分重要的意义。

现有的农作物病害识别方法主要包括 2 种：一类是利用数字图像处理技术通过背景分割、病斑分割、特征提取和分类等步骤实现对早期病害识别。基于传统数字图像处理技术的农作物病害诊断方法取得了较好识别效果^[1]，但在实际应用方面仍存在一定的局限性。大量的图像预处理工作费时费力，以及基于人工设计的特征忽略了病害图像中的全局信息和上下文信息，容易产生语义鸿沟。这些局限性直接导致该方法不能很好地满足早期农作物病害识别的精度要求。另一类是基于深度学习的农作物病害识别方法。随着深度学习技术在计算机视觉领域的快速发展，越来越多的研究者关注深度学习在农业领域的应用^[2]。卷积神经网络 (Convolutional neural network, CNN) 与传统方法相比，能够自动地提取鲁棒性强的叶部病害特征^[3]。许景辉等^[4]针对田地背景下小样本玉米叶部病害识别问题，构建了基于迁移学习的 CNN 模型，实现了对玉米锈病和大斑病染病叶片图像 95.33% 的平均识别准确率。任守纲等^[5]对于目前深度模型在识别病害的过程中易受环境因素干扰，在提取病害特征的过程中存在不确定性等问题，设计了一种基于反卷积引导的 VGG 网络模型，实验结果表明模型精度高达 99.19%。孙云云等^[6]采用 AlexNet 卷积神经网络

对茶轮斑病、炭疽病和云纹叶枯病的染病图像进行识别，实验结果显示模型的识别精度达到 94.29%。马浚诚等^[7]构建了一个基于图像处理的温室黄瓜霜霉病诊断系统，采用的病斑分割方法能够克服复杂背景和光照条件的影响，准确地提取病斑图像，黄瓜霜霉病识别准确率达到 90%。总体来说，目前对于病害程度的识别研究较少，识别准确率有待提高。

本文将深度学习应用于苹果叶部病害程度的识别，提出一种基于快照集成的早期苹果病害识别的卷积神经网络模型。以期提高苹果叶部病害程度识别的准确性，弥补传统方法和单一模型的不足，满足实际苹果生长过程中对病害早期诊断的实际要求。

1 材料与方法

1.1 病害图像采集与获取

本文实验采用的图像数据为 2018 年 AI Challenger 农作物病害检测比赛中的苹果病害数据集，并取样一部分西北农林科技大学白水苹果试验示范站苹果叶部病害图像作为补充。数据集最终包含简单背景条件下健康叶部、黑星病和锈病 3 大类样本共计 2 011 幅图像，其中病害的程度分为一般和严重 2 种状态。数据集中每幅图像都包含一片苹果叶子，即苹果叶子占据着每幅图像的主要位置。数据集中图像文件大小主要集中于 157 ~ 540 kB 之间，尺寸约为 256 像素 × 352 像素。数据集中包含健康叶片图像 671 幅、黑星病早期图像 426 幅、黑星病晚期图像 352 幅、锈病早期图像 275 幅和锈病晚期图像 287 幅(图像样本如图 1 所示)，可知该数据集存在数据分布不平衡的问题。

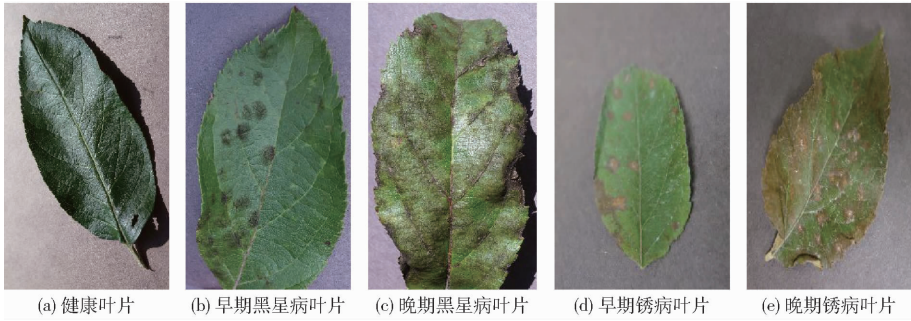


图 1 各类样本图像

Fig. 1 Various sample images

1.2 数据增强

提高神经网络模型泛化能力的最佳方法是采用更多的数据进行训练，但由于获取更多的具有可靠标注的苹果叶部病害图像耗时费力，原始数据集包含的样本图像是非常有限的。对于该问题主流的解决方法是对数据集进行数据增强，可以使有限数据产生相当于更多数据价值的同时，不

必大量增加原始数据量^[8]。本文主要使用 4 类方法进行数据增强：图像旋转与镜像、图像亮度与对比度调整、FANCY PCA^[9]和 MIXUP 方法^[10]。基于这 4 类增强方法生成的样本数据如图 2 所示。增强后的数据集按比例 3:1:1 划分为训练集、验证集和测试集，各样本集中的样本数量情况如表 1 所示。

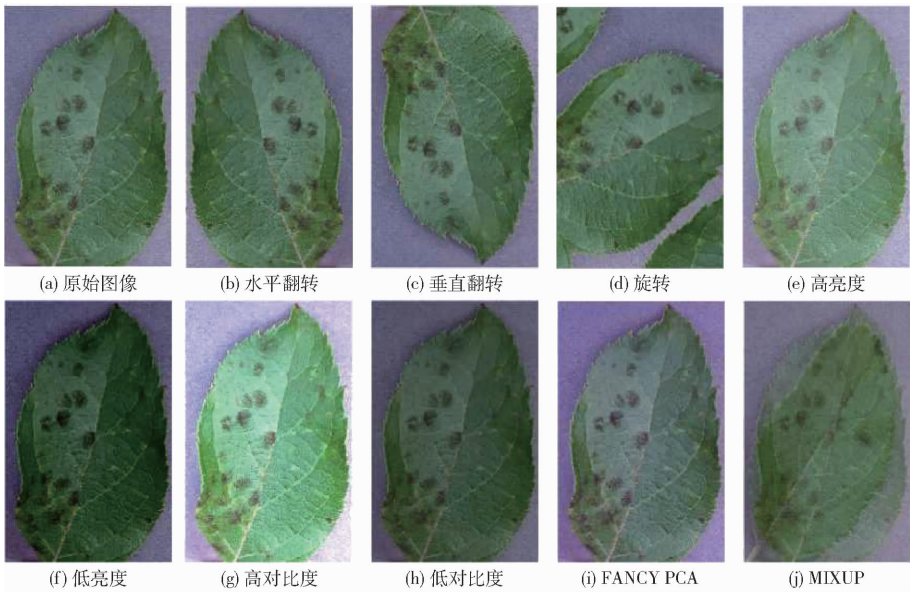


图 2 数据增强结果

Fig. 2 Results of data augmentation

表 1 苹果叶部病害数据集情况

Tab.1 Apple leaf disease data set

样本病害类别	总数 (原始)	总数 (增强)	训练集	验证集	测试集
健康叶片	671	6 710	4 026	1 342	1 342
黑星病早期叶片	426	4 260	2 556	852	852
黑星病晚期叶片	352	3 520	2 112	704	704
锈病早期叶片	275	2 750	1 650	550	550
锈病晚期叶片	287	2 870	1 722	574	574
总数	2 011	20 110	12 066	4 022	4 022

1.3 基于快照集成的苹果叶部病害程度识别模型

1.3.1 基于卷积块注意力模块的注意力机制

由于苹果叶片病害的病斑通常具有尺寸小、分布密集等特性,因此相比其他视觉分类任务,苹果叶片病害识别任务的难度更大。针对上述问题,本文通过加入注意力机制能更有效地对细小差异病斑图像进行特征提取与分类,以此改进模型对苹果叶部病害的识别精度。

卷积块注意力模块 (Convolutional block attention module,CBAM) 是一个轻量级的注意机制模块^[11],可以很容易嵌入到现有的主流卷积神经网络结构中,并且不会带来较多的额外运算量。CBAM 同时采用了最大池化和平均池化两种池化方式,并且通过特征通道和空间 2 个维度的信息来生成权重,实现对原特征的重标定操作。假设给定一个特征图 $X \in \mathbf{R}^{H \times W \times C}$ (其中 H 为图像高度, W 为图像宽度, C 为图像通道数),卷积块注意力模块依次沿着特征通道和特征空间 2 个独立的维度分别构建出通道注意力图 $M_c \in \mathbf{R}^{1 \times 1 \times C}$ 和空间注意力图 $M_s \in \mathbf{R}^{H \times W \times 1}$,整体结构如图 3 所示。

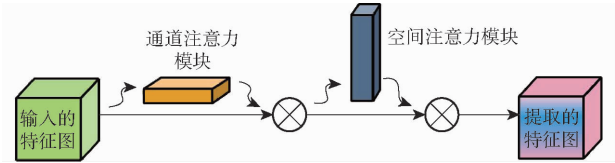


图 3 CBAM 模块整体结构图

Fig. 3 Overall structure diagram of CBAM module

其实现过程为

$$\begin{cases} X' = M_c(X) \otimes X \\ X'' = M_s(X') \otimes X' \end{cases} \quad (1)$$

式中 $\otimes$ ——矩阵点乘运算

X' ——通道注意力模块输出结果与原特征图进行点乘后的输出结果

X'' ——CBAM 的最终输出

CBAM 模块由通道和空间 2 个注意力模块构成,其中通道注意力模块具体实现过程如图 4a 所示。与 SE 模块^[12]类似,CBAM 模块在利用通道相

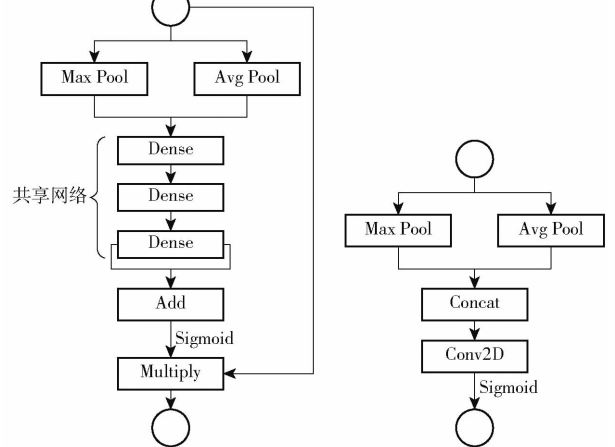


图 4 CBAM 各部分结构

Fig. 4 Structure of each part in CBAM

关性对通道注意力进行计算时,同样需要先对给定的特征图 $\mathbf{X} \in \mathbf{R}^{H \times W \times C}$ 进行挤压操作,然后同时利用平均池化和最大池化聚合特征图的空间信息,以获取更加精细的通道注意力。上述 2 种池化操作后分别生成 2 个不同的空间上下文描述符 $\mathbf{X}_{\text{avg}}^c$ 和 $\mathbf{X}_{\text{max}}^c$,将 2 个描述符传递到共享网络中,然后使其输出进行逐元素相加求和,最终生成通道注意力特征权重 $\mathbf{M}_c \in \mathbf{R}^{1 \times 1 \times C}$ 。其中,共享网络由具有 1 个隐藏层的多层感知机组成,为了减少参数计算量,共享网络中的隐藏层激活大小设置为 $1 \times 1 \times \frac{C}{r}$,其中 r 是降维比。

通道注意力的计算过程为

$$\begin{aligned} \mathbf{M}_c(\mathbf{X}) &= \\ \sigma(\text{MLP}(\text{AvgPool}(\mathbf{X})) + \text{MLP}(\text{MaxPool}(\mathbf{X}))) &= \\ \sigma(\mathbf{W}_1(\mathbf{W}_0(\mathbf{X}_{\text{avg}}^c)) + \mathbf{W}_1(\mathbf{W}_0(\mathbf{X}_{\text{max}}^c))) \end{aligned} \quad (2)$$

式中 $\sigma(\cdot)$ ——S 型函数

$\mathbf{W}_0$ 、 $\mathbf{W}_1$ ——输入共享的 MLP 权重

$\text{MLP}(\cdot)$ ——多层感知机模块

$\text{AvgPool}(\cdot)$ ——平均池化层

$\text{MaxPool}(\cdot)$ ——最大池化层

由于苹果叶部病斑形态各异,并且同一片叶片上可能在多个位置出现小而密集的病斑,因此仅仅使用通道注意力进行特征提取容易造成对小目标信息的丢失。空间注意力模块作为对通道注意力模块的补充,主要用于获取重要特征的空间位置,实现对特征间空间相关性的描述,最终实现对苹果叶部病斑的准确定位。空间注意力模块的具体结构如图 4b 所示。

在计算空间注意力特征图时,首先通过最大池化和平均池化操作沿特征通道轴方向汇聚特征图的通道信息生成 2 组特征描述符,分别表示特征通道的最大池化特征和平均池化特征。然后将 2 个特征图进行串联拼接并通过一个标准卷积层生成空间注意力特征图 $\mathbf{M}_s(\mathbf{X}) \in \mathbf{R}^{H \times W}$,从而实现对强化或抑制的空间位置的编码。空间注意力计算公式为

$$\begin{aligned} \mathbf{M}_s(\mathbf{X}) &= \sigma(f([\text{AvgPool}(\mathbf{X}); \text{MaxPool}(\mathbf{X})])) = \\ &\sigma(f([\mathbf{X}_{\text{avg}}^s; \mathbf{X}_{\text{max}}^s])) \end{aligned} \quad (3)$$

式中 $f(\cdot)$ ——卷积运算

$\mathbf{X}_{\text{avg}}^s$ ——经过平均池化处理的特征图

$\mathbf{X}_{\text{max}}^s$ ——经过最大池化处理的特征图

1.3.2 焦点损失函数

苹果叶部病害样本存在获取困难且分布不平衡的问题,因此模型对不同病害样本的分类难度具有极大的差异性。上述问题原因在于训练过程中采用的原始交叉熵损失函数很难刻画这种不均衡分布特征,因而导致分类效果较差。为了解决这个问题,本

文引入了焦点损失函数(Focal loss)^[13],在一定程度上解决苹果叶部病害数据分布不平衡问题。焦点损失函数是在交叉熵损失函数(Cross entropy)基础上进行修改得到的。多分类交叉熵损失公式为

$$\text{CE}(y, y') = - \sum y'_i \text{lb} y_i \quad (4)$$

式中 $\text{CE}(\cdot)$ ——多分类交叉熵损失函数

y ——样本标签

y' ——经过 Softmax 激活函数的输出概率

焦点损失函数本质上是在交叉熵损失函数的基础上添加调节因子,其中 γ 通过降低易分类样本的权重使得模型在训练时更加关注难分类的样本。

$$\text{FL}(y, y') = - \sum \alpha_i (1 - y'_i)^{\gamma} \text{lb} y'_i \quad (5)$$

其中

$$\alpha_i = 1 - \frac{n_i}{N} \quad (6)$$

式中 $\text{FL}(\cdot)$ ——焦点损失函数

α_i ——平衡因子,用来平衡各类样本比例

n_i ——第 i 类样本数量

N ——数据集中样本总数

1.3.3 快照集成

由于单一卷积网络模型在病害分类识别上的局限性和苹果叶部病害病斑特征的复杂性,将多个能够独立进行判别分析的网络模型进行集成获得的效果一般会比单网络模型效果好。然而,由于训练深度神经网络耗时耗力的特性,训练后集成不同模型的代价很高。

神经网络可能具有多个局部极小值,每个局部极小值对不同的特征赋予不同的权重,并包含一些有价值的信息。然而,传统的神经网络训练方法往往收敛到某个局部极小值,导致信息缺失。在这种情况下,为了提高模型的性能,本文拟采用集成方式来整合不同极小值的信息。文献[14]提出了一种在不增加训练代价的情况下实现多神经网络集成的方法。首先训练单个神经网络,沿其最优路径收敛到局部极小值并保存模型参数;此后,为了获得重复的快速收敛性,采用余弦函数控制学习率,该策略被称为余弦退火策略(Cosine annealing),其学习率公式为

$$\alpha(t) = \frac{\alpha_0}{2} \left(\cos \left(\frac{\pi \text{mod}(t-1, \lfloor T/M \rfloor)}{\lfloor T/M \rfloor} \right) + 1 \right) \quad (7)$$

式中 $\alpha(t)$ ——第 t 次迭代时的学习率

α_0 ——初始学习率

T ——迭代次数 M ——循环次数

$\text{mod}(\cdot)$ ——模运算函数

$\lfloor \cdot \rfloor$ ——向下取整操作符

结合图 5 可以看出,在传统的训练过程中学习率往往呈指数衰减趋势,因此模型会逐渐找到局部最优点。该过程中,由于学习率在初始阶段具有较大的值,因此模型不会进入陡峭的局部最优点,当学习率下降时会很快向平坦的局部最优点移动。然而采用余弦退火策略的模型将经历多个学习率的退火循环周期,在收敛到多个局部最小值后,学习率将大大提高,并从极值点逃逸。模型的多样性很大,具有不同的局部最优点,所以集成后的模型效果会更好。

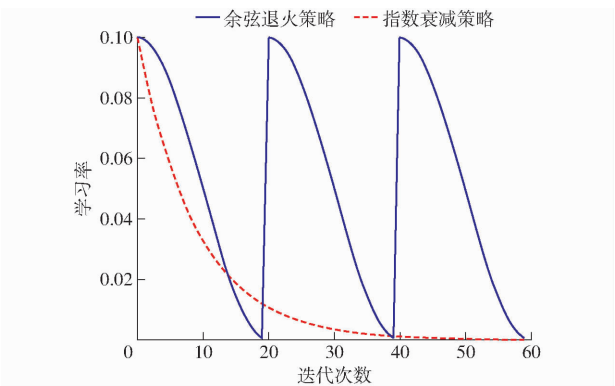


图 5 指数衰减策略与余弦退火策略

Fig. 5 Comparison of exponential decay strategy and cosine annealing strategy

这种优化集成方法称为快照集成 (Snapshot ensembling),与单一模型相比,通过快照集成方法获得的多个集成子模型有以下 2 个性质:①每个模型都有比较低的错误率。②模型两两之间对识别错误的例子没有交集。基于这 2 个性质就可以进行模型集成。

经过 M 次循环后,算法在训练过程中保存了 M 个快照模型

$$h_{\text{Ensemble}} = \frac{1}{m} \sum_{i=0}^{m-1} h_{M-i}(x) \quad (8)$$

式中 x ——测试样本

$h_{M-i}(x)$ ——第 $M-i$ 个快照模型经过 Softmax 激活函数的输出

h_{Ensemble} ——集成模型的输出,是最后 m 个模型 ($m \leq M$) 的简单平均值

针对每幅图像,将每个“快照”模型的概率预测进行串联,形成一个新的数据项。考虑到使用 XGBoost 模型作为元模型可以从预测中提取有用的信息,将这些新数据构成的特征集输入到 XGBoost 模型中,以构建最终的苹果叶部病害程度识别模型。图 6 为该模型的整体结构(以 3 次循环为例)。

2 实验

2.1 环境配置和超参数设置

实验采用表 1 所示的苹果叶部病害数据集,实

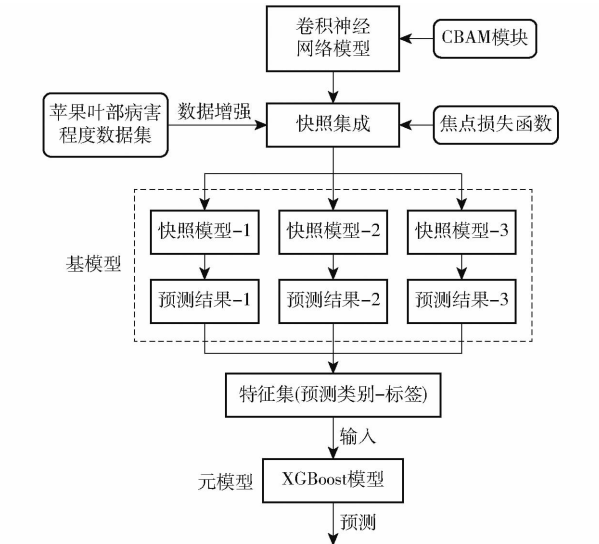


图 6 基于快照集成的苹果叶部病害程度识别模型原理图

Fig. 6 Diagram of apple leaf disease identification model based on snapshot ensembling

验环境配置如表 2 所示。

表 2 实验环境配置

Tab. 2 Experimental configuration

参数	类型/数值
CPU	Intel(R) Xeon(R) CPU E5-2650 v4 @
	2.20GHz(X2)
GPU	GP100GL [Tesla P100 PCIe 16GB]
操作系统	Ubuntu 16.04.2 LTS (64-bit)
内存容量/GB	503(系统内存)
硬盘容量/TB	2
语言	Python 3.6
深度学习框架	TensorFlow 2.0.0
超参数调优框架	Ray.tune 0.84

为了能快速计算导数以及使损失函数脱离鞍点和较差的局部最优点,网络参数更新方法采用随机梯度下降 (Stochastic gradient descent, SGD)^[15],即每次计算梯度时随机选择样本库中的一个样本(或少量样本)进行计算。但是若只采用 SGD 的训练方法,权值更新将完全依赖当前批次,导致更新过程非常不稳定,因此需要引入动量方法。动量方法权值更新过程为

$$\begin{cases} v_{dw} \leftarrow \beta v_{dw} + (1 - \beta) dw \\ w \leftarrow w - \alpha v_{dw} \end{cases} \quad (9)$$

式中 α ——学习率

β ——动量系数,取 0.9

v_{dw} ——动量 dw ——权值衰减

为了评估模型的性能,采用多分类场景下的准确率 (Accuracy) 为评价指标。其他神经网络训练过程所需超参数的设置为:基础学习率设为 0.1,批次

数量为 32,迭代次数设为 60。为确保实验的可复现性,随机种子固定设置为 42。此外,在输入网络前所有叶部病害图像尺寸统一被裁剪为 256 像素 × 256 像素。

2.2 最优识别模型

苹果叶部病害的大部分病斑形状相似,且具有小而密集的特点,因此首先分析了不同卷积神经网络模型的效果对比,例如 VGG19^[16]、AlexNet^[17]、

ResNet101、ResNet152、GoogLeNet V2^[18]、改进 DenNet121^[19]、XDNet^[20] 和 Inception - ResNet V2^[21]。首先基于 2.1 节描述的实验环境和数据增强后苹果病害数据集进行训练,然后比较上述模型在同等训练超参数设置时在测试集上的表现。本文采用测试集上对不同程度病害的识别准确率和识别 1 幅苹果叶部病害图像所需的推理时间作为苹果病害程度识别模型的测评指标,最终结果如表 3 所示。

表 3 各模型在测试集上的性能指标
Tab.3 Performance of selected models on test set

模型	准确率/%					总体准确	推理时间/
	健康	黑星病早期	黑星病晚期	锈病早期	锈病晚期	率/%	ms
VGG19	63. 52	75. 41	59. 22	69. 38	71. 79	67. 86	12. 4
AlexNet	78. 22	75. 64	76. 06	75. 83	79. 9	77. 13	11. 6
ResNet101	74. 73	81. 64	72. 02	75. 64	77. 98	76. 24	11. 2
ResNet152	76. 66	73. 63	71. 06	72. 3	79. 23	74. 58	11. 2
GoogLeNet V2	84. 08	83. 26	83. 31	84. 26	86. 33	84. 25	11. 1
改进 DenNet121	86. 84	85. 82	87. 01	86. 69	86. 71	86. 61	11. 5
XDNet	85. 86	86. 99	87. 83	87. 52	89. 77	87. 59	11. 3
Inception - ResNet V2	86. 53	87. 63	88. 67	89. 49	89. 26	88. 32	10. 8

由实验结果和表 3 可知,对锈病晚期识别准确率仅比 XDNet 网络低 0.51 个百分点的情况下,Inception - ResNet V2 模型对其他病害程度识别准确率和总体准确率均为最佳,并在同等实验条件下识别 1 幅病害图像推理所耗时间最短,在苹果病害程度识别任务中性能优越。由于苹果叶部不同的病害具有不同的特性,如锈病的病斑较小且处于早期时比较稀疏,大多低于 28 像素 × 28 像素,晚期锈病的病斑又变得密集;黑星病的病斑大多高于 46 像素 × 56 像素,通常比较稀疏,所以固定尺度的卷积核在感知不同尺寸的病斑时效果不佳。加之苹果叶部病害识别是一个细粒度图像分类任务,其子类别间类间差异较小,因此与普通分类任务相比本文的病害程度图像分类难度更大。

由以上分析可知,使用单一尺度的卷积核来识别多尺度的病斑效果较差且效率较低。通过分析 Inception - ResNet V2 网络的结构可知,其中的 Inception 模块选取了 2 种不同尺度(3 × 3 和 7 × 7)的卷积核组合进行多尺度特征提取的方式,这使得网络能够感知不同尺度的特征,因此能够较好地感知不同尺寸的病斑,提高网络的特征提取能力。同时,Inception - ResNet V2 引入了残差网络跳跃连接的方法,加深网络深度的同时避免了梯度消失、梯度爆炸的问题,进一步提高了网络性能。另外,Inception - ResNet V2 网络将使用的 3 × 3 和 7 × 7 卷积核进行分解,在加深网络的同时减少了涉及的参

数量,加快了网络的推理速度。

综上所述,Inception - ResNet V2 由于其出色的特征提取能力与随之带来的较之其他网络更高的识别准确率,可以作为后续快照集成工作的基模型。

2.3 数据增强对模型识别性能的影响

以 2.2 节实验结果中性能最优的 Inception - ResNet V2 模型为基础网络模型,探究数据增强操作对最终分类效果性能的影响。模型在训练集和验证集上的准确率对比曲线如图 7 所示。

由图 7 可知,在使用原始数据集进行训练的过程中,网络存在着严重的过拟合现象,且验证集上识别准确率只有 83.17%。当使用增强后的数据集样本进行训练后,可以发现过拟合现象有极大改善,验证集上识别准确率也显著上升,达到了 87.84%。

通过以上实验,可以看出当数据集样本量较少时,模型的识别准确率较低,并且模型容易过拟合,模型性能差;而随着数据集样本量的增大,过拟合现象逐渐消失,模型性能也逐渐得以提升,识别准确率及模型的泛化能力也得到一定程度的提升。因此,基于数据增强方法去扩充苹果病害样本数量是解决模型过拟合和提高准确率的一个有效方法。

2.4 卷积块注意力模块和焦点损失函数对模型性能的影响

在 Inception - ResNet V2 模型基础上,进一步验证注意力模块以及焦点损失函数对最终分类效果的影响。首先对其引入基于卷积块注意力模块的注意

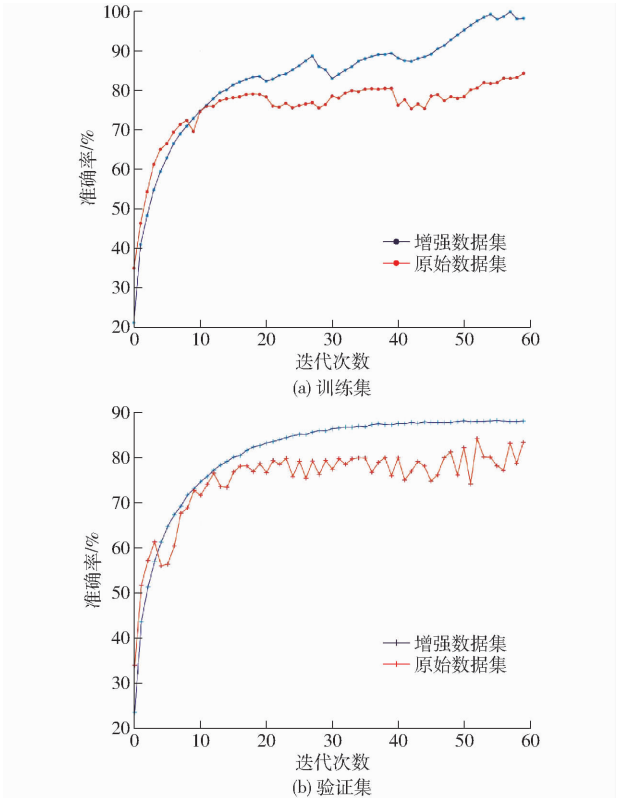


图 7 准确率对比曲线

Fig. 7 Comparison curves of accuracy before and after data augmentation

力机制(记为 CBAM),与原模型(记为 Vanilla)的实验结果进行对比,验证卷积块注意力模块是否对模型性能有提升效果。然后将原先的交叉熵损失函数(记为 CE)更换为焦点损失函数(记为 FL),验证其有效性。以测试集准确率为评估指标,最终消融实验可得 Vanilla + CE、Vanilla + FL、CBAM + CE 和

CBAM + FL 的准确率分别为 88.32%、88.78%、88.91% 和 89.79%。

由消融实验结果可知,引入 CBAM 模块的 Inception - ResNet V2 模型比原模型的准确率高 1.59 个百分点。另外,类激活图(Class activation map, CAM)可视化有助于了解让模型做出正确分类决策的苹果叶部病害图像的部位。因此,为了进一步分析 CBAM 模块对模型病害特征提取能力的提升效果,实验提取了部分测试集图像在经过模型最后卷积层后的类激活图,如图 8 所示。可以看出,引入 CBAM 模块后的 Inception - ResNet V2 模型可以有效捕捉到发病部位及提取病害特征,从而提高苹果病害程度图像识别这种细粒度分类任务下的模型性能。

当使用原模型时,损失函数更换为焦点损失函数后准确率相比交叉熵损失函数提升 0.46 个百分点,说明焦点损失函数的引入缓解了数据集样本不平衡的问题。二者组合的准确率相比基线模型(Baseline)提升了 1.47 个百分点,说明使用卷积块注意力模块和焦点损失函数对最终分类效果的提升有积极影响。

2.5 快照集成对模型性能的影响

选取 2.4 节中的改进 Inception - ResNet V2 模型为研究对象,先将余弦退火策略中的学习率循环次数作为一个超参数(循环 N 次便保存 N 个快照模型),以测试集准确率为评估标准,探究快照集成对最终模型性能的影响。

通过图 9 可以看出,当集成 3 个快照模型时整体模型的性能最好。说明合理地对快照模型的数量

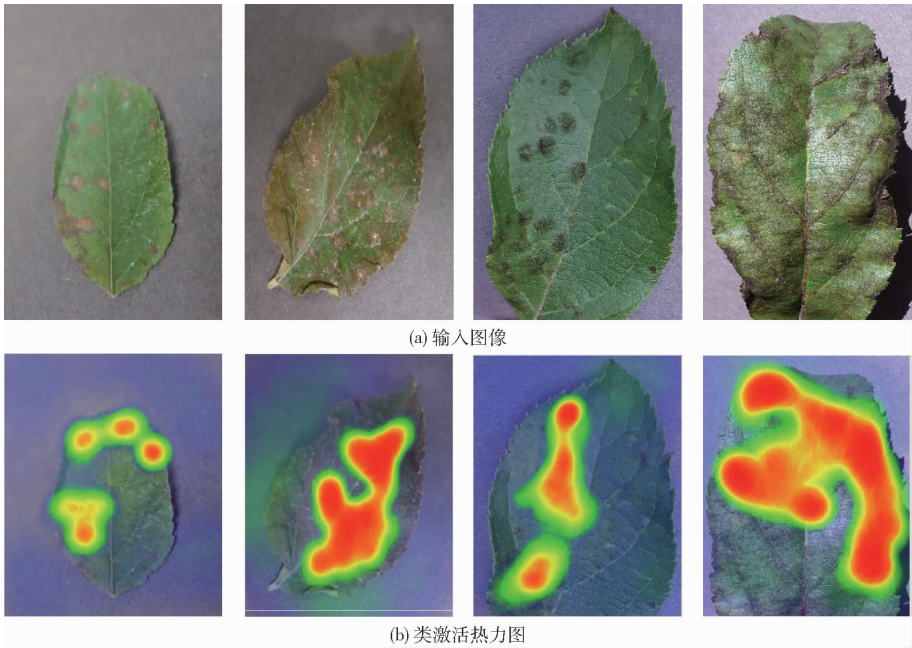


图 8 苹果叶部病害图像类激活热力图

Fig. 8 Class activation maps of apple leaf diseases

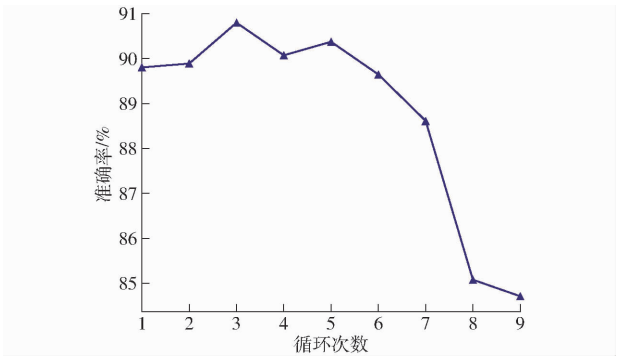


图 9 不同循环次数下快照集成模型的测试集准确率
Fig.9 Accuracy on test set of snapshot ensemble model under different cycles

进行调整也能够改善最终集成模型的性能。集成模型的核心是要求各个子模型必须具有多样性,且每个子模型本身要具有超过随机猜测的准确性。这个理论反映在本实验中即为快照集成方法中的快照子网络并不是越多越好,越多的快照表示相邻 2 次保存操作间隔的时间越短,意味着保存的快照模型之间的相关性越高,集成模型的性能也就越差。快照模型的数量应该根据使用的基础模型和具体的任务进行适当调整,从而找到最佳循环次数。同时,循环次数为 1 表示只保存 1 个快照模型,这等同于使用单一模型进行训练,此时的准确率只有 89.79%,相

比集成最好结果(90.82%)降低了 1.03 个百分点,侧面验证了快照集成方法相比单一模型对模型性能有提升效果。另外,最好结果(90.82%)与单一 Inception – ResNet V2 模型的准确率(88.32%)相比,提高了 2.50 个百分点,表明了注意力机制与快照集成等改进方法的有效性。

3 结论

(1)以苹果黑星病和锈病的早期与晚期病害图像为研究对象,构建了基于快照集成卷积神经网络的苹果叶部病害程度识别模型,实现了对苹果叶部病害程度高准确率的识别。

(2)基于数字图像处理技术对原始数据集进行数据增强,有效地缓解了过拟合现象且提升了识别准确率。

(3)选取了与本文病害识别任务契合度最高的 Inception – ResNet V2 为基础网络,同时引入卷积块注意力模块(CBAM)和焦点损失函数,构建快照集成模型,利用苹果黑星病和锈病的早期和晚期病害数据集进行了模型验证,准确率高达 90.82%,与单一 Inception – ResNet V2 模型 88.32% 的准确率相比,提高了 2.50 个百分点。实验表明该集成模型具有高准确率和良好的泛化能力。

参 考 文 献

[1] 鲍文霞,赵健,张东彦,等.基于椭圆型度量学习的小麦叶部病害识别[J].农业机械学报,2018,49(12):20–26.
BAO Wenxia, ZHAO Jian, ZHANG Dongyan, et al. Recognition of wheat leaf diseases based on elliptic metric learning[J]. Transactions of the Chinese Society for Agricultural Machinery, 2018, 49(12): 20–26. (in Chinese)

[2] 翟肇裕,曹益飞,徐焕良,等.农作物病虫害识别关键技术研究综述[J].农业机械学报,2021,52(7):1–18.
ZHAI Zhaoyu, CAO Yifei, XU Huanliang, et al. Review of key techniques for crop disease and pest detection[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(7): 1–18. (in Chinese)

[3] 杨森,冯全,张建华,等.基于深度学习与复合字典的马铃薯病害识别方法[J].农业机械学报,2020,51(7):22–29.
YANG Sen, FENG Quan, ZHANG Jianhua, et al. Identification method for potato disease based on deep learning and composite dictionary[J]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(7): 22–29. (in Chinese)

[4] 许景辉,邵明烨,王一琛,等.基于迁移学习的卷积神经网络玉米病害图像识别[J].农业机械学报,2020,51(2):230–236,253.
XU Jinghui, SHAO Mingye, WANG Yichen, et al. Recognition of corn leaf spot and rust based on transfer learning with convolutional neural network[J]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(2): 230–236, 253. (in Chinese)

[5] 任守纲,贾馥玮,顾兴健,等.反卷积引导的番茄叶部病害识别及病斑分割模型[J].农业工程学报,2020,36(12):186–195.
REN Shougang, JIA Fuwei, GU Xingjian, et al. Recognition and segmentation model of tomato leaf diseases based on deconvolution-guiding[J]. Transactions of the CSAE, 2020, 36(12): 186–195. (in Chinese)

[6] 孙云云,江朝晖,董伟,等.基于卷积神经网络和小样本的茶树病害图像识别[J].江苏农业学报,2019(1):48–55.
SUN Yunyun, JIANG Zhaohui, DONG Wei, et al. Image recognition of tea plant disease based on convolutional neural network and small samples[J]. Jiangsu Journal of Agricultural Sciences, 2019(1): 48–55. (in Chinese)

[7] 马浚诚,温皓杰,李鑫星,等.基于图像处理的温室黄瓜霜霉病诊断系统[J].农业机械学报,2017,48(2):195–202.
MA Juncheng, WEN Haojie, LI Xinxing, et al. Downy mildew diagnosis system for greenhouse cucumbers based on image processing[J]. Transactions of the Chinese Society for Agricultural Machinery, 2017, 48(2): 195–202. (in Chinese)

[8] SHORTEN C, KHOSHGOFTAAR T M. A survey on image data augmentation for deep learning[J]. Journal of Big Data, 2019, 6(1):1–48.

[9] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks[J]. Communications of the ACM, 2017, 60(6):84–90.

[10] ZHANG H, CISSE M, DAUPHIN Y N, et al. Mixup: beyond empirical risk minimization[J]. arXiv e-prints, 2017, 30(3):

arXiv – 1710.

[11] WOO S, PARK J, LEE J Y, et al. CBAM: convolutional bock attention module[C]//Proceedings of the European Conference on Computer Vision (ECCV), 2018: 3 – 19.

[12] HU J, SHEN L, SUN G. Squeeze-and-Excitation networks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 7132 – 7141.

[13] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection[C]//Proceedings of the IEEE International Conference on Computer Vision, 2017: 2980 – 2988.

[14] HUANG G, LI Y, PLEISS G, et al. Snapshot ensembles: train 1, get m for free[J]. arXiv preprint arXiv:1704.00109, 2017.

[15] BOTTOU L. Stochastic gradient descent tricks[M]//Neural networks: tricks of the trade. Springer, Berlin, Heidelberg, 2012: 421 – 436.

[16] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[J]. arXiv preprint arXiv: 1409.1556, 2014.

[17] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition: Volume 12, 2016: 770 – 778.

[18] IOFFE S, SZEGEDY C. Batch normalization: accelerating deep network training by reducing internal covariate shift[C]//International Conference on Machine Learning. PMLR, 2015: 448 – 456.

[19] CHAO X, SUN G, ZHAO H, et al. Identification of apple tree leaf diseases based on deep learning models[J]. Symmetry, 2020,12(7):1065.

[20] YONG Z, MING Z. Research on deep learning in apple leaf disease recognition[J]. Computers and Electronics in Agriculture, 2020, 168(1):105146.

[21] SZEGEDY C, IOFFE S, VANHOUCKE V, et al. Inception – v4, Inception – ResNet and the impact of residual connections on learning[C]//Proceedings of the AAAI Conference on Artificial Intelligence, 2017.

(上接第 208 页)

[12] 陶惠林,徐良骥,冯海宽,等. 基于无人机高光谱遥感的冬小麦株高和叶面积指数估算[J]. 农业机械学报,2020,51(12):193 – 201.

TAO Huilin, XU Liangji, FENG Haikuan, et al. Estimation of plant height and leaf area index of winter wheat based on UAV hyperspectral remote sensing[J]. Transactions of the Chinese Society for Agricultural Machinery,2020,51(12):193 – 201. (in Chinese)

[13] 牛庆林,冯海宽,杨贵军,等. 基于无人机数码影像的玉米育种材料株高和 LAI 监测[J]. 农业工程学报,2018,34(5):73 – 82.

NIU Qinglin, FENG Haikuan, YANG Guijun, et al. Monitoring plant height and leaf area index of maize breeding material based on UAV digital images[J]. Transactions of the CSAE, 2018,34(5):73 – 82. (in Chinese)

[14] 董文,范祺祺,叶亦心,等. 施氮和施钾水平对冬播马铃薯生长及产量的影响[J]. 湖南农业大学学报(自然科学版), 2021,47(4):392 – 398.

DONG Wen, FAN Qiqi, YE Yixin, et al. Effect of nitrogen and potassium rates on the growth and yield of winter potato crop [J]. Journal of Hunan Agricultural University(Natural Sciences), 2021,47(4):392 – 398. (in Chinese)

[15] 杨福芹,李天驰,冯海宽,等. 基于无人机数码影像的冬小麦氮素营养诊断研究[J]. 福建农业学报,2021,36(3):369 – 378.

YANG Fuqin, LI Tianchi, FENG Haikuan, et al. UAV digital image-assisted monitoring on nitrogen nutrition of winter wheat in the field[J]. Fujian Journal of Agricultural Sciences,2021,36(3):369 – 378. (in Chinese)

[16] 何彩莲,郑顺林,万年鑫,等. 马铃薯光谱及数字图像特征参数对氮素水平的响应及其应用[J]. 光谱学与光谱分析, 2016,36(9):2930 – 2936.

HE Cailian, ZHENG Shunlin, WAN Nianxin, et al. Photo spectrum and the digital image feature parameters on the response of the nitrogen level and its application[J]. Spectroscopy and Spectral Analysis, 2016,36(9):2930 – 2936. (in Chinese)

[17] GITELSON A A, VINA A, ARKEBAUER T J, et al. Remote estimation of leaf area index and green leaf biomass in maize canopies[J]. Geophysical Research Letters, 2003, 30(5):1248.

[18] MEYER G E, NETO J C. Verification of color vegetation indices for automated crop imaging applications[J]. Computers and Electronics in Agriculture, 2008, 63(2): 282 – 293.

[19] BENDIG J, YU K, AASEN H, et al. Combining UAV-based plant height from crop surface models, visible, and near infrared vegetation indices for biomass monitoring in barley[J]. International Journal of Applied Earth Observation and Geoinformation, 2015, 39: 79 – 87.

[20] 陶惠林,徐良骥,冯海宽,等. 基于无人机数码影像的冬小麦株高和生物量估算[J]. 农业工程学报,2019,35(19):107 – 116.

TAO Huilin, XU Liangji, FENG Haikuan, et al. Estimation of plant height and biomass of winter wheat based on UAV digital image[J]. Transactions of the CSAE,2019,35(19):107 – 116. (in Chinese)

[21] YUE J, YANG G, TIAN Q, et al. Estimate of winter-wheat above-ground biomass based on UAV ultrahigh-ground-resolution image textures and vegetation indices[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2019, 150: 226 – 244.

[22] YUE J, YANG G, LI C, et al. Estimation of winter wheat above-ground biomass using unmanned aerial vehicle-based snapshot hyperspectral sensor and crop height improved models[J]. Remote Sensing, 2017, 9(7): 708.