

基于卷积神经网络的农作物病虫害检测研究

白雪松¹, 吴建平^{1,2,3}, 景文超¹, 何旭鑫¹, 余咏¹

(1. 云南大学 信息学院, 云南 昆明 650504;

2. 云南省电子计算中心, 云南 昆明 650223;

3. 云南省高校数字媒体技术重点实验室, 云南 昆明 650223)

摘要: 农作物病虫害图像采集困难, 且公共数据集较少, 因此使用单一网络模型识别准确率不高。常用的数据增强方法只能对图像进行像素空间的变换, 不能进行语义转换。对此, 研究并提出基于隐式语义数据增强算法的 CBAM-Res2Net50 模型: 该模型使用 Res2Net50 主干网络, 从多尺度学习图像信息, 加载预训练模型部分参数, 提高模型的收敛速度; 在网络残差块中添加混合注意力模块, 提取并保留关键特征; 训练过程中使用隐式语义数据增强算法对提取的深层网络空间特征进行语义扩充增强, 提高模型的泛化能力。改进模型与现有模型在 AI Challenger 2018 农作物病虫害数据集上的对比实验结果表明: 改进模型具有较高的识别率, 其分类准确率达 88.33%。改进后的模型通过挖掘相似病虫害图像的语义信息, 在一定程度上解决了深度网络中由于数据不足导致的过拟合等问题。

关键词: 农作物病虫害; 卷积神经网络; CBAM-Res2Net50; 迁移学习; 注意力模块; 隐式语义数据增强

中图分类号: TP312

文献标识码: A

文章编号: 1673-629X(2022)12-0200-06

doi:10.3969/j.issn.1673-629X.2022.12.030

Research on Crop Diseases and Pests Detection Based on Convolutional Neural Network

BAI Xue-song¹, WU Jian-ping^{1,2,3}, JING Wen-chao¹, HE Xu-xin¹, YU Yong¹

(1. School of Information Science & Engineering, Yunnan University, Kunming 650504, China;

2. Yunnan Provincial Electronic Computing Center, Kunming 650223, China;

3. Key Laboratory of Digital Media Technology of Universities and Colleges in Yunnan Province, Kunming 650223, China)

Abstract: It is difficult to collect images of crop diseases and pests, and there are few common data sets, so the recognition accuracy of single network model is not high. The commonly used data augmentation methods can only transform image pixel space, but not semantic transformation. Therefore, CBAM-Res2Net50 model based on implicit semantic data augmentation algorithm is studied and proposed. The model uses Res2Net50 backbone network to learn image information from multi-scale and load some parameters of pre-training model to improve the convergence speed of the model. A mixed attention module is added to the network residual block to extract and retain key features. In the training process, implicit semantic data augmentation algorithm is used to enhance semantic expansion and generalization ability of the model. The experimental results of comparison between the improved model and the existing model on AI Challenger 2018 Crop Diseases and Pests Data Set show that the improved model has a high recognition rate, and its classification accuracy is 88.33%. The improved model can solve the problem of over-fitting caused by insufficient data in deep network to some extent by mining semantic information of similar images of diseases and pests.

Key words: crop diseases and pests; convolutional neural network; CBAM-Res2Net50; transfer learning; attention module; implicit semantic data augmentation

收稿日期: 2022-01-14

修回日期: 2022-05-19

基金项目: 云南省重大科技专项计划项目(202002AD080001); 云南省科技厅应用基础研究计划重点项目(2019FA044); 云南大学第一届专业学位研究生实践创新项目(2021Y183)

作者简介: 白雪松(1996-), 男, 硕士研究生, 研究方向为图像处理、AI 技术; 通信作者: 吴建平(1965-), 男, 高级工程师, 硕士, 研究方向为智能化、信息化技术。

0 引言

农作物病虫害严重影响农产品的质量和产量,给农业生产带来了重大损失。传统的农作物病虫害检测主要是通过现场检测或实验室化学分析,对每一株农作物进行检查使得农技人员面临着大量的重复性工作,不仅费时费力而且要求检测人员具备该领域的专业知识及丰富的经验^[1]。此外,受个人主观因素影响可能无法诊断某些具有高变异性的病虫害,而导致错误的结论和治疗方案^[2]。

人工智能(Artificial Intelligence, AI)技术在农业检测领域处理复杂问题时表现优秀。卷积神经网络(Convolutional Neural Network, CNN)在国内外农作物病虫害检测研究中已形成热点。Sladojevic 等^[3]最早采用 AlexNet 神经网络检测农作物病虫害,并采用多种训练方法提高模型的泛化能力。Liu 等^[4]将 AlexNet 网络和 GoogLeNet 网络融合成一种新模型,新模型经过训练可以有效地识别 4 种不同病虫害的苹果叶病。Too 等^[5]在农作物病虫害检测实验中对各类模型进行评估训练,包括 ResNet50、VGG16、InceptionV4 等主流网络。王东方等^[6]对 SE-ResNeXt-101 模型进行改进,并结合迁移学习提出了 TL-SE-ResNeXt-101 模型,在试验中取得较好的识别效果。刘阳等^[7]优化轻量级网络 SqueezeNet 的部分结构,在不降低识别精度的情况下,大大减少了模型计算量和参数存储量。樊湘鹏等^[8]建立了基于迁移学习的 VGG16 改进模型,对网络部分结构进行优化训练,最后将 VGG16 改进模型移植到手机端。

上述研究工作的识别效果非常不错,在农作物病虫害识别任务中取得较大进展,但依旧存在一些问题亟待解决。首先是数据样本不足导致某些具有变异性的疾病经常与相似疾病混淆,导致分类准确率降低。其次,大多数样本都是在受限环境下采集的,没有考虑到复杂环境对样本的干扰。此外,相当多的研究采用模型融合的方法,大大增加了模型参数量及计算量。目前,大部分研究者采用传统数据增强技术进行处理,但没有能对图片进行角度、纹理等语义变换,效果受限。虽然文献[9]采用生成对抗网络(Generative Adversarial Net, GAN)来扩增数据集,但训练 GAN 增加了研究人员的工作量。

针对上述问题,该文提出了一种基于隐式语义数据增强算法的残差网络优化模型。综合考虑性能、参数数量、计算复杂度等因素,选择 Res2Net50 网络进行改进优化。训练之前,对样本数据进行传统数据增强。为提取数据图像的关键特征信息,在网络中添加混合注意力模块。在训练过程中,使用隐式语义数据增强算法进行网络优化。

1 相关网络模型

1.1 Res2Net50 残差模块

网络层次加深会增强网络的特征提取能力。随着网络深度增加到一定程度,网络可能会出现退化现象。在 2015 年,何恺明等^[10]提出了性能表现优越的残差网络(Residual Network, ResNet),利用恒等映射的思想创建残差块,避免发生网络退化现象。最近, Gao 等^[11]提出了新的多尺度网络架构(A New Multi-scale Backbone Architecture, Res2Net),在残差网络的基础上使用分层、层叠的特征组卷积(scale)取代原来的 3×3 卷积,增加每一层的感受野,提高了网络的特征表达能力。

该文采用 Res2Net50 网络模型,用 4 组 scale 架构替换一般的 3×3 卷积,以模型迁移的方式加载网络权重,修改全连接层分类数目。此外, Res2Net50 网络残差块中加入混合注意力模块后,能在多个通道提取特征再合并输出特征,结合通道和空间维度自适应调整特征图,保留关键特征。混合注意力模块将在 1.3 节进行详细介绍。

改进后的残差块结构如图 1 所示。

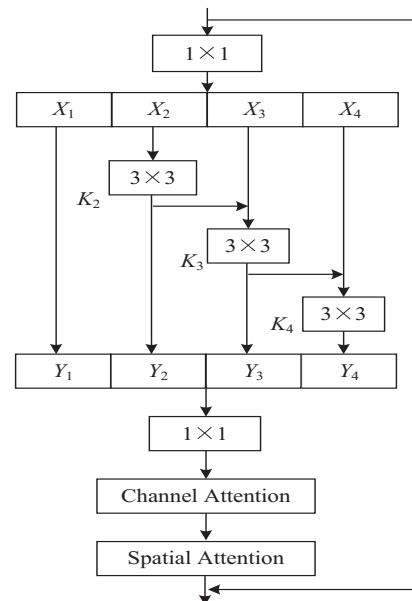


图 1 改进后的残差块结构

1.2 隐式语义数据增强算法

农作物病虫害数据集的样本种类分布不均匀且图像在获取时受到拍摄角度、光照和其他因素的影响。为了缓解样本不平衡问题,避免出现网络过拟合现象,对数据集进行数据增强是一种简单有效的方法。传统数据增强方法虽然在一定程度提高了模型鲁棒性,但只能对图像进行像素空间的变换。

为了在传统数据增强的基础上实现图像的语义增强,该文采用一种新型的损失函数——隐式语义数据增强算法(Implicit Semantic Data Augmentation,

ISDA)^[12]。ISDA 算法在交叉熵损失函数的基础上优化,通过动态估计样本的协方差矩阵来进行语义增强。

协方差矩阵通过特征的空间位置捕获样本的差异信息^[13],选择有意义的语义方向进行增强。语义信息估计方法如下:

$$\mathbf{u}_j^{(t)} = \frac{n_j^{(t-1)} \mathbf{u}_j^{(t-1)} + m_j^{(t)} \mathbf{u}_j'^{(t)}}{n_j^{(t-1)} + m_j^{(t)}} \quad (1)$$

$$\Sigma_j^{(t)} = \frac{n_j^{(t-1)} \Sigma_j^{(t-1)} + m_j^{(t)} \Sigma_j'^{(t)}}{n_j^{(t-1)} + m_j^{(t)}} + \frac{n_j^{(t-1)} m_j^{(t)} (\mathbf{u}_j^{(t-1)} - \mathbf{u}_j'^{(t)}) (\mathbf{u}_j^{(t-1)} - \mathbf{u}_j'^{(t)})^T}{(n_j^{(t-1)} + m_j^{(t)})^2} \quad (2)$$

$$n_j^{(t)} = n_j^{(t-1)} + m_j^{(t)} \quad (3)$$

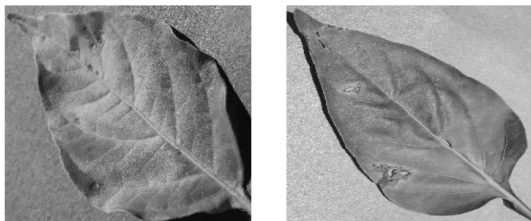
在公式(1)~公式(3)中, $\mathbf{u}_j^{(t)}$ 和 $\Sigma_j^{(t)}$ 分别对应于 j 类特征在第 t 次迭代时的均值和协方差矩阵的估计值。 $n_j^{(t)}$ 是所有 t 个批次处理中属于 j 类的样本总数,而 $m_j^{(t)}$ 是仅在第 t 个批次处理中属于 j 类的样本数量。 $\mathbf{u}_j'^{(t)}$ 和 $\Sigma_j'^{(t)}$ 分别是 j 类特征在第 t 个批次中的均值和协方差矩阵估计值。语义增强方向 \bar{a}_i ,在 $N(a_i, \lambda \Sigma_{y_i})$ 中进行选择,其中 λ 是控制语义数据增强的强度系数。

此外,对交叉熵损失函数进行优化得到计算上限,避免因语义增强而造成的计算浪费。优化后的交叉熵损失函数如下:

$$\mathcal{L}_\infty \leq \frac{1}{N} \sum_{i=1}^N -\log\left(\frac{e^{w_i^T a_i + b_i}}{\sum_{j=1}^C e^{w_j^T a_i + b_j + \frac{1}{2} v_j^T \Sigma_{y_j} v_j}}\right) \triangleq \bar{\mathcal{L}}_\infty \quad (4)$$

1.3 注意力机制

如图2所示,虽然农作物的病虫害特征有不同之处,但某些病害特征却极为相似,极易造成误判。为了更好地区分农作物病虫害,该文在网络中加入混合注意力机制获取图像的关键信息。Woo等^[14]提出卷积块注意力模块(Convolutional Block Attention Module, CBAM),混合通道维度和空间维度计算注意力权重,然后调整输入特征图,保留加强图像的关键特征。CBAM-Res2Net就是在Res2Net网络残差块中加入CBAM模块。



(a) 苹果黑星病晚期

(b) 辣椒疮痂病早期

图2 病虫害图片

1.3.1 通道注意力模块

通道注意力模块沿着通道维度计算注意力系数,

方法见图3。首先,输入特征分别经过全局最大池化、平均池化获得两个一维特征,然后将两个一维特征分别输入两层共享参数的神经网络。将输出的两个特征相加,通过Sigmoid函数激活得到一维通道注意力 $M_c(F)$ 。最后,将输入特征与 $M_c(F)$ 相乘,生成调整后的新特征。

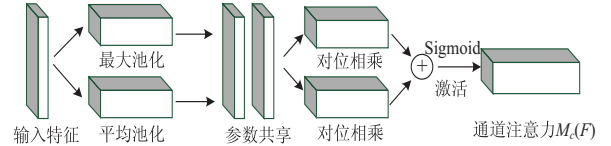


图3 通道注意力模块

通道注意力计算公式为:

$$M_c(F) = \sigma(\text{MLP}(\text{AvgPool}(F)) + \text{MLP}(\text{MaxPool}(F))) = \sigma(W_1(W_0(F_{\text{avg}}^c)) + W_1(W_0(F_{\text{max}}^c))) \quad (5)$$

式中, σ 为Sigmoid激活函数,MLP为参数共享的神经网络, W_0 和 W_1 分别对应第一层和第二层的网络权重,输入特征 F 分别经全局平均池化和最大池化得到新特征 F_{avg}^c 和 F_{max}^c 。

1.3.2 空间注意力模块

空间注意力模块沿着空间维度计算注意力系数,方法见图4。输入特征进行全局最大池化和平均池化得到两个二维特征后再进行拼接操作,然后使用 7×7 卷积降维后再经Sigmoid函数激活得到二维空间注意力 $M_s(F)$ 。最后,将输入特征与 $M_s(F)$ 相乘,得到新特征。

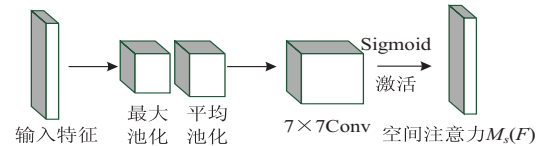


图4 空间注意力模块

空间注意力计算公式为:

$$M_s(F) = \sigma(f^{7 \times 7}([\text{AvgPool}(F); \text{MaxPool}(F)])) = \sigma(f^{7 \times 7}([F_{\text{avg}}^s; F_{\text{max}}^s])) \quad (6)$$

式(6), σ 为Sigmoid激活函数, $f^{7 \times 7}$ 表示 7×7 卷积, F_{avg}^s 和 F_{max}^s 对应于输入特征 F 经过全局平均池化和最大池化后得到的新特征。

1.4 农作物病虫害检测改进模型

提出的改进模型是保持Res2Net50网络主干不变,采用模型迁移来加载参数,并对部分结构微调训练。改进后的农作物病虫害检测模型如图5所示。

算法实现过程如下:

(1)在训练前,对数据集进行在线预处理,增加样本数量。

(2)网络加载经过ImageNet数据集训练的

Res2Net50 预训练模型,迁移底层网络参数。

(3) 训练过程中使用 ISDA 算法进行语义增强,优化网络输出。

(4) 选取动量梯度下降 (Momentum) 算法,反向传

播,更新网络参数。

(5) 通过全连接层 Softmax 函数预测分类。

(6) 多次试验,对比效果,选取最优超参数。

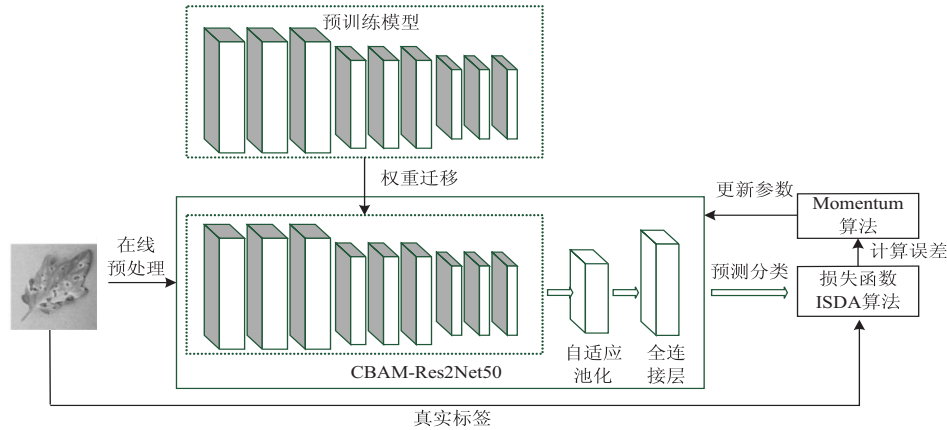


图 5 农作物病虫害检测改进模型

2 实验设计

2.1 数据集准备

实验采用 AI Challenger 2018 农作物病虫害数据集。该数据集包含苹果、玉米、樱桃、马铃薯、草莓、番茄、葡萄、柑橘、桃树和辣椒 10 种农作物、27 种病害类别 (分一般或严重) 和 10 种健康类别。数据集共有 47 393 张图片,其中训练集 31 718 张,验证集 4 540 张,测试集 (未标注) 9 027 张。由于数据集中第 44、45 类样本较少,故删除。该实验使用的数据集按照“物种-疾病-发病程度”分为 59 类,其中训练集有 31 716 个样本,测试集有 4 539 个样本。

2.2 数据预处理

训练神经网络需要海量且全面的数据来保证其稳定性。对数据集进行预处理不仅能增加样本量,还能提高模型的泛化能力。通过试验对比将图片尺寸重定义为 224×224 ,此时模型识别效果最好且节省计算资源。实验采用在线预处理的方式,首先对训练集图片尺寸重新定义,然后再进行随机亮度调整、对比度增强、 30° 随机旋转、上下翻转、水平翻转、 45° 仿射变换、像素去均值化与归一化处理等操作。

测试集图片只做尺寸重定义为 224×224 、像素去均值化与归一化处理操作。

2.3 超参数设计

合适的学习率可提高收敛速度。经过试验对比,将初始学习率设置为 0.001。在训练过程中,学习率根据验证集 loss 进行动态调整以便更好的收敛。当 loss 连续两次不再下降时,学习率便缩小 10 倍。

单次训练样本量 (Batch-Size) 会对模型性能造成一定影响。考虑到显卡容量,选取 16、32、64 进行试验

后确定为 32。

通过试验对比不同的梯度优化算法,综合考虑收敛速度和不容易陷入局部极值等因素,实验最终采用动量梯度下降算法 (Gradient Descent with Momentum, Momentum)。实验将 momentum 设为 0.9, L2 正则化系数设为 0.000 5。

损失函数使用 ISDA 算法,其中超参数 λ 是控制语义数据增强的强度系数。通过试验对比效果后 λ 确定为 1.0。

2.4 实验环境

实验环境: Windows10 OS, 32 GB 内存, AMD Ryzen7 5800X 处理器和 NVIDIA 3070 显卡。实验编程语言: PyTorch1.71 框架, Python3.6。

3 实验及结果分析

3.1 对比实验及分析

相同实验条件下,将改进模型和其他 4 种模型进行对比试验。考虑到实验数据集的病虫害种类较多,实验以平均准确率 (Average accuracy rate) 作为评价指标。为保证试验合理性,所有实验模型均采用模型迁移的方法,训练 60 轮。在测试集上的实验结果如图 6 所示。

结果分析:改进后的模型可达到更高的准确率和更低的损失值。改进模型收敛后,其准确率达到 88.3% 左右,损失值为 0.325。改进模型收敛速度最快,大约 10 次训练后趋于收敛。ResNet50 模型收敛速度稍慢于改进模型。ResNet101 和 InceptionV4 模型在 15 轮训练后趋于收敛。DenseNet121 模型的收敛速度较慢,且识别准确率低于其他模型。由于改进模型加了注意力机制,初始运行时性能表现低于其他模型,但

经过两次迭代后,性能表现优于其他模型。

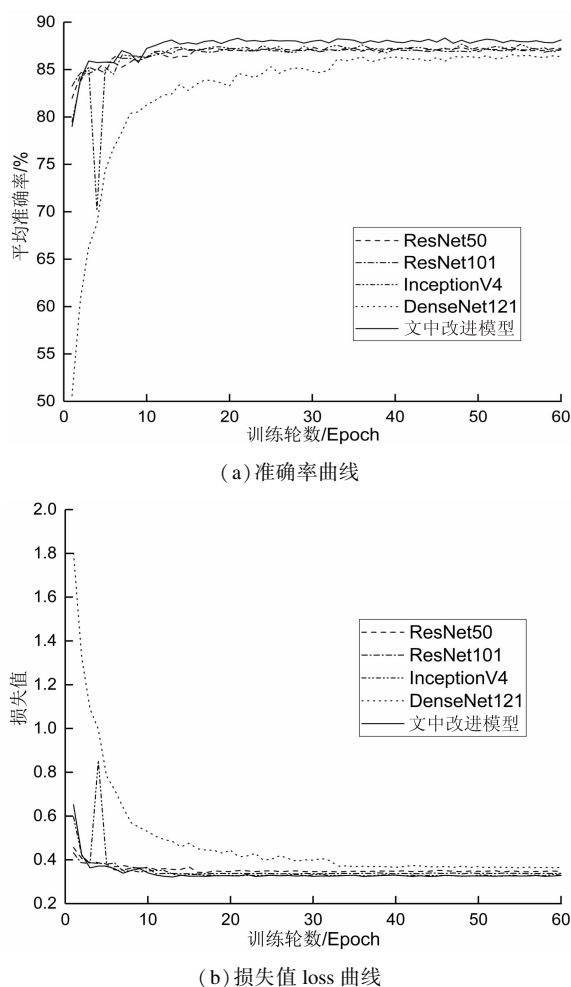


图 6 对比模型的准确率和损失值

表 1 为对比模型在测试集的平均准确率及其参数量。改进模型将 CBAM 模块添加到 Res2Net50 网络残差块中,使用 ISDA 算法进行训练。改进模型比 DenseNet121 和 ResNet50 模型参数略多,但比其他模型少,且识别准确率最高。改进模型在测试集上的识别准确率达 88.33%,分别比其他 4 种模型提高了 0.95%、1.04%、0.63% 和 1.75%。

表 1 对比模型的表现

模型	模型参数量(10^6)	平均准确率/%
ResNet50	23.6	87.38
ResNet101	42.6	87.29
InceptionV4	41.2	87.70
DenseNet121	7.0	86.58
改进模型	26.2	88.33

3.2 消融实验及分析

Res2Net50 为初始模型;模型 1 表示初始模型仅用 ISDA 算法进行训练;模型 2 仅在初始模型残差块中加入 CBAM 模块,即 CBAM-Res2Net50 模型。所有实验模型均采用模型迁移的方法,训练 60 轮。在测试集上

的实验结果如图 7 所示。

在消融实验中,模型 2 最先达到收敛状态,在 10 次迭代左右就趋于收敛。改进模型和模型 1 经过 10 次迭代后趋于收敛,而初始模型收敛速度最慢。模型 1 和模型 2 在测试集上的准确率均高于初始模型。通过对比模型的准确率和损失收敛曲线,改进模型比初始模型先收敛,识别准确率提高。

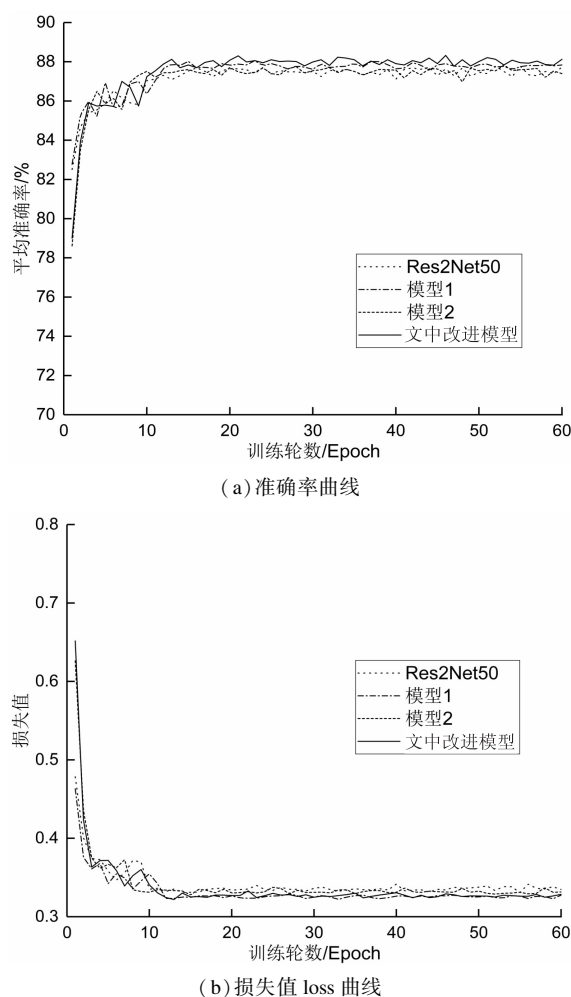


图 7 不同模型的准确率和损失值

如表 2 所示,改进模型收敛后,识别准确率可达 88.33%。与 Res2Net50、模型 1 和模型 2 相比,改进模型的识别准确率分别提高了 0.56%、0.31% 和 0.49%。

表 2 消融实验

模型	模型参数量(10^6)	平均准确率/%
Res2Net50	23.7	87.77
模型 1	23.7	88.02
模型 2	26.2	87.84
改进模型	26.2	88.33

3.3 不同文献对比分析

由于文献[15-16]均采用 AI Challenger 2018 农作物病虫害数据集进行实验,选择其最优模型进行对比,

结果如表 3 所示。

表 3 不同文献模型对比

模型	识别准确率/%
文献[15]	87.19
文献[16]	86.76
改进模型	88.33

文献[15]采用模型融合的方法,即 RestNext50、SE-ResNet50、SE-ResNext50、ResNet101 等 4 种网络模型融合成新模型。虽然模型效果相较于单一模型有所提升,但增加了模型训练时间和参数量。文献[16]对深度残差网络进行微调,使用 Focal Loss 进行训练,分类准确率较差。改进模型的识别准确率比文献[15]和文献[16]分别提高了 1.14%、1.57%。

4 结束语

该文提出了基于隐式语义数据增强算法的 CBAM-ResNet50 模型,保持 ResNet50 网络结构不变,使用模型迁移加载部分参数,并在神经网络残差块中添加 CBAM 模块。研究采用 AI Challenger 2018 农作物病虫害数据集进行对比实验。实验结果表明,改进模型对 59 类病虫害的分类准确率为 88.33%。与 ResNet50、ResNet101、InceptionV4 和 DenseNet121 模型相比,改进模型的识别准确率分别提高了 0.95%、1.04%、0.63% 和 1.75%。改进模型从多尺度学习图像特征,挖掘图像的语义信息,在一定程度上可解决样本不足的问题,从而提高网络模型的泛化能力。

目前,实验所用图片背景简单,没有考虑复杂环境的干扰。后续研究会将不断收集农作物图片,丰富农作物病虫害数据集,对神经网络进行剪枝以减少参数数量,建立移动终端的农作物分类病虫害检测系统等。

参考文献:

- [1] 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.
- [2] FERENTINOS K P. Deep learning models for plant disease detection and diagnosis[J]. Computers and Electronics in Agriculture, 2018, 145: 311-318.
- [3] SLADOJEVIC S, ARSENOVIC M, ANDERLA A, et al.

Deep neural networks based recognition of plant diseases by leaf image classification[J]. Computational Intelligence and Neur-oscience, 2016, 2016: 3289801.

- [4] LIU B, ZHANG Y, HE D J, et al. Identification of apple leaf diseases based on deep convolutional neural networks[J]. Symmetry-Basel, 2018, 10(1): 1-16.
- [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, 2019, 161: 272-279.
- [6] 王东方,汪军.基于迁移学习和残差网络的农作物病害分类[J].农业工程学报,2021,37(4):199-207.
- [7] 刘阳,高国琴.采用改进的 SqueezeNet 模型识别多类叶片病害[J].农业工程学报,2021,37(2):187-195.
- [8] 樊湘鹏,许燕,周建平,等.基于迁移学习和改进 CNN 的葡萄叶部病害检测系统[J].农业工程学报,2021,37(6):151-159.
- [9] 贾俊杰,李捷.基于半监督生成对抗网络的病虫害图像识别算法[J].武汉轻工大学学报,2019,38(4):45-52.
- [10] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]//IEEE conference on computer vision and pattern recognition (CVPR). Las Vegas: IEEE, 2016: 770-778.
- [11] GAO S H, CHENG M M, ZHAO K, et al. Res2Net: a new multi-scale backbone architecture[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021, 43(2): 652-662.
- [12] WANG Y, HUANG G, SONG S, et al. Regularizing deep networks with semantic data augmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 44(7): 3733-3748.
- [13] BENGIO Y, MESNIL G, DAUPHIN Y, et al. Better mixing via deep representations[C]//International conference on machine learning (ICML). Atlanta: ACM, 2013: 552-560.
- [14] WOO S, PARK J, LEE J Y, et al. CBAM: convolutional block attention module[C]//Proceedings of the European conference on computer vision (ECCV). Munich: Springer, 2018: 3-19.
- [15] 龚安,井晓萌.多卷积神经网络模型融合的农作物病害图像识别[J].计算机技术与发展,2020,30(8):134-139.
- [16] 史冰莹,李佳琦,张磊,等.基于 CNN 的农作物病虫害图像识别模型[J].计算机系统应用,2020,29(6):89-96.