

基于 Res2Net 和双线性注意力的番茄病害时期识别方法

贾兆红 张袁源 王海涛 梁 栋
(安徽大学互联网学院, 合肥 230039)

摘要: 针对番茄叶片型病害在早晚期具有类内差异大、类间差异小的特点,常规神经网络对此类病害的分类效果不佳的问题,提出了基于 Res2Net 和双线性注意力的番茄病害时期识别方法,通过多尺度特征和注意力机制,提高网络的细粒度表征能力。首先,提出 EFCA 通道注意力模块,在不降维的基础上,使用二维离散余弦变换代替全局平均池化,以减少常规通道注意力获取时的信息丢失。其次,在外积之后加入最大池化和 concat 操作,避免双线性融合后因维度过高导致的特征冗余。在 7 种不同种类和 14 种不同程度病害番茄叶面型病害数据集实验中,本文方法分类准确度分别为 98.66% 和 86.89%。

关键词: 神经网络; 细粒度分类; 番茄病害时期; 注意力机制; 双线性卷积神经网络

中图分类号: TP391.4 **文献标识码:** A **文章编号:** 1000-1298(2022)07-0259-08 **OSID:** 

Identification Method of Tomato Disease Period Based on Res2Net and Bilinear Attention Mechanism

JIA Zhaohong ZHANG Yuanyuan WANG Haitao LIANG Dong
(School of Internet, Anhui University, Hefei 230039, China)

Abstract: Tomato leaf-type diseases have the characteristics of large intra-class differences and small inter-class differences in the early and late stages. The conventional neural network is not effective in classifying such diseases. Therefore, based on the fine-grained weakly supervised classification method, a Res2Net bilinear attention network, combining the bilinear model and attention mechanism, was proposed. The fine-grained representation ability was improved through extracting multi-scale features and combining the attention mechanism. First of all, for the problem of information loss in the process of conventional channel attention acquisition, the EFCA channel attention module was proposed. On the basis of no dimensionality reduction, two-dimensional discrete cosine transform was used instead of global average pooling to avoid some features from being lost in downsampling. Secondly, by adding the maximum pooling after the outer product, and the concat module designed by drawing on the shortcut idea in the residual network, the problem of redundant features caused by the excessively high dimensionality after bilinear fusion was solved. The obtained classification accuracies of the proposed model on the data set with 7 types and 14 different degrees of tomato leaf type diseases were 98.66% and 86.89%, respectively.

Key words: neural network; fine grained classification; tomato disease period; attention mechanism; bilinear convolutional neural network

0 引言

番茄病害是影响番茄产量的重要因素之一,如果可以提早发现并防治,对于提升作物品质,增加经济效益有着不可或缺的重要作用^[1]。而倾向于细

粒度识别(Fine-grained recognition)^[2]的叶片病害早晚期识别,可以在病害发生初期就得以发现并解决,减少经济损失,但早期病害的特征不明显且占据叶片的比例较小,部分病害早晚期区分差异小等都给识别判断带来了挑战。

收稿日期: 2021-07-16 修回日期: 2021-08-25

基金项目: 国家自然科学基金项目(71971002)、安徽省重大专项(202003a06020016)和安徽省教育厅科学研究项目(YJS20210029)

作者简介: 贾兆红(1976—),女,教授,博士,主要从事人工智能、决策支持、多目标优化算法及其应用研究,E-mail: Jiazhaozhong001@163.com

通信作者: 王海涛(1978—),男,教授,博士,主要从事软件工程、智能决策及数据工程研究,E-mail: htwang@ahu.edu.cn

随着人工智能技术的兴起,相比于传统人工识别植物病害的低效和费时,通过机器学习和深度学习自动提取特征并进行分类的效果更好^[3]。柴阿丽等^[4]通过图形处理手段人工提取了番茄早疫病、晚疫病、叶霉病和棒孢叶斑病等4种叶部病害的病变区域的颜色、纹理和形状特征参数,并构建贝叶斯判别模型对病害识别分类。因为人工提取特征常受到主观因素和先验知识的限制,因此利用神经网络自动提取病害特征进行迭代学习,在农作物病害分类中取得了更好的分类效果^[5-9]。袁培森等^[10]以 Inception ResNet V2 作为基础模型,将双线性卷积神经网络用于细粒度菌菇表型数据集上,并取得了较好的分类识别率。除双线性模型^[11]外,注意力机制也常常用于细粒度图像分类。胡志伟等^[12]提出一种注意力与残差思想结合的卷积神经网络模型,通过残差网络提取叶片特征,注意力机制分配网络权重,过滤掉噪声信息并提升模型效果。ZHENG 等^[13]通过多注意力实现的信道分组将图像分为 N 个部分,并使用两个损失函数相互加强,不断迭代达到了分组和分类相互强化的细粒度分类效果。LIANG 等^[14]将 ResNet 作为基础网络,加入 Shuffle 单元,在实现植物病害的3个不同功能分类的同时,降低了模型的计算复杂度。

上述研究都没有在农作物数据集中尝试结合细粒度识别中的不同方法,相互促进学习,弥补单一网络的不足。本文提出采用 Res2Net 作为基础网络,提升多尺度特征提取能力,以使网络获得更丰富特征信息,增强模型的表征能力。同时提出一种新的通道注意力模块,解决注意力权重获取过程中信息丢失和模型复杂度问题。改进双线性网络的特征融合,避免外积后出现大量特征冗余,且将空间和通道注意力模块分别引入双线性网络,解决串联注意力模块造成的准确率不升反降的问题。

1 番茄病害识别模型

1.1 理论方法

1.1.1 Res2Net 核心理念

番茄叶片上的病害程度不同且类型多样,采用基础的 ResNet^[15]提取特征往往难以充分表达叶片不同部位的语义信息,因此需要考虑多尺度特征以解决细粒度叶片分类中差异性局部关键信息的提取问题。由 GAO 等^[16]提出的一种新的多尺度主干网络 (Res2Net),利用特征分组的思想,在残差块内以卷积组代替原来单一卷积,并以层级残差方式连接,从而达到增加网络的感受野区域,多个特征尺度产生组合效应,提升网络的多尺度特征提取能力。

Res2Net 的核心网络结构如图 1 所示,主要修改了 Bottleneck block 中间的 3×3 卷积层结构,引入超参数 s ,将经过 1×1 卷积层输出的特征图 $F(h \times w \times d)$ 按通道划分为 s 组,即每一组特征 X_i 的形状为 $h \times w \times d/s, i \in \{1, 2, \dots, s\}$ 。在保持空间信息不变的同时,记经过 3×3 卷积层输出的操作为 $K_i()$,第 1 组特征 X_1 不经过卷积操作直接输出 $Y_1 = X_1$, X_2 经过 3×3 卷积层输出 $Y_2 = K_2(X_2)$, X_3 与 Y_2 作残差连接通过卷积层后输出 $Y_3 = K_3(X_3 + Y_2)$,以此类推,最后得到输出 Y_s 。再将这些特征组进行通道拼接后输入 1×1 卷积层,输出即可,具体输入公式为

$$Y_i = \begin{cases} X_i & (i = 1) \\ K_i(X_i + Y_{i-1}) & (2 \leq i \leq s) \end{cases} \quad (1)$$

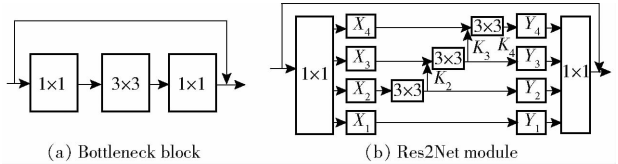


图 1 Res2Net 网络结构
Fig. 1 Structure of Res2Net network

1.1.2 Bilinear CNN 模型

细粒度图像分类中除使用类别标签,还通过额外的 bounding box 和 part locations 等人工标注信息,以更加精确地获取目标区域的位置和大小,提升局部与全局区域的关联性,这类算法称为强监督学习^[17]。只用类别标签而不用其它标签信息的算法称为弱监督学习^[18]。虽然强监督学习已取得不错的分类效果,但人工标注代价高且算法通用性不强。因此本文采取弱监督图像分类算法中的端到端模型,即由 LIN 等^[11]提出的双线性卷积神经网络。

一个双线性模型主要由一个四元组函数 $T = (f_x, f_y, p, c)$ 构成,其中 f 是指去除 softmax 层和全连接层的卷积网络特征提取函数, x 与 y 相同时称为同源双线性模型, x 与 y 不同时则称为多模双线性模型, p 代表池化函数, c 代表由全连接层构成的分类器。对于 f_x 提取的特征向量 $A \in \mathbf{R}^{l \times m}$, f_y 提取的特征向量 $B \in \mathbf{R}^{l' \times n}$, 通过双线性池化进行外积相乘得到 $M \in \mathbf{R}^{m \times n}$, 这样得到的融合特征计算方便且便于梯度反向传播。

1.2 模型结构

由于原 Bilinear CNN 模型采取 VGG 作为特征提取子网络,而 VGG 网络多尺度特征提取能力较弱,无法完成番茄病害叶片上的差异性局部关键信息的捕捉,不利于提高网络的表征能力,所以本文采用 Res2Net 作为特征提取子网络,获得灵活的感受

野,以识别多尺度范围的叶片病害。

针对同一种番茄病害在早晚期的表现特征差异大,且拍摄角度、光照阴影等环境因素影响较大,使用传统卷积神经网络进行识别准确率不高,本文提出一种基于 Res2Net 的双线性注意力网络模型用于番茄病害叶片早晚期识别。模型主要包括两部分:结合高效频域和瓶颈结构设计的通道注意力模块(EFCA)和空间注意力模块(SA)。分别将这两个注意力模块添加到 Res2Net50 和 ResNet34 的残差结

构中,然后使用双线性融合两个子网络提取的特征。由于双线性外积得到的特征维度较高,存在特征冗余,严重影响分类精度和模型效率,所以在融合后先使用一维卷积捕获通道间的交互关系,然后追加最大池化层降低特征维度,减少冗余信息。同时借鉴残差网络的 shortcut 思想,再与外积前的特征张量做 concat 操作,使得融合后模型的表征能力较融合前至少没有降低,最后使用全连接层达到分类的目的。模型整体结构如图 2 所示。

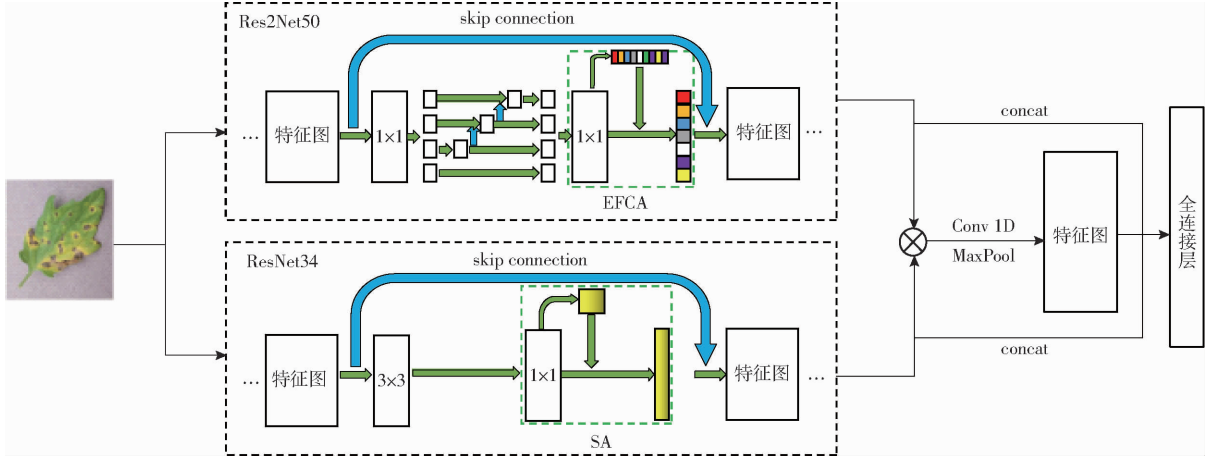


图 2 双线性注意力网络模型结构

Fig. 2 Structure of bilinear attention network model

1.3 注意力机制

卷积神经网络通过卷积滤波器捕捉局部感受域的空间和通道融合信息,从而获得全局感受域的分层模式来作为图像特征的代表。但是不同的空间和通道信息对应图像中的不同信息,有些背景干扰信息会影响分类性能,所以采用空间^[19]和通道^[20]注意力机制为特征层添加不同权重,以降低特征冗余产生的影响。

(1)通道注意力

EFCA 模块是结合模块 ECANet block^[21] 和 FcaNet block^[22] 各自优势的一种通道注意力模块,避免了权重获取过程中部分特征在降采样时仅保留不充分信息及降维给通道注意力预测带来的副作用,而且在不增加模型复杂度和计算量的同时性能得到明显提升,模块结构如图 3 所示。

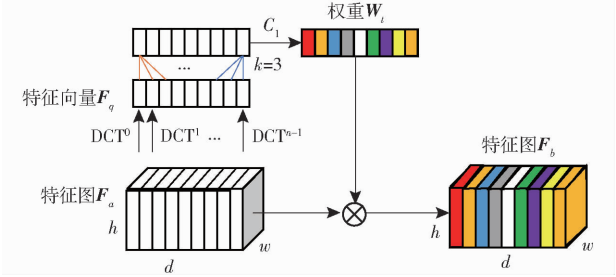


图 3 EFCA 模块结构

Fig. 3 Structure of EFCA module

获取通道注意力的过程描述如下:①特征提取函数某一卷积层捕捉得到的特征图 $F_a \in \mathbf{R}^{h \times w \times d}$ 作为输入,其中 h, w 代表特征图的空间维度, d 代表通道维度的数量。为了避免全局平均池化降低特征维度,增大 kernel 感受野范围时导致信息丢失,对 F_a 使用 DCT (Discrete cosine transform) 进行降采样处理,得到由多个频域组合提取得到的特征向量 $F_q \in \mathbf{R}^{1 \times 1 \times d}$ 。②将 F_q 通过内核大小为 k 的一维卷积,使得在提取非线性特征时,实现无需降维的局部跨通道交互,再通过 Sigmoid 激活函数生成通道权重 $W_i \in \mathbf{R}^{1 \times 1 \times d}$ 。③将权重 W_i 和输入的特征图 F_a 采用对应元素相乘的策略进行加权融合,得到带通道注意力的融合特征图 F_b 。相较于 F_a ,使用 F_b 可以使模型更关注于显著区分特征,而忽视冗余的背景信息,增强模型的有效特征提取能力。

步骤①~③的通道注意力获取过程可表示为

$$F_b = \delta(C_1[D_{ct}(F_a)]) \otimes F_a \tag{2}$$

式中 δ ——sigmoid 激活函数
 C_1 ——卷积核大小为 k 的一维卷积操作
 D_{ct} ——2 维离散余弦变换
 $\otimes$ ——对应元素乘积操作

(2)空间注意力

对于二维图像来说,仅关注图像的通道维度是不够的,还需要基于图像的尺度空间即长宽的注意

力机制。本文基于 CBAM^[23] 算法,结合使用空间注意力寻找图像中与任务相关的区域,使得网络具有一定的尺度不变性,模块结构如图 4 所示。

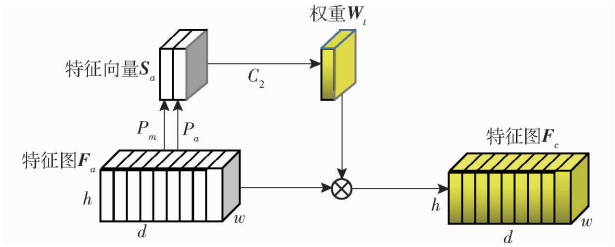


图 4 空间注意力模块结构图

Fig.4 Structure diagram of spatial attention module

获取空间注意力的过程描述如下:①将特征提取函数得到的特征图 $F_a \in \mathbf{R}^{h \times w \times d}$ 作为输入,通过最大池化和平均池化在通道维度上进行降采样,得到两个特征向量 $S_a \in \mathbf{R}^{h \times w \times 2}$ 。②使用 concat 函数对两个特征向量 S_a 级联后,采用大小为 7 的二维卷积进行空间维度的特征提取,并通过 sigmoid 函数激活得到对应的空间权重 $W_t \in \mathbf{R}^{h \times w \times 1}$ 。③对 W_t 与 F_a 进行对应元素相乘,得到带权重信息的空间注意力特征图 F_c 。

步骤①~③的通道注意力获取过程可表示为

$$F_c = \delta \{ C_2 [P_{ma}(F_a)] \} \otimes F_a \tag{3}$$

式中 C_2 ——卷积核大小为 7 的二维卷积操作

P_{ma} ——通道维度上的最大池化和平均池化

2 实验与结果分析

2.1 数据集

所提出的网络模型可用于不同农作物病害的分类,本文使用番茄病害作为实验对象。采用的数据集来自 AI Challenger 2018 比赛中的农作物病害数据集,单独将其中的番茄病害分离出来,并筛选掉其中个别图像数目极不平衡的类别。最终得到番茄的 14 个病害类别数据集,共 11 359 幅图像,分别是白粉病早晚期、早疫病早晚期、晚疫病早晚期、叶霉病早晚期、斑枯病早晚期、红蜘蛛损伤早晚期、黄化曲叶病毒病早晚期,具体类别与对应数目如表 1 所示。

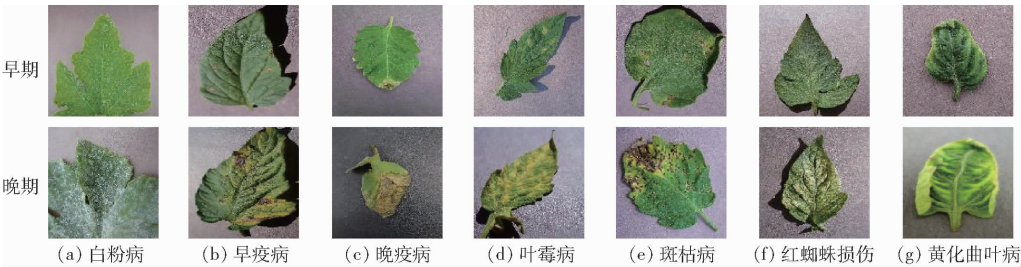


图 5 不同病害类别早晚期对比图像

Fig.5 Contrast images of different disease categories in early and late stages

表 1 番茄数据集病害种类与数量分布		
Tab.1 Disease types and quantity distribution of tomato dataset		
序号	病害类别	数量/幅
0	番茄白粉病早期	365
1	番茄白粉病晚期	1 104
2	番茄早疫病早期	287
3	番茄早疫病晚期	505
4	番茄晚疫病早期	302
5	番茄晚疫病晚期	1 267
6	番茄叶霉病早期	371
7	番茄叶霉病晚期	384
8	番茄斑枯病早期	481
9	番茄斑枯病晚期	922
10	番茄红蜘蛛损伤早期	619
11	番茄红蜘蛛损伤晚期	310
12	番茄黄化曲叶病早期	1 616
13	番茄黄化曲叶病晚期	2 826

病害不同时期的表现如图 5 所示,部分病害主要区别为:①晚疫病早期从叶片边缘、叶尖处开始,为暗绿色水渍状半圆形病斑;晚期病斑扩大后中间黑褐色、周围暗绿色。②红蜘蛛损伤早期叶片呈现密集苍白的小斑点,卷曲发黄;晚期病害面积扩散,叶片变黄,出现卷叶、焦叶。可以发现叶片病害早晚期的表现不同,病状差异明显。符合细粒度图像识别的“类内差异大,类间差异小”的基本特征^[24],这也给使用常规卷积神经网络进行叶片分类带来了困难。

2.2 数据预处理

由于深度学习依赖于海量数据集的训练,所以使用数据增强技术扩展数据集是取得良好分类效果且避免模型训练过拟合的重要途径。本实验中采取在训练过程中获得 batch 后,对 batch 数据进行在线增强,对数据进行随机旋转角度、水平平移和上下平移、错切变换、随机缩放比例、水平翻转等预处理方式,增强后的部分图像与原始图像对比见图 6。

2.3 实验设计

本文在深度学习框架 tensorflow2 基础上,使用 NVIDIA GeForce RTX 2080Ti GPU 加速网络训练。

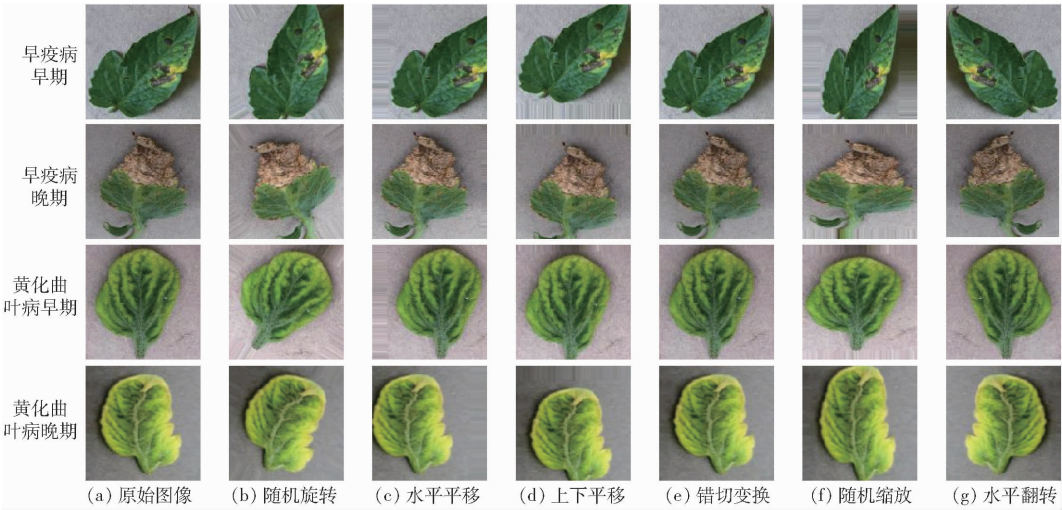


图 6 部分病害图像及增强图像示例

Fig. 6 Examples of partial disease image and enhanced image

为减少模型计算量,将图像缩放至 224 像素 × 224 像素,另外考虑降低图像几何变化带来的影响,需要对图像进行归一化处理,并将图像按 8: 2 的比例划分为训练集、验证集,保证模型的泛化性和鲁棒性。

由于数据集中各类别数量较不平衡,因此采用 focal loss^[25] 取代传统的交叉熵损失函数,解决难易样本和正负样本不平衡给模型训练带来的过拟合问题。针对数据集中异常的样本标注,使用 label smoothing^[26] 缓解错误标签给模型训练带来的影响。

2.4 参数设置

使用自适应矩估计 (Adam) 优化器优化网络并在一定程度上解决数据噪声较高的问题,初始学习率为 0.001,使用 keras 中的回调函数 ReduceLROnPlateau 动态调整学习率,当 val_loss 连续 3 次不下降时,调整学习率为 0.2。最大训练轮数为 50,批量大小为 16,使用 EarlyStopping 函数监测 val_loss 指标,当连续 5 次下降低于 0.000 1 时停止训练。Res2Net 中的超参数 width 设为 16, scale 为 4,FECA 模块中一维卷积核大小 k 设为 3,用于捕获模型的跨通道交互,双线性融合后的最大池化层 pool_size 为 16,达到降低特征维度,避免特征冗余

对分类效果的影响。

2.5 消融与对比实验

为了验证本文提出的模型在番茄叶片病害识别上的有效性,首先对模型进行消融实验,分别使用 Res2Net50、+ 通道注意力 (EFCA)、+ 空间注意力 (SA)、+ Bilinear CNN 在数据集上从头训练,采用准确率作为评价指标,结果如表 2 所示,训练过程中验证集损失和准确率曲线如图 7 所示。

表 2 不同模块消融实验结果

Tab. 2 Ablation results of different modules

基础网络	分类方法	准确率/%
Res2Net50	backbone	84.78
+ EFCA	通道注意力	85.62
+ SA	通道、空间注意力	84.71
+ Bilinear CNN	通道、空间注意力、双线性池化	86.89

从表 2 可以看出,使用 EFCA 通道注意力可以在基础模型的基础上将准确率提升 0.84 个百分点。将空间和通道注意力串联使用后,准确率反而有所下降,可能原因是通过通道注意力后,特征图缺失了部分空间特征,因此再使用空间注意力后效果不升反降。因此本文提出的模型将通道和空间注意力并行实现在 Res2Net50 和 ResNet34 中,再通过双线性

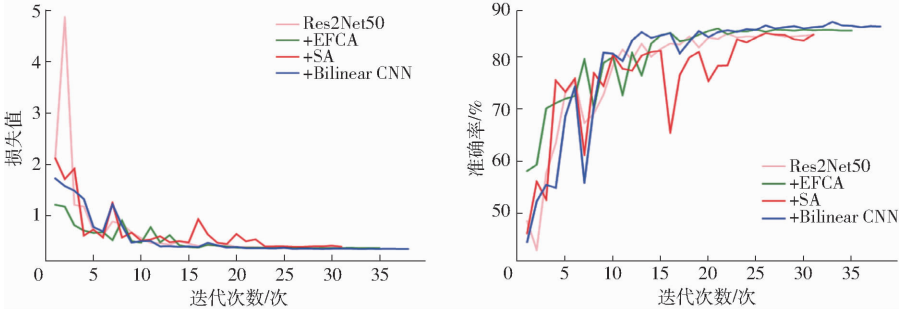


图 7 消融实验验证集损失值与准确率曲线

Fig. 7 Loss and accuracy curves of verification set in ablation experiment

融合以便更好地提取病害特征,最终模型准确率可以达到 86.89%,较基础模型提升 2.11 个百分点。观察图 7 发现,串联空间和通道注意力后导致模型的损失值和准确率曲线震荡严重,表现较差。而本文使用的双线性注意力模型触发 EarlyStopping 最晚,且曲线后期表现平稳,说明该模型具有较高的稳定性。

为了验证本文方法的优越性,首先将 backbone 模型与 ShuffleNetV2^[27]、InceptionV4^[28]、DenseNet121^[29] 做对比实验;同时为了实验的公平性,与同样使用注意力机制的 CBAM^[23]、ARNet^[12] 模型进行对比;最后,与原文中提出的双线性模型 Bilinear CNN^[11] 和自定义的 Bilinear EfficientNet 比

较,结果见表 3,验证集损失值和准确率曲线如图 8 所示。

表 3 不同分类方法对比实验结果
Tab.3 Comparative experiment results of different methods

基础网络	分类方法	准确率/%
Res2Net50	backbone	84.78
ShuffleNetV2		81.04
InceptionV4		84.21
DenseNet121		82.31
CBAM	空间、通道注意力	83.44
ARNet	残差注意力	83.72
Bilinear CNN	双线性融合	84.78
Bilinear EfficientNet	双线性融合	85.74

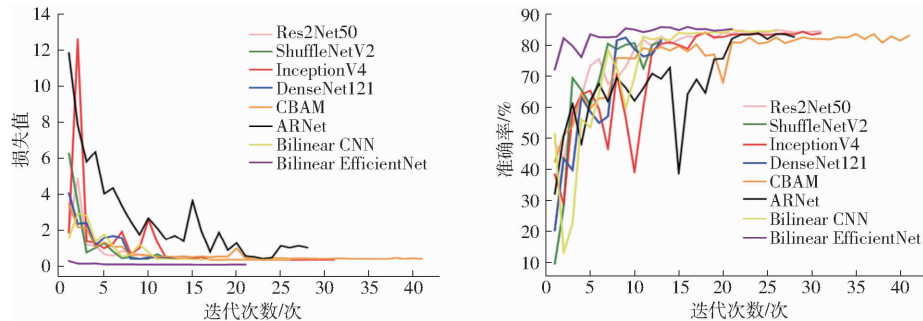


图 8 对比实验验证集损失值与准确率曲线

Fig.8 Loss and accuracy curves of verification set in comparative experiment

从表 3 可看出,与 ShuffleNetV2、InceptionV4 和 DenseNet121 相比,Res2Net50 取得了最佳结果,准确率分别提升 3.74、0.57、2.47 个百分点。对比同样使用空间注意力和通道注意力的 CBAM 模型,以及针对番茄病害分类所提出的残差注意力 ARNet 模型,本文提出的双线性注意力模型准确率分别提升 3.45、3.17 个百分点,比 Bilinear CNN 和自定义的 Bilinear EfficientNet 准确率分别提升 2.11、1.15 个百分点。原双线性网络表现不好的原因是,双线性融合后大量的特征冗余影响了模型决策。在图 8 的 backbone 模型比较中可以发现,只有 InceptionV4 的最终表现接近 Res2Net,这是因为二者都利用多个不同大小的卷积核提取多尺度特征,这有利于模型捕捉番茄叶面病害的差异性局部关键特征,但 InceptionV4 的曲线波动较大,稳定性较差,这也是本文选取 Res2Net 作为基础模型的原因之一。自定义的 Bilinear EfficientNet 模型表现较稳定的主要原因是迁移学习了 ImageNet 的预训练参数,但同时触发 EarlyStopping 较早,说明迁移学习在细粒度病害分类问题上仅仅加速了模型的收敛速度,但对于模型的准确率提升并没有较大的帮助,因此本文并没有基于迁移学习进行模型训练。

本文模型在 14 种不同程度病害分类上的准确率

为 86.89%,如果不区分病害时期,仅关注病害类别,在 7 种不同种类病害数据集上,本文模型平均准确率为 98.66%,说明该模型具有良好的泛化能力。由图 9 计算可知每个类别的准确率,其中图 9a 中行列号 0~6 分别对应白粉病、早疫病、晚疫病、叶霉病、斑枯病、红蜘蛛损伤、黄化曲叶病这 7 种病害类别,图 9b 中的行列号 0~13 对应表 1 中的序号 0、9、10、11、12、13、1、2、3、4、5、6、7、8。可以看出区分病害时期的分类中,红蜘蛛损伤晚期和黄化曲叶病晚期的准确率只达到 74.36% 和 75.74%,较其它种类分类效果较差,主要原因是这两种病害早晚期界定较难,人工分类存在部分标签异常,影响模型分类效果。

2.6 注意力可视化分析

通过将卷积层的特征映射与原图在空间位置上进行对应,利用类激活可视化 (grad-cam)^[30] 方法,将模型最后一个卷积层在叶片上的注意力可视化,更清晰地说明注意力机制在模型判别时所关注的叶片病害特征。分别将 Res2Net50、空间注意力、通道注意力生成的热力图进行可视化,如图 10 所示。

热力图中红色区域代表特征图与预测类别的响应值高,这些区域越多代表模型决策依据越多,判别更准确。观察红蜘蛛损伤以及早疫病的早晚期在空间和通道注意力中的热力图可以发现,使用注意力

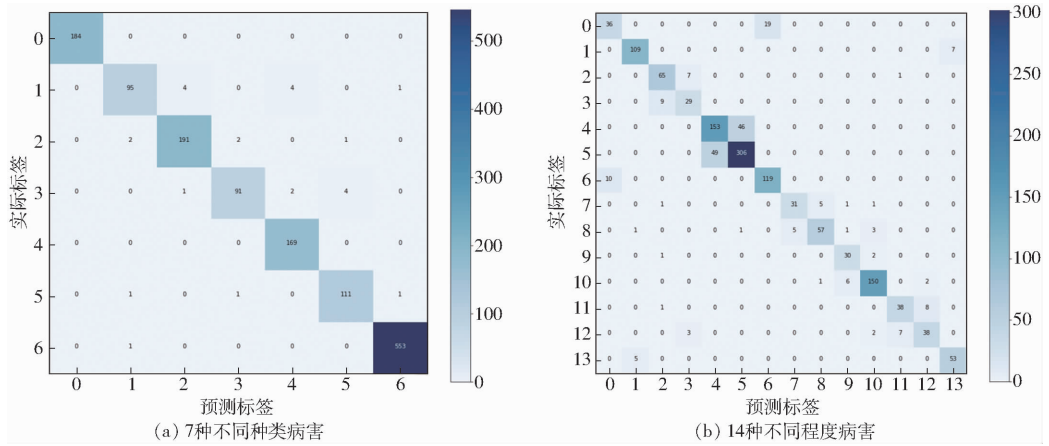


图 9 验证集混淆矩阵
Fig.9 Validation set confusion matrix

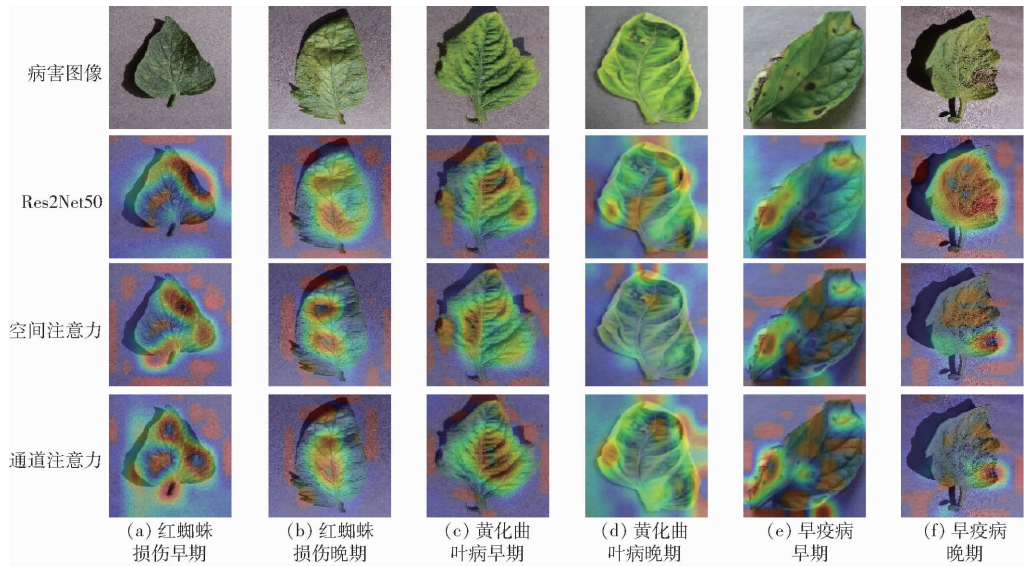


图 10 注意力特征图可视化对比
Fig.10 Visual comparison of attention characteristic map

后模型对于病害判别性区域,即叶片中出现的小斑点和大面积病斑识别更精准。对于黄化曲叶病,相较于早期时关注的叶肉变厚特征,注意力机制在判别晚期时更加关注叶片卷曲、皱缩的边缘区域,使得模型的决策依据更加丰富多样化,增强了模型的病害表征能力。

3 结束语

通过结合细粒度图像分类中的弱监督分类算

法,提出一种基于 Res2Net 和双线性注意力模型对番茄叶面型病害时期进行分类。首先,针对常规通道注意力获取过程中出现的信息丢失问题,在不降维的基础上,提出使用 DCT 代替 GAP 的 EFCA 通道注意力模块,避免部分特征在降采样中缺失。其次,针对双线性融合后维度过高导致的特征冗余,通过在外积后增加最大池化和基于 shortcut 思想设计的 concat 模块进行了有效解决。

参 考 文 献

[1] 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. WANG Xiangyu,WEN Haojie,LI Xinxing,et al. Research progress analysis of mainly agricultural diseases detection and early warning technologies [J]. Transactions of the Chinese Society for Agricultural Machinery,2016,47(9):266-277. (in Chinese)

[2] WANG Y, WANG Z. A survey of recent work on fine-grained image classification techniques [J]. Journal of Visual Communication and Image Representation,2019,59:210-214.

[3] 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)

- [4] 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.
CHAI Ali,LI Baoju,SHI Yanxia,et al. Recognition of tomato foliage disease based on computer vision technology[J]. Acta Horticulturae Sinica,2010,37(9):1423-1430. (in Chinese)
- [5] CRR C,PSA A,MEA A,et al. Identification and recognition of rice diseases and pests using convolutional neural networks[J]. Biosystems Engineering,2020,194:112-120.
- [6] LU Y,YI S,ZENG N,et al. Identification of rice diseases using deep convolutional neural networks[J]. Neurocomputing, 2017,267:378-384.
- [7] FERENTINOS K P. Deep learning models for plant disease detection and diagnosis[J]. Computers and Electronics in Agriculture,2018,145(1):311-318.
- [8] 黄林生,罗耀武,杨小冬,等. 基于注意力机制和多尺度残差网络的农作物病害识别[J]. 农业机械学报,2021,52(10):264-271.
HUANG Linsheng,LUO Yaowu,YANG Xiaodong,et al. Crop disease recognition based on attention mechanism and multi-scale residual network[J]. Transactions of the Chinese Society for Agricultural Machinery,2021,52(10):264-271. (in Chinese)
- [9] 王美华,吴振鑫,周祖光. 基于注意力改进 CBAM 的农作物病虫害细粒度识别研究[J]. 农业机械学报,2021,52(4):239-247.
WANG Meihua,WU Zhenxin,ZHOU Zuguang. Fine-grained identification research of crop pests and diseases based on improved CBAM via attention[J]. Transactions of the Chinese Society for Agricultural Machinery,2021,52(4):239-247. (in Chinese)
- [10] 袁培森,申成吉,徐焕良. 基于迁移学习和双线性 CNN 的细粒度菌菇表型识别[J]. 农业机械学报,2021,52(7):151-158.
YUAN Peisen,SHEN Chengji,XU Huanliang. Fine-grained mushroom phenotype recognition based on transfer learning and Bilinear CNN[J]. Transactions of the Chinese Society for Agricultural Machinery,2021,52(7):151-158. (in Chinese)
- [11] LIN T, ROYCHOWDHURY A, MAJI S. Bilinear CNN models for fine-grained visual recognition[C]//IEEE International Conference on Computer Vision,2015:1449-1457.
- [12] 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 Jimin,et al. Fine-grained tomato disease recognition based on attention residual mechanism [J]. Journal of South China Agricultural University,2019,40(6):124-132. (in Chinese)
- [13] ZHENG H, FU J, MEI T,et al. Learning multi-attention convolutional neural network for fine-grained image recognition[C]// 2017 IEEE International Conference on Computer Vision (ICCV), 2017:5219-5227.
- [14] LIANG Q, XIANG S, HU Y,et al. PD2SE-Net: computer-assisted plant disease diagnosis and severity estimation network [J]. Computers and Electronics in Agriculture,2019,157(3):518-529.
- [15] HE K,ZHANG X, REN S,et al. Deep residual learning for image recognition[C]// Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016:770-778.
- [16] GAO S,CHENG M M,ZHAO K,et al. Res2Net: a new multi-scale backbone architecture[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2019,43(2):652-662.
- [17] 李思瑶,刘宇红,张荣芬. 多尺度特征融合的细粒度图像分类[J]. 激光与光电子学进展,2020,57(12):91-97.
LI Siyao,LIU Yuhong,ZHANG Rongfen. Fine-grained image classification based on multi-scale feature fusion[J]. Laser & Optoelectronics Progress,2020,57(12):91-97. (in Chinese)
- [18] 朱阳光,刘瑞敏,黄琼桃. 基于深度神经网络的弱监督信息细粒度图像识别[J]. 电子测量与仪器学报,2020,34(2): 115-122.
ZHU Yangguang,LIU Ruimin,HUANG Qiongtao. Fine-grained image recognition of weak supervisory information based on deep neural network[J]. Journal of Electronic Measurement and Instrument,2020,34(2):115-122. (in Chinese)
- [19] FU J,ZHENG H, MEI T. Look closer to see better: recurrent attention convolutional neural network for fine-grained image recognition[C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition,2017:4476-4484.
- [20] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]// Proceedings of IEEE Conference on Computer Vision and Pattern Recognition,2018:7132-7141.
- [21] WANG Q,WU B,ZHU P,et al. ECA-Net: efficient channel attention for deep convolutional neural networks[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE,2020:11531-11539.
- [22] QIN Z, ZHANG P, WU F, et al. Fcanet: frequency channel attention networks[C]//Proc. of the IEEE/CVF International Conference on Computer Vision, 2021: 783-792.
- [23] WOO S,PARK J,LEE J,et al. CBAM:convolutional block attention module[C]//European Conference on Computer Vision, 2018:3-19.
- [24] 罗建豪,吴建鑫. 基于深度卷积特征的细粒度图像分类研究综述[J]. 自动化学报,2017,43(8):1306-1318.
LUO Jianhao,WU Jianxin. A survey on fine-grained image categorization using deep convolutional features[J]. Acta Automatica Sinica,2017,43(8):1306-1318. (in Chinese)
- [25] 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.
- [26] HE T, ZHANG Z, ZHANG H, et al. Bag of tricks for image classification with convolutional neural networks[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2019.
- [27] MA N,HANG X,ZHENG H T, et al. ShuffleNet V2: practical guidelines for efficient CNN architecture design[C]//European Conference on Computer Vision. Springer, Cham,2018:122-138.
- [28] SZEGEDY C,IOFFE S,VANHOUCKE V,et al. Inception-v4,Inception-ResNet and the impact of residual connections on learning[C]//Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence, 2017:4278-4284.
- [29] HUANG G,LIU Z,LAURENS V,et al. Densely connected convolutional networks[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR),2017:2261-2269.
- [30] SELVARAJU R R, COGSWELL M, DAS A, et al. Grad-cam: visual explanations from deep networks via gradient-based localization[C]//Proceedings of the IEEE International Conference on Computer Vision,2017:618-626.