

肖天赐,陈燕红,李永可,等. 基于改进通道注意力机制的农作物病害识别模型研究[J]. 江苏农业科学,2023,51(24):168-175.  
doi:10.15889/j.issn.1002-1302.2023.24.023

# 基于改进通道注意力机制的农作物病害识别模型研究

肖天赐<sup>1</sup>, 陈燕红<sup>1,2,3</sup>, 李永可<sup>1,2,3</sup>, 李雨晴<sup>1</sup>, 罗玉峰<sup>4</sup>

(1. 新疆农业大学,新疆乌鲁木齐 830052; 2. 智能农业教育部工程研究中心,新疆乌鲁木齐 830052;  
3. 新疆农业信息化工程技术研究中心,新疆乌鲁木齐 830052; 4. 社旗县中等职业学校,河南社旗 473399)

**摘要:**准确地识别农作物病害种类、病害程度,是能够正确防治病害的基础,对农作物的高质量生产有重要意义。针对传统深度学习模型对图像的细粒度分类不够精准的问题,提出不参与残差计算的通道注意力(efficient channel attention without participating in residual calculation,EWPRC)结构,该结构将改进的通道注意力机制 ECANet3 放在残差块之后,增加模型对通道维度的权重学习能力,并将 EWPRC 结构用于骨干网络为 ResNet50 的迁移学习模型中,通过替换模型中的 layer3、layer4 层得到了 EWPRC-ResNet-t 模型。试验使用了 AICChallenger 2018 数据集,在数据预处理、数据增强、超参数相同的情况下,首先对比了固定核大小为 3、5、7、11、13 的通道注意力机制对模型准确率的影响,在此试验中,模型的准确率随卷积核变大呈下降趋势,其中一维卷积核大小为 3 的模型准确率最高,达到了 87.42%,比核大小为 5、7、11、13 的模型分别提高了 0.03、0.42、0.51、0.64 百分点。再将 EWPRC-ResNet-t 模型与经过微调的迁移学习模型 ResNet-t 以及 GoogLeNet、MobileNet-v3、ResNet50 模型进行对比,以准确率、精确率、召回率以及  $F_1$  值作为评价指标,试验结果证明 EWPRC-ResNet-t 模型取得了最好的效果,比传统深度学习模型中准确率最高的 ResNet50 模型提高了 0.99 百分点,比 ResNet-t 模型提高了 0.75 百分点。且相比传统的深度学习模型,EWPRC-ResNet-t 模型有更高的精度、召回率与  $F_1$  得分。

**关键词:**农作物病害识别;通道注意力机制;残差网络;迁移学习;数据增强

**中图分类号:**S126;TP391.41 **文献标志码:**A **文章编号:**1002-1302(2023)24-0168-07

作物在生产过程中,常由于其他生物的侵染或不适宜的环境导致生长异常、生理机能或结构发生恶性变化,最终导致作物的产量与质量降低,更为严重时甚至导致作物整株死亡。及早发现病害,对作物病害的防治至关重要,其中对作物病害种类、病害程度的识别,是对症、对量下药的关键。然而在种植产业中,依靠人工检测,费时费力,对检测人员也有很高的要求,且准确率不高,因此病虫害检测的信息化是必不可少的。随着互联网的飞速发展,数据的获取变得更加容易,数据量也更为庞大,深度学习等技术获益颇丰,很多学者开展了基于深度学习病害识别方法的研究<sup>[1-5]</sup>。在此基础上,一些学者使用迁移学习的思想,使用模型在其

他数据集上学习的参数作为初始化参数,从而使模型达到更好的效果<sup>[6-8]</sup>。

Agarwal 等对比了支持向量机(support vector machine,SVM)等许多机器学习算法与卷积神经网络,构建了一种由 8 个隐藏层组成的简化 CNN 模型(proposed model)<sup>[9]</sup>。该模型相比于预训练模型参数更少,准确率高于传统机器学习模型与预训练模型,并且在公开数据集 PlantVillage 上达到了 98.7% 的准确率。Wang 等提出 MB-ResNet-18 模型,通过将 SE-Net 嵌入 ResNet-18 中,并使用特征图分层合并方法进一步增加模型的准确率,增加纵向连接,将 1 个输出节点的模型变为 3 个输出节点;此外还改进了损失函数,将 3 个分类任务的损失函数合并为 1 个联合损失函数<sup>[10]</sup>。此模型在部分 AICChallenger 2018 数据集上表现优异,其中对作物的分类准确率较原始的 ResNet-18 模型下降了 0.67%,但对作物病害与病害程度的预测准确率分别提高了 0.51% 和 2.11%。Thakur 等提出了一种轻量级卷积神经网络“VGG-ICNN”,该模型由 1 个 Max Pool、3 个 Inception 块、1 个 GAP 和 1 个 FC

收稿日期:2023-05-09

基金项目:新疆维吾尔自治区重大科技专项(编号:2022A02011)。

作者简介:肖天赐(2000—),男,河南安阳人,硕士研究生,研究方向为计算机视觉。E-mail:2351175028@qq.com。

通信作者:陈燕红,硕士,副教授,研究方向为视觉-语言智能表征学习,E-mail:cyh@xjau.edu.cn;李永可,硕士,副教授,研究方向为智慧农业,E-mail:lyk@xjau.edu.cn。

Layer 加 Softmax 组成,只有约 600 万个参数组成,远少于大多数可用的高性能深度学习模型<sup>[11]</sup>。该模型在 PlantVillage 等 5 个数据集上都表现出一致的良好性能,其中在 PlantVillage 上达到了 99.16% 的准确率。高荣华等提出了 SMLP 模块,该模块将 SENet 中的 2 个线性层替换为多层感知机,并将改进的通道注意力机制融入到 ResNet 中得到 SMLP - ResNet 模型<sup>[12]</sup>。经验证 SMLP - ResNet 在 PlantVillage、AICChallenger 2018 这 2 种数据集上准确率分别达到了 99.32% 与 86.93%,相比原始的 ResNet 与融入了 SENet 的 SE - ResNet 模型,在 PlantVillage 上分别提高了 0.27%、0.13%。在 AICChallenger 2018 上分别提高了 3.00% 与 1.45%,均有不同程度的提升。姜红花等选取 ResNet18 作为骨干网络,通过引入混合注意力机制与随机裁剪分支的方式对其进行改进,得到了 ResNet18 - CBAM - RC1 模型<sup>[13]</sup>。该模型在苹果常见的 5 种病害分类上准确率达到 98.25%,较 ResNet18 模型提高了 5.06%,较 VGG16 提高了 2.12%。

随着深度学习的发展,很多成果通过对空间维度进行改进从而提升了网络的性能,压缩和激励网络(squeeze - and - excitation networks, SENet)注意到了通道上的关系,使用挤压(squeeze)和激励(excitation)的思想学习每个卷积块的通道注意来提升网络性能<sup>[14]</sup>。后有许多基于通道注意力的研究,皆致力于开发更复杂的注意力模块,通过捕获更复杂的信道依赖关系或与额外的空间注意结合来改进 SENet 模块,以获得更好的性能,这不可避免地增加了模型的复杂性<sup>[15-18]</sup>。Wang 等研究证明 SENet 采用的降维操作会对通道注意力的预测产生负面影响,获取依赖关系效率低且不必要<sup>[19]</sup>。基于此对 SENet 进行改良,构造了高效通道注意力网络(efficient channel attention network, ECANet)模块,该模块避免了 SENet 中的降维操作,通过一维卷积的方式有效地实现了跨通道交互,产生了轻量级的、副作用更小的高效通道注意力模块。

为了构建在作物病害分类上有良好性能模型,选用分类难度较大的 AICChallenger 2018 数据集对模型进行训练、测试与对比。试验的骨干网络选取 ResNet50 模型,引入不参与残差计算的通道注意力(efficient channel attention without participating in residual calculation, EWPRC)结构,该结构引入了改进后的通道注意力机 ECANet3,使模型拥有对通道

维度的注意力并增强了注意力在网络中的影响,增强了模型对相似图片的区分能力,再引入迁移学习并进行冻结与微调,最后生成改进后的 EWPRC - ResNet - t 模型。经过对比 EWPRC - ResNet - t 模型与传统深度学习模型,其他学者提出模型的试验结果,以准确率、精度、召回率与  $F_1$  分数为参考指标,验证了改进的模型相比于传统深度学习模型有更高的准确率。使用 EWPRC 结构的模型与原模型有着几乎相同大小的参数,证明了 EWPRC 结构有助于提升模型的准确率,且 EWPRC 结构中的 ECANet3 有轻量级、易添加的特质。

## 1 材料与方法

### 1.1 数据来源

为了测试模型对图像的细粒度识别能力,本试验使用了 AICChallenger 2018 数据集进行训练。该数据集包含 36 254 张作物病害图像,其中训练集有 31 715 张图像,测试集有 4 539 张图像。该数据集包含苹果、樱桃、玉米、葡萄、柑橘、桃、辣椒、马铃薯、草莓、番茄共 10 种作物的 27 种病害,其中包括 10 种健康分类,3 种病害只有 1 种病害程度,24 种病害有一般、严重 2 种病害程度的分类共计 61 个类别,由于编号为 44、45 的种类数据只有 1 张,不具备测试意义,所以在实际训练与测试时没有这 2 种分类,共计 59 个类别,详细数据如表 1 所示。

在 AICChallenger 2018 数据集中,根据病害发病的特点以及严重程度,对病害进行了一般和严重 2 种程度的分类,通过对比观察,不同程度的病害相似度较高,识别难度大,特别是一般程度病害与严重程度病害,某些健康程度与一般程度的图像也不易区分。由于种类较多、区分难度较大,模型能够准确识别的难度也大大增加,如图 1 所示,3 种病害示例分别为柑橘黄龙病、苹果黑星病、葡萄黑腐病,每种病害按照健康、一般病害、严重病害 3 种程度划分为 3 类。

### 1.2 数据预处理与数据增强

图片在输入模型之前先进行预处理操作。先将每张图片重置为 224 像素 × 224 像素大小,再经过以下数据增强操作,包括随机剪裁为 192 像素 × 192 像素大小、概率为 0.5 的水平翻转、概率为 0.5 的垂直翻转的操作,经此操作增强模型的泛化能力,避免出现过拟合现象。图 2 为苹果黑星病图片经过预处理以及图像增强后的结果。

表 1 AI Challenger 2018 数据集

标签编号	标签名称	标签编号	标签名称
0	苹果健康	31	辣椒疮痂病一般
1	苹果黑星病一般	32	辣椒疮痂病严重
2	苹果黑星病严重	33	马铃薯健康
3	苹果灰斑病	34	马铃薯早疫病一般
4	苹果雪松锈病一般	35	马铃薯早疫病严重
5	苹果雪松锈病严重	36	马铃薯晚疫病一般
6	樱桃健康	37	马铃薯晚疫病严重
7	樱桃白粉病一般	38	草莓健康
8	樱桃白粉病严重	39	草莓叶枯病一般
9	玉米健康	40	草莓叶枯病严重
10	玉米灰斑病一般	41	番茄健康
11	玉米灰斑病严重	42	番茄白粉病一般
12	玉米锈病一般	43	番茄白粉病严重
13	玉米锈病严重	44	番茄疮痂病一般
14	玉米叶斑病一般	45	番茄疮痂病严重
15	玉米叶斑病严重	46	番茄早疫病一般
16	玉米花叶病毒病	47	番茄早疫病严重
17	葡萄健康	48	番茄晚疫病一般
18	葡萄黑腐病一般	49	番茄晚疫病严重
19	葡萄黑腐病严重	50	番茄叶霉病一般
20	葡萄轮斑病一般	51	番茄叶霉病严重
21	葡萄轮斑病严重	52	番茄斑点病一般
22	葡萄褐斑病一般	53	番茄斑点病严重
23	葡萄褐斑病严重	54	番茄斑枯病一般
24	柑橘健康	55	番茄斑枯病严重
25	柑橘黄龙病一般	56	番茄红蜘蛛损伤一般
26	柑橘黄龙病严重	57	番茄红蜘蛛损伤严重
27	桃健康	58	番茄黄化曲叶病毒病一般
28	桃疮痂病一般	59	番茄黄化曲病毒病严重
29	桃疮痂病严重	60	番茄花叶病毒病
30	辣椒健康		

1.3 模型搭建

在深度学习模型变深时,参数更新的链式法则变长,模型变得难以训练,很容易遇到梯度消失、梯度爆炸、过拟合甚至网络退化等问题。He 等提出残



图1 不同程度的柑橘黄龙病、苹果黑星病、葡萄黑腐病示例

差学习(residual learning),这种结构通过拟合残差映射的方式将浅层的输出与深层的输出相关联,减少了信息的丢失,同时使参数的更新法则变为,倒数恒大于1,这使参数在反向传播时,即使面对很长的链式求导也不会出现梯度消失的情况,更容易被优化<sup>[20]</sup>。传统模型对通道这一维度没有特殊处理,意味着对所有通道的重视程度是一样的,但有些通道对结果的影响较小,而有些通道则富含重要的分类信息,因此增加模型对通道的注意力,有利于模型捕捉包含关键信息的通道,从而提升模型的分类能力。且通道的交互范围对于模型的识别能力也有影响,交互范围不当,则通道之间起不到相互的促进作用,从而影响模型识别的准确率。为提高模型对于细粒度分类的准确率,增加模型作物病害的细粒度识别能力,及以上深度学习可能遇到的问题,基于对模型结构影响最小、更方便迁移学习实现的角度出发,提出了基于残差结构与通道注意力的 EWPRC 模块,该模块对 ECANet 进行了改进,将改进后的 ECANet3 放在残差块之后构成新的结构,增加了模型对通道的注意能力,改善了通道注意力



图2 苹果黑星病图像增强示例

在进行细粒度分类任务时的感受野,再将其应用到 ResNet50 的迁移模型中,通过冻结与微调,搭建了 EWPRC - ResNet - t 模型。

### 1.3.1 固定核大小的通道注意力机制 ECANet3

ECANet 使用池化与一维卷积的操作进行了通道维度的注意,其中一维卷积核的大小采用自适应的方式自动生成,使得卷积核与通道数成比例。此处不采用自适应卷积核大小的操作,而使用固定大小为 3 的一维卷积,其结构如图 3 所示。

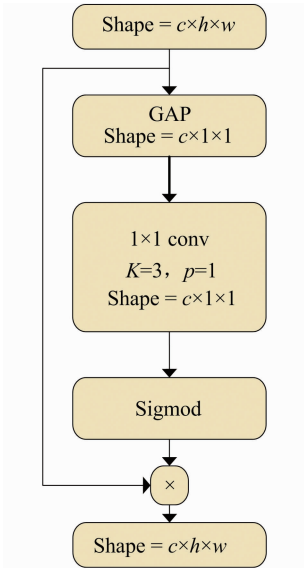


图3 通道注意力——ECANet3

其中 GAP(global average pooling)层计算了每张特征图中所有像素点的均值,使每个通道中高为  $H$ 、宽为  $W$  的特征图形状变为了  $1 \times 1$ ,所有通道的形状为  $c \times 1 \times 1$ 。GAP 层对每张特征图  $u$  的计算如公式(1)所示。

$$F(u) = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W u_{(i,j)} \quad (1)$$

第2步通过固定核大小为3、步长为1、填充为1的一维卷积操作,这一步即对通道进行权重学习的过程,计算如公式(2)所示。

$$Y_{i-2} = \sum_{i=2j=0}^{w+1} \sum_{j=0}^2 W_j \times x_{i+j} \quad (2)$$

式中: $W$ 表示长度为3的一维特征; $W_j$ 表示一维卷积核中第 $j$ 个参数。经过固定核大小的一维卷积操作之后,映射特征的形状没有发生改变。与原始 ECANet 的区别是,原始的 ECANet 取自适应大小的奇数卷积核,而在通道数较大时,自适应的卷积核数也会较大,如 2 048 大小的通道由自适应公式得到卷积核大小为 7。而大的感受野对通道的注意力

并不是完全友好的。在下面不同通道交互范围对比试验中得出结论:模型的性能随卷积核变大总体呈下降趋势,其中卷积核的大小固定为 3 时效果最优。

再通过 Sigmoid 函数将像素值非线性值映射到 0~1 之间,其计算如公式(3)所示。

$$\text{Sigmoid}(x) = \frac{1}{1 + e^{-x}} \quad (3)$$

最后将输出  $y$  与原始特征  $x$  相乘,相乘后的特征便有了权重信息,即对每个通道的注意力,计算如公式(4)所示。

$$\text{Out}_i = y_i \times x_i \quad (4)$$

式中: $i$ 表示第 $i$ 个通道; $\text{Out}_i$ 表示第 $i$ 个通道的输出。

1.3.2 基于 ECANet3 与残差计算的 EWPRC 网络结构 在残差结构中,浅层的输出会与深层的输出通过建立残差映射的方式建立链接,这也意味着与普通结构相比,残差结构中增加了上一层的信息,这使每个残差块参数影响变小。因此将残差层与 ECANet3 做相连而非嵌入,得到 EWPRC 结构,其结构如图 4 所示,图中残差层的结构取 ResNet50 中 layer4 残差层的结构。在残差结构中,每层的输出都与上一层的输出  $h(x)$  相关联,使得整个残差块的输出为  $x + h(x)$ 。每个残差块的第 1 个残差层额外对上一层的输出做了卷积与批标准化操作。在整个残差块之后加入固定核大小为 3 的注意力机制 ECANet3。加入的通道注意力机制不参与任何残差计算,而是单独作为一层,跟在整个残差块之后形成新的 EWPRC 结构。这样的结构不仅添加的参数少,且对原模型结构的影响微小,便于迁移模型的微调。与原模块相比,该模块只加入了 1 个核大小固定为 3 的通道注意力机制,共包含 1 个全局最大池化层,1 个一维卷积,1 个 Sigmoid 层和 1 个相乘操作。

1.3.3 基于 EWPRC 结构与迁移学习的 EWPRC - ResNet - t 模型 为减少模型的训练参数、加速模型的训练,选取骨干网络为 ResNet50 的迁移学习模型,引入了基于 EWPRC 结构与迁移学习的 EWPRC - ResNet - t 模型结构如图 5 所示。该模型的初始化参数为 ResNet50 在 ImageNet 数据集上训练好的参数,将 ResNet50 的 Layer3、layer4 层的结构换为 EWPRC 结构,并冻结了包括 layer2 层之前所有的参数,使之不参与更新,只更新 EWPRC3、EWPRC4、avgpool 与



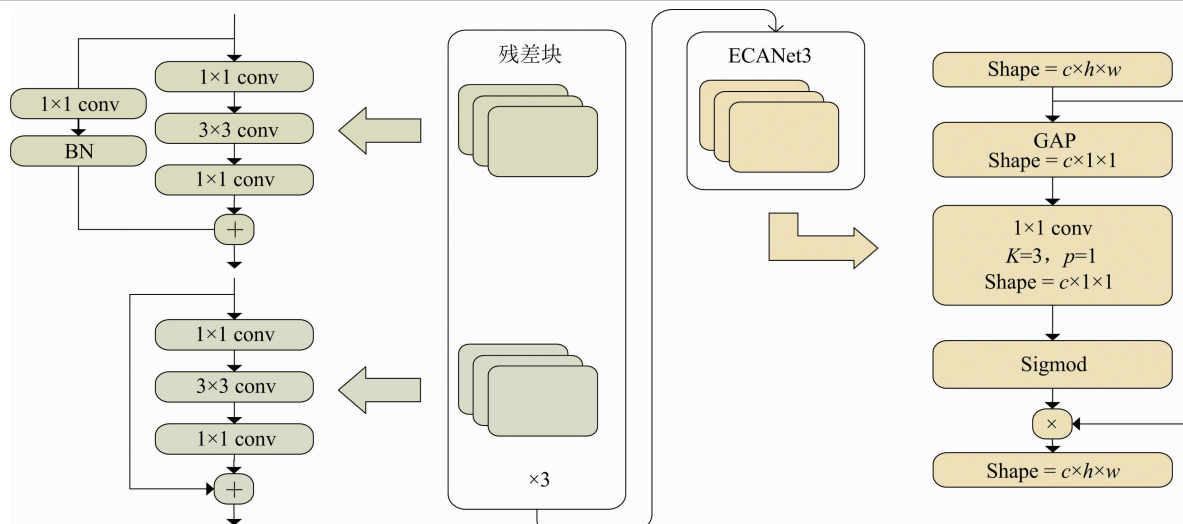


图4 EWPRC 结构

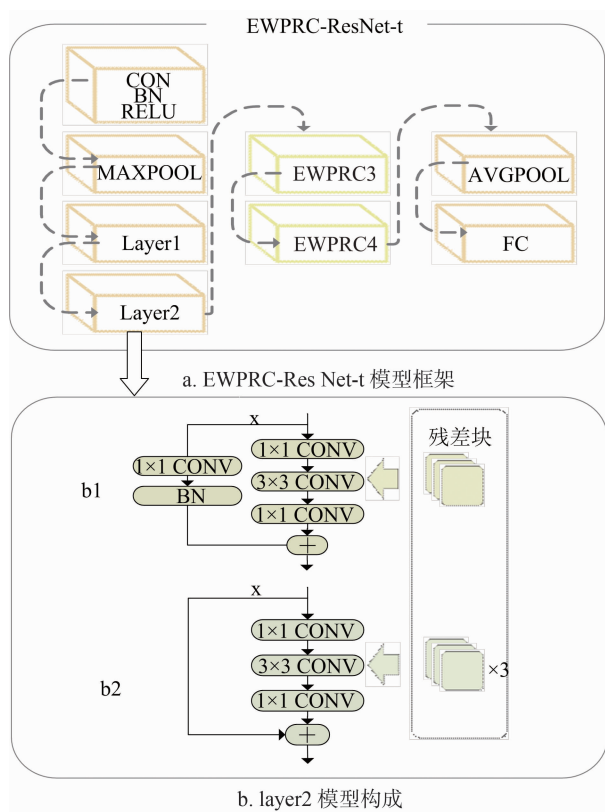


图5 EWPRC-ResNet-t 模型

t 模型中,每批量张图片经过模型后得到对应数量张图片的种类预测结果。

## 2 结果与分析

### 2.1 试验条件

本试验在操作系统为 Window 11 的条件下进行。计算机内存为 16 GB, GPU 为 NVIDIA GeForce RTX 3060, 显存为 6 GB, 使用了 python 3.8 的版本, 模型的实现均使用了 pytorch 1.13.1 版本, torchvision 0.14.1 版本, CUDA 版本为 11.6。

### 2.2 试验数据及参数

本试验使用 AICChallenger 2018 数据集。为防止模型对固定顺序的数据出现依赖从而导致模型不具有泛化能力的现象, 训练数据随机打乱。将图片初始化为统一大小, 并对其中像素进行了均值与方差均为 0.5 的标准化操作, 此外对训练数据集还额外采用了随机剪裁、水平翻转、垂直翻转来进行数据增强操作。使用 Adam 优化器对网络参数进行更新。使用指数衰减的学习率衰减算法, 初始学习速率为 0.001, 每进行 7 次训练学习率变为当前的 0.1 倍。由于显存为 6 GB, 试验均采用小批量训练方法, 批大小设置为 64, 共训练 20 次。每次训练迭代所有的数据集, 其中训练数据为 31 716 张图片, 测试数据为 4 539 张图片, 测试集不参与模型的训练, 只进行模型 4 种评价指标的测试。评价指标采用模型在 AICChallenger 2018 数据集上的准确率、召回率、精度与  $F_1$  分数。

### 2.3 不同通道交互范围对比试验

此部分对模型中通道交互范围进行测试, 在所

最后的全连接层 FC, 并将 FC 层的 outfeature 由原来的 1 000 分类改为 59 分类。layer 模块的结构如图 5-b 所示, 以 layer2 模块为例, 每个残差块的第 1 个残差层 (如图 5 中 b1 所示) 的  $x$  映射经过  $1 \times 1$  的卷积与 BN (batch normalization) 剩余的 3 层残差层 (如图 5 中 b2 所示) 直接将  $x$  映射到结果中, 共包含 4 层残差层。其中 layer1、layer2、EWPRC3、EWPRC4 模块的残差层数分别为 3、4、6、3 层。图片在经过数据预处理与数据增强后, 输入到 EWPRC-ResNet-

有条件均相同的情况下,对比通道注意力机制中不同通道感受范围对模型准确率的影响,并与自适应大小注意力机制比较。具体参数与对比结果如表 2 所示。

表 2 不同通道交互范围对比

通道注意力 机制模型	准确率 (%)	EWPRC3 模块 通道交互范围	EWPRC4 模块 通道交互范围
Eca3	87.42	3	3
Eca5	87.39	5	5
Eca7	87.00	7	7
Eca11	86.91	11	11
Eca13	86.78	13	13
Eca	86.71	5	7

表 2 中每个模型中的 EWPRC3 和 EWPRC4 这 2 个模块使用了通道注意力机制,前 5 个模型使用了固定感受范围的通道注意力机制,通道感受范围从小到大依次设置为 3、5、7、11、13,最后一个模型是自适应大小的通道注意力机制,EWPRC3 模块中为 5,EWPRC4 模块中为 7。通过对试验结果的对比分析发现,模型的准确率随卷积核数变大呈下降趋势,验证了在进行图像细粒度分类任务时,通道交互范围越大,模型的识别效果越不佳,且固定核大小的模型在学习和捕捉特征方面更加稳定,能够更好地适应数据的特征,使用较小的卷积核大小能够更好地捕捉通道间的相关性。此试验中 Eca3 的模型表现最好,EWPRC3、EWPRC4 中的通道注意力机制均使用了大小为 3 的一维卷积,准确率达到 87.42%。

#### 2.4 基于 EWPRC - ResNet - t 的作物病害识别试验

多个传统深度学习模型、微调的迁移学习模型与改进后的迁移学习模型在 AICChallenger 2018 数据集上的结果如表 3 所示,其中结尾为“-t”的模型均为使用了迁移学习的模型,且都进行了如下操作:FC 层的微调、冻结包括 layer2 层之前的参数。ResNet - t 为经过微调的 ResNet50 迁移学习模型;ECA - ResNet - t 模型在 ResNet - t 模型中引入了 EWPRC 结构与迁移学习,模型中的通道注意力机制为未改进前的 ECANet;EWPRC - ResNet - t 模型在 ResNet - t 模型中引入了 EWPRC 结构与迁移学习,模型中的通道注意力机制为改进后的 ECANet3;SMLP - ResNet 为高荣华等提出的改进注意力机制的病害识别模型<sup>[12]</sup>。

表 3 不同病害识别模型试验结果展示

模型	准确率 (%)	精度 (%)	召回率 (%)	$F_1$ 分数 (%)
GoogleNet	86.30	84.16	82.16	83.16
MobileNet - V3	84.01	81.35	78.12	79.70
ResNet18	86.21	83.75	82.77	83.26
ResNet50	86.43	84.42	83.02	83.72
ResNet - t	86.67	85.47	83.56	84.50
ECA - ResNet - t	86.72	84.67	83.00	83.83
SMLP - ResNet	86.93	84.15	83.42	83.60
EWPRC - ResNet - t	87.42	85.36	84.23	84.79

从表 3 可以看出,在传统的深度学习模型中,准确率最高的 ResNet50 模型达到了 86.43%,而 ResNet - t 模型达到了 86.67%,比没有经过迁移学习的 ResNet50 模型提高了 0.24 百分点。ECA - ResNet - t 模型准确率达到 86.72%,比只经过迁移学习的 ResNet - t 模型效果高了 0.05 百分点,而 EWPRC - ResNet - t 模型准确率达到 87.42%,比原始的 ResNet50 模型提高了 0.99 百分点,比 ResNet - t 模型提高了 0.75%,比 ECA - ResNet - t 模型高了 0.70 百分点,比高荣华等提出的 SMLP - ResNet 模型提高了 0.49 百分点<sup>[12]</sup>。在所有测试模型中,EWPRC - ResNet - t 有最优的召回率 (Recall) 和  $F_1$  分数 ( $F_1$  score),分别达到了 84.23% 和 84.79%,在精度 (Precision) 指标上达到了 85.36%,比精度最高的 ResNet - t 模型低 0.11 百分点,在所有模型中排第二。由此可见,迁移学习会提高模型的准确率,在引进 ECANet 之后,模型提取了通道上的特征,但在效果上并不明显,而固定且较小的通道感受范围,可以使模型更好地学习通道上的特征提取能力,在图像细粒度分类任务中有较大幅度地提升。

图 6 为模型的得分情况,图中  $x$  轴是第  $k$  次训练, $y$  轴是第  $k$  次训练模型的准确率 (最优)。图 7 是模型的准确率随训练次数的变化情况,图中  $x$  轴是训练次数, $y$  轴是模型的准确率,实线代表的是引入了迁移学习的模型,虚线代表的是传统深度学习模型。从图 6 可以看出,红色点所代表的 EWPRC - ResNet - t 模型准确率为 87.42%,在所有模型中最高,并且在第 16 次训练中获取了最优的参数,达到了最高的准确率。由图 7 可以看出,相比于虚线代表的传统的深度学习模型,以 EWPRC - ResNet - t

为代表的引进了迁移学习的模型,在第 1 次训练时的准确率就达到了 77% 以上,而传统的深度学习模型在第 1 次训练时准确率都在 55% 以下。

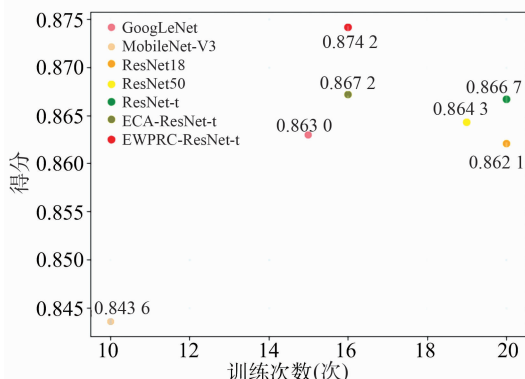


图6 模型得分

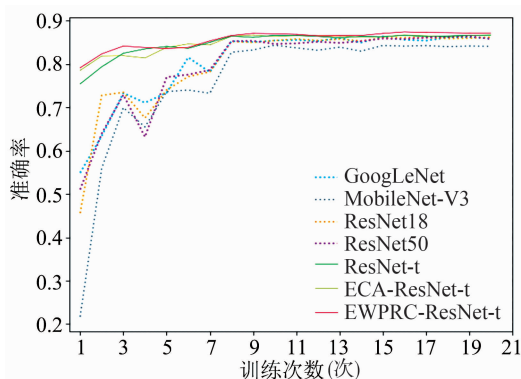


图7 模型训练准确率

### 3 讨论与结论

本研究针对图像细粒度分类任务,对 ECANet 进行通道交互范围的优化,并提出 EWPRC 结构,该结构对原模型影响微小,易于整合到迁移学习模型中,使迁移学习模型的微调便捷,且在图像的细粒度分类上较其他模型有较大提升。将其应用到结合了引进迁移学习的 ResNet50 模型中,得到 EWPRC - ResNet - t 模型。通过在分类难度较大的 AIChallenger 2018 数据集上进行试验,发现在进行图像细粒度分类任务时,固定通道交互范围的通道注意力机制更能提取出有用的通道特征,在整合了固定通道交互范围的通道注意力机制之后,模型对图像的细粒度分类有了较大幅度的提升,弥补了残差结构缺失通道特征表达能力的缺陷。对比 EWPRC - ResNet - t 模型与传统深度学习模型、近期其他学者提出的模型,在作物病害程度的图像细粒度分类任务试验中,EWPRC - ResNet - t 模型的

准确率、精度、召回率与  $F_1$  分数分别达到了 87.42%、85.36%、84.23%、84.79%,在作物病害细粒度分类任务上整体性能更优,且使用了 EWPRC 结构的 EWPRC - ResNet - t 模型比未使用 EWPRC 结构的 ResNet - t 模型准确率提高了 0.75 百分点,比使用了 ECANet 的 ECA - ResNet - t 模型提高了 0.70 百分点,验证了 EWPRC 可以有效提升模型对作物病害识别的准确率。

### 参考文献:

- [1] 张珂,冯晓晗,郭玉荣,等. 图像分类的深度卷积神经网络模型综述[J]. 中国图象图形学报,2021,26(10):2305-2325.
- [2] Huo M Y, Tan J. Overview: research progress on pest and disease identification[C]//Suen, Ching Y. Pattern recognition and artificial intelligence. Cham: Springer International Publishing, 2020: 404-415.
- [3] Xu W X, Sun L, Zhen C, et al. Deep learning - based image recognition of agricultural pests[J]. Applied Sciences, 2022, 12(24):12896.
- [4] Bondre S, Patil D. Recent advances in agricultural disease image recognition technologies: a review[J]. Concurrency and Computation (Practice and Experience), 2023, 35(9):e7644.
- [5] Yuan Y, Chen L, Wu H R, et al. Advanced agricultural disease image recognition technologies: a review[J]. Information Processing in Agriculture, 2022, 9(1):48-59.
- [6] Huang M L, Chuang T C, Liao Y C. Application of transfer learning and image augmentation technology for tomato pest identification[J]. Sustainable Computing (Informatics and Systems), 2022, 33:100646.
- [7] Kathiresan G, Anirudh M, Nagharjun M, et al. Disease detection in rice leaves using transfer learning techniques[J]. Journal of Physics (Conference Series), 2021, 1911(1):012004.
- [8] Tirkey D, Singh K K, Tripathi S. Performance analysis of AI - based solutions for crop disease identification, detection, and classification[J]. Smart Agricultural Technology, 2023, 5:100238.
- [9] Agarwal M, Gupta S K, Biswas K K. Development of efficient CNN model for tomato crop disease identification[J]. Sustainable Computing (Informatics and Systems), 2020, 28:100407.
- [10] Wang C F, Ni P, Cao M Y. Research on crop disease recognition based on Multi - Branch ResNet - 18[J]. Journal of Physics (Conference Series), 2021, 1961(1):012009.
- [11] Thakur P S, Shorey T, Ojha A. VGG - ICNN: a lightweight CNN model for crop disease identification[J]. Multimedia Tools and Applications, 2023, 82(1):497-520.
- [12] 高荣华,白强,王荣,等. 改进注意力机制的多叉树网络多作物早期病害识别方法[J]. 计算机科学,2022,49(增刊1):363-369.
- [13] 姜红花,杨祥海,丁睿柔,等. 基于改进 ResNet18 的苹果叶部病害多分类算法研究[J]. 农业机械学报,2023,54(4):295-303.

施 杰,林双双,罗建刚,等. 基于 YOLO v5s 改进模型的玉米作物病虫害检测方法[J]. 江苏农业科学,2023,51(24):175-183.  
doi:10.15889/j.issn.1002-1302.2023.24.024

# 基于 YOLO v5s 改进模型的玉米作物病虫害检测方法

施 杰,林双双,罗建刚,杨琳琳,张毅杰,顾丽春

(云南农业大学机电工程学院,云南昆明 650201)

**摘要:**针对玉米作物病虫害目标检测中由于图像背景复杂、无关因素干扰较多而导致的对病虫害目标检测效果不理想等问题,采用深度学习技术,提出一种基于 YOLO v5s 改进模型的玉米作物病虫害检测方法。该方法通过引入无参 SimAM 注意力机制对 YOLO v5s 模型进行改进,利用该机制能更全面地评估模型特征权重,以增强玉米病虫害的特征,削弱复杂背景和无关信息的干扰,从而提高模型的检测性能。同时,将模型的损失函数由 CIOU 替换为 EIOU,可使病虫害预测框更加接近真实框,以提升模型检测的准确率。通过自建的玉米病虫害数据集对模型进行试验对比,结果表明,YOLO v5s(EIOU + SimAM)改进模型的精确率为 94.6%,召回率为 83.4%,平均精度均值为 90.1%。经比较发现,改进模型比原始模型在上述 3 个指标方面分别提升了 5.4、0.5、1.9 百分点,说明 YOLO v5s(EIOU + SimAM)模型对玉米病虫害检测具有较强的鲁棒性和泛化性。

**关键词:**玉米作物病虫害;YOLO v5s 改进模型;注意力机制;EIOU;目标检测

**中图分类号:**S126;TP391.41 **文献标志码:**A **文章编号:**1002-1302(2023)24-0175-09

玉米是我国主要的粮食作物,其种植面积和产量长期稳居三大谷物之首。近年来,玉米饲用消费和工业消费均持续增长,玉米产业的健康稳定发展对国家粮食安全意义重大<sup>[1]</sup>。研究发现,玉米生长期间的病虫害问题是影响其产量、质量的主要因素。因此,对玉米常见病虫害进行早期诊断与防治是获得玉米丰产的重要措施<sup>[2-3]</sup>。

农作物在种植期间的传统管理模式主要依靠人工来观测病虫害,存在主观性强、工作量大等问题。随着人工智能和大数据技术的不断发展,利用基于深度学习的机器视觉技术对图像形状、颜色及纹理等特征信息进行提取,能够有效解决农作物种植期间病虫害检测准确率低、效率低等问题,也是农作物种植管理智能化的重要研究方向<sup>[4-7]</sup>。目前,国内外研究者在将深度学习和机器视觉技术应用于农作物管理方面开展了大量研究。Woebbecke 等通过图像数字化对叶片形态特征进行识别与检测,实现了对单子叶植物与双子叶植物的分类,其准确率达到 60%~80%<sup>[8]</sup>。El-Faki 等研究了一种基于颜色特征的杂草识别方法,具有不易受拍摄距离、遮挡等因素影响的优点,试验结果表明,该方

收稿日期:2023-03-16

基金项目:国家自然科学基金(编号:32260438);云南省教育厅科学研究基金(编号:2022J0307)。

作者简介:施 杰(1981—),男,云南昆明人,博士,副教授,硕士生导师,主要从事农业智能装备、机械故障诊断技术研究。E-mail: km\_shijie@126.com。

通信作者:顾丽春,硕士,实验师,主要从事农业信息化研究。E-mail:307524527@qq.com。

[14] Hu J, Shen L, Sun G, et al. Squeeze-and-excitation networks [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42(8): 2011-2023.

[15] Woo S, Park J, Lee J Y, et al. CBAM: convolutional block attention module [C]//European Conference on Computer Vision. Cham: Springer, 2018: 3-19.

[16] Fu J, Liu J, Tian H J, et al. Dual attention network for scene segmentation [C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach, 2020: 3141-3149.

[17] Chen Y P, Kalantidis Y, Li J S, et al. A2-nets: double attention networks [EB/OL]. [2023-04-12]. [https://arxiv.org/abs/](https://arxiv.org/abs/1810.11579)

1810.11579.

[18] Gao Z L, Xie J T, Wang Q L, et al. Global second-order pooling convolutional networks [C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach, 2020: 3019-3028.

[19] Wang Q L, Wu B G, Zhu P F, et al. ECA-net: efficient channel attention for deep convolutional neural networks [C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle, 2020: 11531-11539.

[20] He K M, Zhang X Y, Ren S Q, et al. Deep residual learning for image recognition [C]//2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, 2016: 770-778.