

王忠培,董伟,朱静波,等.简单三维注意力机制水稻病害识别模型[J].江苏农业科学,2023,51(20):186-193.

doi:10.15889/j.issn.1002-1302.2023.20.027

简单三维注意力机制水稻病害识别模型

王忠培¹,董伟¹,朱静波¹,谢成军²

(1.安徽省农业科学院农业经济与信息研究所,安徽合肥 230001;2.中国科学院合肥智能机械研究所,安徽合肥 230031)

摘要:准确、快速地识别水稻病害并及时采取防治措施,是减少水稻产量损失和提高水稻质量的有效途径之一。以生产上常见的 6 种水稻病害为研究对象,提出一种简单的三维注意力机制水稻识别模型。不同于通道注意力或空间注意力方法将研究对象特征分开考虑而导致研究对象本身固有的三维特性丢失的现象,本研究借鉴人类观察物体时将观察主体作为三维整体考虑的特点,提出算法。不同于 SimAM 算法将输入图像中的激活像素人为设置为 +1 作为正样本、不激活像素设置为 -1 作为负样本的假定,本研究不对输入图像的每个像素作人为硬性阈值的设定,而是保留其本身输入特征大小;这种设定不会破坏研究对象本身的固有属性,更符合研究主题自身的特性。研究结果表明,在自建水稻病害识别数据集达到的最高准确率为 98.6%,比 SimAM 算法提高 0.84 百分点;相比经典网络模型 ResNet50、MobileNetV2、EfficientNet_B0、DenseNet 分别提高 1.71、1.93、1.93、0.84 百分点;相比通道注意力机制模型 SENet、ECA 模型分别提高 1.20、1.28 百分点,表明本模型能够为自然环境下水稻病害的智能识别提供技术支持。

关键词:水稻病害;识别;三维注意力;注意力机制

中图分类号:TP391.41 **文献标志码:**A **文章编号:**1002-1302(2023)20-0186-07

水稻是我国第二大粮食作物,种植面积约占我国耕地面积的 1/4,水稻产量关系到国计民生。水稻病虫害对水稻产量产生重要影响。据农业技术推广中心的预报,2023 年预计水稻病虫害发生 7 733.3 万次/hm²,同比增加 25.1%;其中,“三虫两病”(稻飞虱、稻纵卷叶螟、二化螟、纹枯病、稻瘟病)发生 6 720 万次/hm²,同比增加 29.0%^[1]。因此,准确识别病虫害并及时进行防治,有助于减少水稻产量的损失。

近年来,卷积神经网络(convolutional neural networks,CNN)技术由于其性能的优越性,已经在计算机视觉领域获得大量的应用,在农业领域也得到学者的关注并取得良好的效果。赵立新等使用迁移学习方法,实现棉花病虫害的智能识别^[2]。樊湘鹏等将迁移学习的方法应用到葡萄叶片病害识别^[3]。刘斌等将深度可分离卷积和通道混洗 ALS 模块引入轻量化模型 ShuffleNetV2,提出轻量化识别模型 ALS-Net,模型大小只有 1.64 MB,但识别精

度达到 99.43%,可实现移动端苹果叶病害的实时识别^[4]。张善文等将可变形卷积神经网络引入 VGG-16 模型,解决识别对象大小、形态、姿态变化等问题^[5]。Espejo-Garcia 等使用迁移学习的方法,基于 EfficientNetB4 模型在自建数据集上进行微调,可以实现作物营养缺乏症状的诊断^[6]。营养缺乏症可以实现作物早期诊断,避免重大性农业损失。Peng 等为了解决稻田杂草和水稻的相似性问题,提出基于 RetinaNet 的 WeedDet 模型,该模型的 mAP 值为 94.1%,比基准模型 RetinaNet 提高 5.5% 的准确率^[7]。由以上研究可知,深度学习技术在农业方面的应用主要包括迁移学习方法、卷积神经网络技术及深度可分离卷积、可变形卷积神经网络等技术。

随着研究的深入,一些学者针对卷积神经网络提出通道注意力机制。Hu 等首先提出的 SENet 网络模型取得 2017 年 ILSVRC(imagenet large scale visual recognition challenge)竞赛图像分类子任务中的第 1 名,并且 top-5 的错误率相对于 2016 年的第 1 名提升了大约 25%,证明通道注意力机制具备有效性^[8]。Zhao 等使用迁移学习结合注意力机制提出 DTL-SE-ResNet50 模型,该模型主要将 SENet 方法引入 ResNet50 模型^[9]。Zhao 等将改进的块注意力模块(convolutional block attention

收稿日期:2023-01-31

基金项目:国家自然科学基金(编号:32171888)。

作者简介:王忠培(1981—),男,安徽金寨人,博士,助理研究员,研究方向为智能农业技术。E-mail:wangzhongpei@aaas.org.cn。

通信作者:董伟,硕士,副研究员,研究方向为植物保护信息化技术。E-mail:dongwei@aaas.org.cn。

module, CBAM) 引入网络, 实现对玉米、土豆、西红柿病害叶片的分类识别, 3 种作物的整体识别准确率达到 99.55%^[10]。赵辉等将高效通道注意力机制 (efficient channel attention, ECA) 引入 DenseNet-121 网络结构, 并应用到田间杂草识别任务, 改进后模型的平均识别准确率提高了 2.09 百分点, 为智能除草机器人的开发奠定了技术基础^[11-12]。孙俊等将轻量型坐标注意力机制引入 MobileNet-V2 模型, 并将分组卷积引入模型, 实现田间农作物叶片病害的识别^[13]。甘雨等将坐标注意力机制 (coordinate attention, CA) 引入 EfficientNet 网络结构, 提出了一种改进的轻量化作物害虫识别模型 CA-EfficientNet^[14-15]。宋怀波等将挤压激发块 (Squeeze-and-Excitation block) 和非局部块 (Non-Local block) 这 2 种注意力方法引入 YOLO v4 网络模型, 提出一种改进的 YOLO v4-SENL 模型, 在苹果幼果的检测任务中可以达到 96.9% 的平均检测精度^[16-17]。

以上研究表明, 现阶段注意力机制在农业方面主要使用通道注意力的应用方法, 部分使用通道注意力结合空间注意力机制方法, 鲜有应用三维注意力机制。本研究提出一种简单的三维注意力实现方法, 不改变输入图像特征的固有属性, 通过对输入图像的三维整体乘以激活函数获得三维图像每个像素的激活值, 将获得的激活值乘以输入图像以获得输入图像的三维注意力。同时探讨不同激活函数对算法性能的影响, 结果表明 Sigmoid 激活函数可以获得最好的性能, 最终选取 Sigmoid 函数作为本研究激活函数。

1 数据集

本研究以常见 6 种对水稻产量影响较大的病害作为研究主体, 分别是水稻胡麻斑病、水稻白叶枯病、水稻细菌性条纹病、稻曲病、稻瘟病。为了真实反应自然场景, 所有病害图像均使用佳能单反相机在田间直接拍摄。

6 种主要水稻病害的部分示例如图 1 所示。6 种病害的样本量如表 1 所示, 共计 6 938 幅病害图像。由表 1 可知, 这些病害样本量分布具有不平衡性, 比如稻瘟病的样本量是水稻纹枯病的近 4 倍。

2 简单三维注意力机制

2.1 三维注意力机制

当前注意力机制实现主要有 2 种方法: 通道注

意力、空间注意力。通道注意力只能激活通道相关特征, 空间注意力只能激活物体二维空间特征, 而人在观察物体时是按照物体固有的三维特性作为观察目标, 因此三维注意力机制更符合人类观察物体的特点。Yang 等提出一种被称为 SimAM 的三维注意力机制实现方式^[18]。SimAM 模型基于神经学理论, 将神经学的空间抑制性转化为二分类问题, 实现三维注意力机制^[19]。该模型是一种无参模型, 且二分类的优化可以求出一个简单的闭式解, 可以实现快速计算以及可解释性。

Webb 等的研究结果证明, 被激活的神经元通常会抑制周围的神经元, 这种特性被称为空间抑制性^[19]。空间抑制性说明被激活的神经元应该被赋予较高的权重, 而在它周围的神经元被赋予较低的权重。为了实现该特性, 需要将激活神经元和非激活神经元进行区分, 一种简单的方法是将其转换为简单的分类任务。因此, 定义如下损失函数:

$$e_t(w_t, b_t, y, x_i) = (y_t - \hat{t})^2 + \frac{1}{M-1} \sum_{i=1}^{M-1} (y_o - \hat{x}_i)^2. \quad (1)$$

式中: e_t 表示能量方程; y 表示标签; y_t 表示目标的真实标签; y_o 表示除了目标之外的标签。 $\hat{t} = w_t t + b_t$ 中, t 表示目标神经元; w_t 表示权重; b_t 表示偏置。 $\hat{x}_i = w_i x_i + b_i$ 中, x_i 表示输入特征的第 i 个神经元, 输入特征表示为 $\mathbf{X} \in R^{C \times H \times W}$, 其中, \mathbf{X} 表示经过卷积神经网络提取的特征, C, H, W 分别表示输入特征的信道数、高度、宽度, R 表示实数集合。 $M = H \times W$, 表示神经元的总数目, 这里表示输入图像的像素总数, i 表示 $(H \times W - 1)$ 个值的总和。

为了简单起见, 将激活神经元设置为 +1, 其周围的神经元设置为 -1, 即 y_t 设置为 +1, y_o 设置为 -1。同时为提高泛化能力, 添加正则项, 可得到最终的损失函数:

$$e_t(w_t, b_t, y, x_i) = [1 - (w_t t + b_t)]^2 + \frac{1}{M-1} \sum_{i=1}^{M-1} [-1 - (w_i x_i + b_i)]^2 + \lambda w_t^2. \quad (2)$$

公式(2)是一个凸函数, 因此有简单的闭式解:

$$w_t = -\frac{2(t - \mu_t)}{(t - \mu_t)^2 + 2\sigma_t^2 + 2\lambda};$$

$$b_t = -\frac{1}{2}(t + \mu_t)w_t;$$

$$u_t = \frac{1}{M-1} \sum_{i=1}^{M-1} x_i;$$

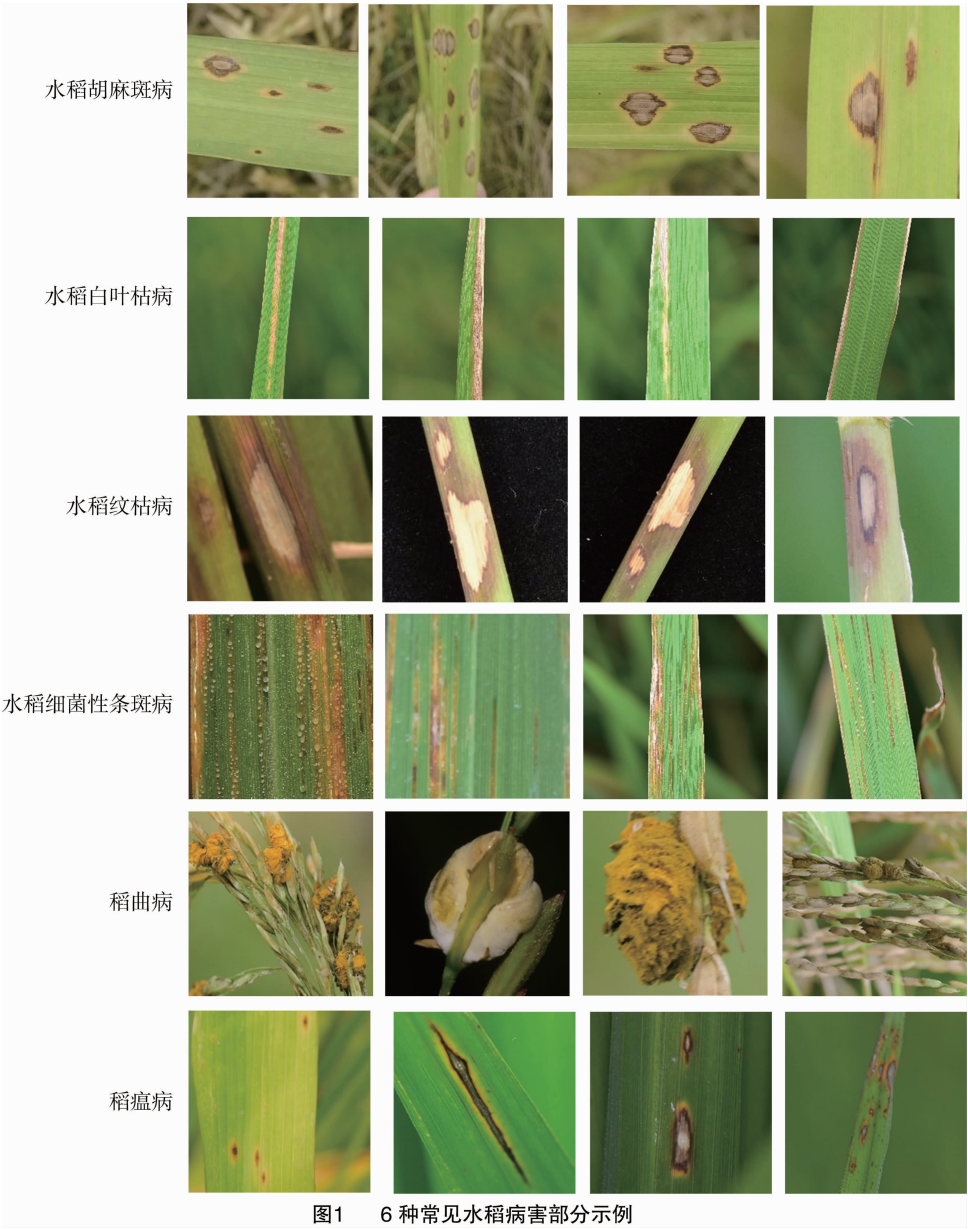


图1 6 种常见水稻病害部分示例

表 1 6 种常见水稻病害样本量

序号	病害名称	样本量 (幅)
1	水稻胡麻斑病	1 458
2	水稻白叶枯病	814
3	水稻纹枯病	548
4	水稻细菌性条斑病	1 498
5	稻曲病	716
6	稻瘟病	1 904

$$\sigma_i^2 = \frac{1}{M-1} \sum_{i=1}^{M-1} (x_i - \mu_i)^2。$$

其中, μ_i 表示除了激活神经元之外的其他神经

元的平均值, σ_i 表示除了激活神经元之外的其他神经元的方差值。

最终可以求得公式(2)的最小值为:

$$e_i^* = \frac{4(\hat{\sigma}^2 + \lambda)}{(t - \hat{\mu}^2 + 2\hat{\sigma}^2 + 2\lambda)}。 \tag{3}$$

公式(3)既是公式(2)的最小值, 其中, $\hat{\mu} = \frac{1}{M} \sum_{i=1}^M$

$x_i, \hat{\sigma}^2 = \frac{1}{M} \sum_{i=1}^M (x_i - \hat{\mu})^2$ 。 e_i^* 表示求得的最小能量值, 值越小表明神经元 t 与周围神经元的区别越大; 当输入的是图像时, 表示激活的像素值和周围的像素值之间的区别越大。所以, 每个神经元的重要性可以通过 $1/e^*$ 得到, 该值表示了每个像素的权重大

小,然后经过 Sigmoid 函数作一次非线性化处理再乘以输入图像,就能得到图像每个像素的激活值大小,最终实现图像的三维注意力机制。

$$\tilde{X} = \text{Sigmoid}\left(\frac{1}{E}\right)X. \quad (4)$$

式中: X 表示输入的图像或者特征值, \tilde{X} 表示经过 Sigmoid 激活以后的值, E 表示对 e^* 分别在空间维度和信道维度进行计算后的值。

2.2 简单三维注意力机制

2.2.1 网络整体结构 本研究提出的简单三维注意力机制模型是以 ResNet50 模型为基础,整体的网络结构如图 2-A 所示。

输入图像首先经过 1 个卷积层,该卷积层的卷积核大小设置为 (7×7) , padding 大小设置为 3, stride 设置为 2,然后经过 16 个被称为 Bottleneck 的单元。其中第 1 个 Bottleneck 的层数是 3 层,输出大小设置为 256;第 2 个 Bottleneck 的层数是 4 层,输出大小设置为 512;第 3 个 Bottleneck 的层数是 6 层,输出大小设置为 1 024;第 4 个 Bottleneck 的层数是 3 层,输出大小设置为 2 048。最后输入 1 个全连接层,最终得出每个类别的概率值。

2.2.2 Bottleneck 结构 Bottleneck 的结构设置如图 2-B 所示。分别由 (1×1) 卷积、 (3×3) 卷积、简单三维注意力机制、 (1×1) 卷积相连接,并使用残差的方式和初始输入连接。

2.2.3 简单三维注意力机制 Yang 等提出的 SimAM 算法对输入图像的三维结构实现整体操作,同时借助神经学理论,将每一个像素转换为二分类问题。受到 SimAM 算法的启发,本研究提出一种简单的三维注意力机制实现方法。SimAM 算法为了实现每个像素的激活值,需要设置正负样本并通过二分类算法求解,例如将正样本设置为 +1,负样本设置为 -1。本研究认为这种假定是一种硬性阈值设置的方式,不符合每个像素的真实值大小。本研究不对每个像素的标签做硬性假定,而是保留每个像素的实际值,经过一个激活函数就可以得到每个像素的权重大小,将得到的激活值乘以输入图像,如公式(5)所示。

$$\tilde{X} = \text{Sigmoid}(X)X. \quad (5)$$

公式(5)是对输入图像的整体操作,是一种三维注意力机制实现方式,保留了原始输入图像特征的固有属性大小,没有作人为预设假定,更能真实反映输入图像本身的固有属性。

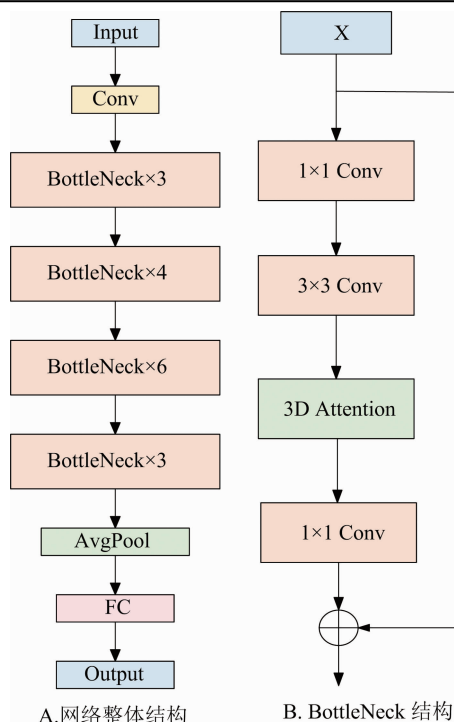


图2 网络整体结构以及 Bottleneck 结构

3 试验设置

3.1 试验环境

本研究使用 Pytorch 软件进行仿真。电脑配置为:CPU, Intel i7 - 9700 @ 3.00 GHz; GPU, 英伟达 GeForce RTX2060, 显存大小为 6 GB。

3.2 试验参数设置

将数据集划分为训练集、验证集,各占 80%、20%;并将批处理大小设置为 16,一共迭代训练 100 次。本研究使用 SGD 优化器,优化器学习率设置为 0.001,动量因子(momentum)设置为 0.9,权重衰减系数(weight_decay)设置为 0.005。

3.3 数据预处理方法

为了保证训练的稳定性,需要对图像做归一化处理,即对图像的 R、G、B 这 3 个通道分别减去均值并除以标准差。均值和标准差的值分别设置为 (0.485, 0.456, 0.406) 和 (0.229, 0.224, 0.225),以上均值和标准差由 ImageNet 数据集中所有图像计算所得。

因为模型训练要求输入的图像尺寸一致,先将图像的短片大小调整为 256 像素,再从图像中心选取 (224×224) 像素作为训练图像的输入尺寸。

4 试验结果与分析

为了公平性,对比分析的所有算法都以

ResNet50 作为基础网络进行仿真试验。

4.1 本研究算法与 SimAM 模型对比试验

因为 SimAM 算法是三维注意力机制模型,而本研究算法是另一种三维注意力机制模型,所以需要在这 2 种模型进行对比分析。2 组试验所有参数设置一致,经过 100 次迭代后,在验证集上的识别准确率曲线如图 3 所示。

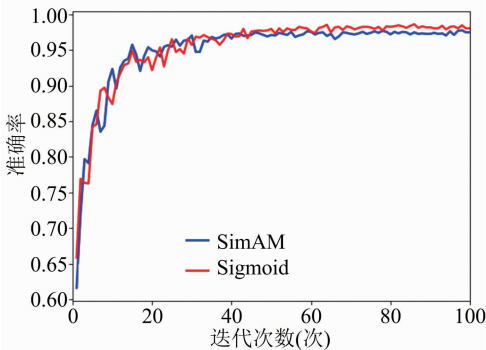


图3 本研究算法与 SimAM 模型准确率对比

由图 3 可知,经过 50 次训练后,本研究所提算法的准确率就开始优于 SimAM,而且随着训练的进行,这种趋势一直保持,证明了本研究算法的优越性和稳定性。

2 种算法的最高准确率以及对应出现的轮次如表 2 所示。本研究模型比 SimAM 模型的最高准确率提高 0.84 百分点。证明虽然本研究算法简单,但因没有破坏三维物体固有的属性特点,反而优于较为复杂的 SimAM 算法。

表 2 本研究模型与 SimAM 模型最高准确率及对应轮次

模型	最高准确率 (%)	对应轮次 (次)
SimAM	97.76	98
本研究算法	98.60	86

4.2 不同激活函数对比试验

为了验证本研究算法的性能表现,尝试使用不同的激活函数进行对比试验。现阶段已提出的激活函数包括:ReLU、ReLU6、PReLU、LeakyReLU、RReLU、ELU、SELU、CELU、GELU、SiLU、Mish、Softplus、Sigmoid、Tanh 等。通过查看以上激活函数的图形,可以将其分为三大类:将 ReLU、ReLU6、PReLU、LeakyReLU、RReLU 统一称为 ReLU 系列;ELU、SELU、CELU、GELU、SiLU、Mish 称为 ELU 系列;而将 Softplus、Sigmoid、Tanh 划分为其他系列,因为这 3 种激活函数既有相似特性又有不同特点,比

如 Softplus、Sigmoid 在 0 轴时,其对应的激活值不等于 0,Sigmoid、Tanh 将激活值限定在一个区间内,Sigmoid 将激活值限定在 $[0,1]$ 区间,Tanh 将激活值限定在 $[-1,1]$ 区间。本研究从以上激活函数类别中选取有各自特点的函数进行仿真验证,经过 100 轮训练后,识别率结果如图 4 所示。

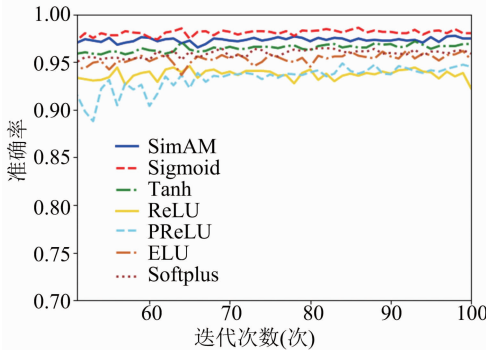


图4 不同激活函数准确率对比曲线图

由图 4 可知,使用 Sigmoid 激活函数取得了最好的结果,而且一直保持优于其他激活函数。使用 Tanh 激活函数取得了第 3 位的效果。相反,ReLU 系列以及 ELU 系列的效果是最差的,说明以上 2 个系列的激活函数不适应于本研究模型。

为了做进一步分析,对不同激活函数在验证集上的最高识别准确率以及对应的轮次进行比较(表 3),结果表明,Sigmoid 激活函数最适用于本研究算法。

表 3 不同激活函数最高准确率及对应轮次

模型	最高准确率 (%)	对应轮次 (次)
SimAM	97.76	98
Sigmoid	98.60	86
Tanh	97.18	93
ReLU	94.58	65
PReLU	94.87	84
ELU	96.24	92
Softplus	96.46	69

4.3 不同网络模型性能对比试验

为了验证本研究算法的性能优势,与三维注意力机制 SimAM 模型和通道注意力机制模型 SENet、ECA 模型^[12]做对比;并同时与经典网络模型 ResNet 模型^[20]、MobileNetV2 模型^[21]、DenseNet 模型^[22]、EfficientNet 模型^[23]做对比试验。为保证对比的公平性,所有试验参数设置一致。各模型经过 100 次训练后,识别准确率如图 5 所示。

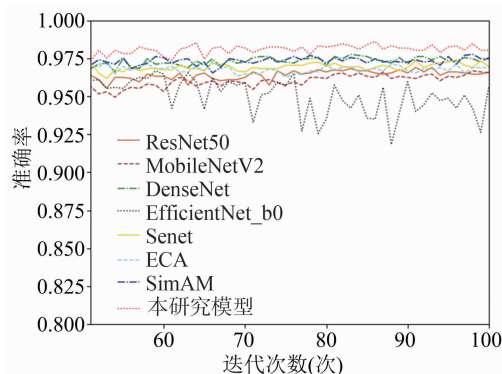


图5 不同模型准确率对比曲线图

试验结果表明,本研究提出的模型性能是最好的,而表现最差的是 EfficientNet_b0 模型。同时,本研究模型优于2个通道注意力机制模型 SENet、ECA 模型,证明了三维注意力机制模型优于单纯的通道注意力机制方法。

为了进一步分析各模型的准确率,分别统计了各模型的最高准确率以及对应的轮次。

表4 不同模型最高准确率及对应轮次

模型	最高准确率 (%)	对应轮次 (次)
ResNet50	96.89	93
MobileNetV2	96.67	98
DenseNet	97.76	83
EfficientNet_b0	96.67	59
SENet	97.40	96
ECA	97.32	53
SimAM	97.76	98
本研究模型	98.60	86

从表4可以看出,8种模型的最高准确率都在96%以上,本研究提出的模型取得了98.60%的最高准确率,比ResNet50、MobileNetV2、EfficientNet_b0、DenseNet模型分别提高1.71、1.93、1.93、0.84个百分点,比通道注意力机制模型 SENet、ECA模型分别提高1.20、1.28个百分点。本研究的模型优于一些经典网络模型和通道注意力机制模型,证明三维注意力机制具有优越性。

4.4 水稻病害识别结果定性分析

混淆矩阵常用来对模型性能做可视化分析,展示每类水稻病害的识别准确率情况。图6为6种常见水稻病害的分类混淆矩阵。

图6中,横坐标编号代表每种水稻病害的ID:10000表示水稻胡麻斑病;10017表示水稻白叶枯病;10018表示水稻纹枯病;10021表示水稻细菌性

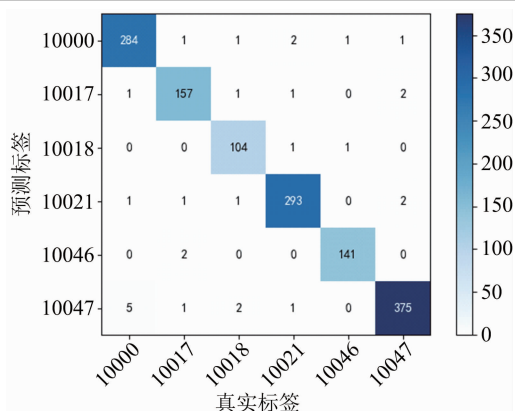


图6 常见水稻病害混淆矩阵图

条斑病;10046表示稻曲病;10047表示稻瘟病。对角线上的数字表示预测正确的个数,非对角线上的数字表示预测为其他类别的数目。比如,水稻细菌性条斑病在测试集上一共有298幅图像,被正确预测的有293幅,5幅水稻细菌性条斑病被预测为其他病害,其中有2幅被预测为水稻胡麻斑病,剩余的3幅分别被预测为水稻白叶枯病、水稻纹枯病、稻瘟病。

4.5 水稻病害识别结果定量分析

为了对每种水稻病害做定量分析,本研究选用以下评价指标:精度、召回率、特异度、 F_1 得分、平均准确率。

精度 P (Precision) 公式定义为:

$$P = \frac{TP}{TP + FP} \times 100\%。 \quad (6)$$

精度表示正样本的数量有多少被正确预测。公式(6)中: TP (true positive)称为真正类,表示样本的真实值为正且预测值也为正的数量; FP (false positive)称为假正类,表示样本的真实值为负却被预测为正样本的数量。

特异度 S (Specificity) 公式为:

$$S = \frac{TN}{FP + TN} \times 100\%。 \quad (7)$$

特异度表示真实标签是负样本同时也被预测为负样本的数量与真实负样本标签的数量比例。公式(7)中 TN (true negative)称为真负类,表示样本的真实标签为负且预测值也为负类的数量。

召回率 R (Recall) 公式为:

$$R = \frac{TP}{TP + FN} \times 100\%。 \quad (8)$$

召回率也称之为查全率,表示预测正确的正样本占有所有真实正样本标签的比例,值越大表示模型的预测能力越好。公式(8)中 FN (false negative)称

为假负类,表示样本的真实标签是正类却被预测为负类。

F_1 得分的公式为:

$$F_1 = \frac{2 \times (P \times R)}{P + R} \times 100\%。$$
 (9)

精度和召回率是一对矛盾的指标,当精度高时召回率往往偏低,而精度低时召回率往往偏高。 F_1 得分综合考虑了精度和召回率,其核心思想是提高精度和召回率的同时,也希望两者之间的差异尽可能小。

准确率 Acc (Accuracy) 的公式为:

$$Acc = \frac{TP + TN}{TP + TN + FP + FN} \times 100\%。$$
 (10)

表示预测正确的样本数占总样本数的比例。但是当正、负样本不平衡时,该指标有一定的缺陷。

根据以上定义,分别计算出每种水稻病害的精度、召回率、特异度、 F_1 得分、平均准确率,结果如表 5 所示。

表 5 6 种常见水稻病害的 5 个指标值

病害 ID	精度 (%)	召回率 (%)	特异度 (%)	F_1 得分 (%)	平均准确率 (%)
10000	97.9	97.6	99.5	97.7	97.90
10017	96.9	96.9	99.6	96.9	
10018	98.1	95.4	99.8	96.7	
10021	98.3	98.3	99.5	98.3	
10046	98.6	98.6	99.8	98.6	
10047	97.7	98.7	99.1	98.2	

表 5 结果表明,精度最低的是水稻白叶枯病(96.9%),召回率最低的是水稻纹枯病(98.1%),特异度各病害差别不大, F_1 得分最低的是水稻纹枯病(96.7%),其次是水稻白叶枯病(96.9%)。该模型对水稻白叶枯病和水稻纹枯病的识别效果不太好。

5 结论

为了快速准确地识别水稻病害,本研究针对 6 种常见水稻病害提出一种简单的三维注意力机制识别模型。首先自建水稻病害数据集,然后针对通道注意力机制以及空间注意力机制的不足,提出三维注意力机制识别模型。不同于 SimAM 方法将输入图像的每个像素按照硬性阈值来划分正负样本的做法,本研究不对每个像素做人为划分而是保留图像本身的固有属性,这种方法更符合图像本身的

特性。通过对比分析,本研究算法的最高准确率可以达到 98.6%,相对于 SimAM 算法提高了 0.84 百分点。同时,相对于通道注意力模型 SENet、ECA 模型,本研究算法分别提高 1.20、1.28 百分点,证明三维注意力机制具备相对优越性,表明本研究模型可以有效解决大田复杂环境下水稻病害的识别问题。

参考文献:

[1] 刘 杰,曾 娟,杨清坡,等. 2023 年全国农作物重大病虫害发生趋势预报[J]. 中国植保导刊,2023,43(1):32-35.

[2] 赵立新,侯发东,吕正超,等. 基于迁移学习的棉花叶部病虫害图像识别[J]. 农业工程学报,2020,36(7):184-191.

[3] 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.

[4] 刘 斌,贾润昌,朱先语,等. 面向移动端的苹果叶部病虫害轻量级识别模型[J]. 农业工程学报,2022,38(6):130-139.

[5] 张善文,许新华,齐国红,等. 基于可形变 VGG-16 模型的田间作物害虫检测方法[J]. 农业工程学报,2021,37(18):188-194.

[6] Espejo - Garcia B, Malounas I, Mylonas N, et al. Using EfficientNet and transfer learning for image - based diagnosis of nutrient deficiencies[J]. Computers and Electronics in Agriculture, 2022, 196:106868.

[7] Peng H X, Li Z H, Zhou Z Y, et al. Weed detection in paddy field using an improved RetinaNet network[J]. Computers and Electronics in Agriculture, 2022, 199:107179.

[8] Hu J, Shen L, Sun G. Squeeze - and - excitation networks[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. June 18 - 23, 2018, Salt Lake City, UT, USA. IEEE, 2018:7132-7141.

[9] Zhao X , Li K Y, Li Y X, et al. Identification method of vegetable diseases based on transfer learning and attention mechanism[J]. Computers and Electronics in Agriculture, 2022, 193:106703.

[10] Zhao Y, Sun C D, Xu X, et al. RIC - Net: a plant disease classification model based on the fusion of Inception and residual structure and embedded attention mechanism[J]. Computers and Electronics in Agriculture, 2022, 193:106644.

[11] 赵 辉,曹宇航,岳有军,等. 基于改进 DenseNet 的田间杂草识别[J]. 农业工程学报,2021,37(18):136-142.

[12] 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 (CVPR). June 13 - 19, 2020. Seattle, WA, USA. IEEE, 2020: 11531-11539.

[13] 孙 俊,朱伟栋,罗元秋,等. 基于改进 MobileNet - V2 的田间农作物叶片病害识别[J]. 农业工程学报,2021,37(22):161-169.

[14] 甘 雨,郭庆文,王春桃,等. 基于改进 EfficientNet 模型的作物害虫识别[J]. 农业工程学报,2022,38(1):203-211.

[15] Hou Q B, Zhou D Q, Feng J S. Coordinate attention for efficient

李 军,李志伟,李艳红. 基于多原型指导的小样本水稻害虫识别与分类[J]. 江苏农业科学,2023,51(20):193–200.

doi:10.15889/j.issn.1002-1302.2023.20.028

基于多原型指导的小样本水稻害虫识别与分类

李 军¹,李志伟²,李艳红³

(1. 山西华澳商贸职业学院科研处,山西晋中 030600; 2. 山西农业大学农业工程学院,山西太谷 030801;

3. 晋中职业技术学院机电工程系,山西晋中 030600)

摘要:精确识别与分类复杂场景下水稻害虫对水稻病害治理与产量提升具有重要的研究意义,提出一种基于多原型指导的小样本水稻害虫识别与分类模型。首先,利用 Vision Transformer 网络作为主干网络进行特征映射,将水稻害虫图片映射到深度特征空间;其次,利用可微分的超像素聚类算法实现深度特征图中类特定原型的聚类分组,构造水稻多类型害虫的特征表达;再次,提出一种双通道特征融合注意力模块实现水稻害虫支持特征和查询特征的深度融合;最后,利用无参数的度量学习算法计算待测水稻病虫害图片特征与深度融合特征之间的距离,根据距离值实现待测病虫害图片的识别与分类。试验结果表明,所提出的水稻害虫识别模型可以实现 0.955 的识别准确率、0.941 的精确率、0.949 的召回率和 0.962 的 F_1 值,与 AlexNet、ResNet、YOLO v5、Vgg-16、Inception v3 和 LeNet5 等模型相比,所提出模型的各项指标表现良好。该方法的提出为农作物害虫的智能化识别提供了新的思路。

关键词:水稻;害虫识别;多原型;小样本学习;超像素聚类;注意力机制

中图分类号:TP391.41 **文献标志码:**A **文章编号:**1002-1302(2023)20-0193-08

水稻是中国乃至世界的主要粮食作物之一,随着全球气候变暖,农作物病虫害发病率逐年上升,严重影响了农作物的产量和质量^[1]。因此,研究水稻病虫害的定位与识别,可以及时预防病虫害对农作物的破坏,对提高农作物的产量和质量具有重要

的意义。

传统的水稻病虫害检测工作主要依靠人工经验完成^[2-4],该方法虽然可以实现部分病虫害的定位与识别,但费时费力、主观性强、效率低、识别精度不高。之后,随着机器学习技术在图像处理领域的成功应用,利用 K 最近邻算法^[5]、支持向量机^[6]、决策树算法^[7]和随机森林^[8]等机器学习模型构造农作物病虫害图片的纹理、颜色、形状等特征的决策模型,并在各类农作物病虫害测试数据集中

收稿日期:2023-01-30

基金项目:山西省重点研发计划(编号:201903D211005)。

作者简介:李 军(1981—),男,山西泽州人,硕士,讲师,研究方向为农业工程。E-mail:lijun198060@163.com。

mobile network design [C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). June 20–25, 2021. Nashville, TN, USA. IEEE, 2021:13713–13722.

[16] 宋怀波,江 梅,王云飞,等. 融合卷积神经网络与视觉注意机制的苹果幼果高效检测方法[J]. 农业工程学报,2021,37(9):297–303.

[17] Wang X L, Girshick R, Gupta A, et al. Non-local neural networks [C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. June 18–23, 2018. Salt Lake City, UT, USA. IEEE, 2018:7794–7803.

[18] Yang L, Zhang R Y, Li L, et al. Simam: a simple, parameter-free attention module for convolutional neural networks [C]//International conference on machine learning. PMLR, 2021: 11863–11874.

[19] Webb B S, Dhruv N T, Solomon S G, et al. Early and late mechanisms of surround suppression in striate cortex of macaque

[J]. The Journal of Neuroscience, 2005, 25(50):11666–11675.

[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 (CVPR). June 27–30, 2016. Las Vegas, NV, USA. IEEE, 2016:770–778.

[21] Sandler M, Howard A, Zhu M L, et al. MobileNetV2: inverted residuals and linear bottlenecks [C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. June 18–23, 2018. Salt Lake City, UT. IEEE, 2018:4510–4520.

[22] Huang G, Liu Z A, van Der Maaten L, et al. Densely connected convolutional networks [C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). July 21–26, 2017. Honolulu, HI. IEEE, 2017:4700–4708.

[23] Tan M, Le Q. Efficientnet: rethinking model scaling for convolutional neural networks [C]//International conference on machine learning. PMLR, 2019:6105–6114.