

王忠培, 谢成军, 董 伟, 等. 基于多维间注意力机制的水稻病害识别模型[J]. 江苏农业学报, 2024, 40(4): 625-635.
doi:10.3969/j.issn.1000-4440.2024.04.006

基于多维间注意力机制的水稻病害识别模型

王忠培¹, 谢成军², 董 伟¹, 管博伦¹

(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: 水稻病害的快速、准确识别是水稻病害防治的前提,也是提高水稻产量和品质的有效途径之一。为了提高水稻病害识别的准确率,本研究提出一种多维间的三维注意力水稻病害识别模型 Inter_3DRiceNet 网络模型,通过 3 个不同维度(通道维度、高度维度以及宽度维度)提取水稻病害特征信息。通道维度主要构建基于通道关系的三维立体注意力机制,通过建立一维的通道间关系注意力机制再结合二维空间关系,最终获得基于通道关系的三维注意力特征信息。高度维度建立的是基于高度维度关系的三维注意力机制,而宽度维度建立的是基于宽度维度关系的立体注意力机制。然后将以上 3 个不同维度的注意力信息进行简单的相加再取平均值作为最终的病害提取特征。通过这种方式,不仅可以获取输入图像更丰富的特征,而且可以获得不同维度的立体空间关系。试验结果表明,在自建的 6 种真实自然环境水稻病害数据集中,本研究提出的 Inter_3DRiceNet 网络模型在测试集取得了 98.32% 的最高准确率,高于经典网络模型 ResNet34、ResNet50、MobileNetV2、DenseNet、EfficientNet_B0 和通道注意力机制模型 SENet 和 GCT。可见本研究方法有效提高了水稻病害的识别准确率,获得了优于经典网络模型和通道注意力模型的识别准确率,有助于提升自然环境下对常见水稻病害的识别性能。

关键词: 水稻病害; 三维注意力; 多维间关系; 注意力机制; 识别

中图分类号: TP391.41 **文献标识码:** A **文章编号:** 1000-4440(2024)04-0625-11

Rice disease identification model based on multi-dimensional attention mechanism

WANG Zhong-pei¹, XIE Cheng-jun², DONG Wei¹, GUAN Bo-lun¹

(1. Institute of Agricultural Economics and Information, Anhui Academy of Agricultural Sciences, Hefei 230001, China; 2. Hefei Institute of Intelligent Machinery, Chinese Academy of Sciences, Hefei 230031, China)

Abstract: Rapid and accurate identification of rice diseases is a prerequisite for controlling rice diseases and is one of the effective ways to improve rice yield and quality. To improve the identification accuracy of rice diseases, a network model of multi-dimensional attention mechanism for rice disease identification named Inter_3DRiceNet was proposed in this study to extract rice disease feature information through three different dimensions (channel dimension, height dimension and width dimension). The channel dimension mainly constructed a three-dimensional attention mechanism based on channel relationship, and finally obtained three-dimensional attention feature information based on channel relationship by estab-

lishing a one-dimensional attention mechanism of inter-channel relationship combined with two-dimensional spatial relationship. The height dimension established a three-dimensional attention mechanism based on the height dimension relationship, while the width dimension established a tridimensional attention mechanism based on the width dimension relationship. The attention information of the a-

收稿日期: 2023-02-24

基金项目: 国家自然科学基金项目(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

通讯作者: 谢成军, (E-mail) cjxie@iim.ac.cn

bove three different dimensions was simply summed and then averaged as the final disease extraction features. Thus, besides more abundant features of the input images could be obtained, stereoscopic spatial relations of different dimensions could also be obtained. The experimental results showed that, the Inter_3DRiceNet network model proposed in the study got the highest accuracy of 98.32% in the test sets of the six self-constructed rice disease datasets in real natural environment, which was higher than the classical network models such as ResNet34, ResNet50, MobileNetV2, DenseNet, EfficientNet_B0, and channel attention mechanism models SENet and GCT. The research method improved the recognition accuracy of rice diseases effectively and obtained better classification accuracy than the classical network model and the channel attention model, which can help improve the performance of common rice diseases recognition in natural environment.

Key words: rice diseases; three-dimensional attention; multi-dimensional relationship; attention mechanism; identification

水稻是中国的主要粮食作物之一,种植面积约占全国耕地面积的四分之一,是中国的第二大粮食作物。由于中国水稻种植具有面积大、分布广的特点,因此不同地区、不同环境等因素(比如天气、气候、土壤、水资源等自然条件)都会影响水稻产量。除了自然条件,水稻病虫害也是影响水稻产量的重要因素之一。全国农业技术推广服务中心对 2022 年全国重大病虫害发生情况的预报结果^[1]显示,2022 年预计水稻“三虫两病”发生面积为 6.87×10^7 hm^2 ,比 2021 年增加 13.6%。因此,水稻病虫害的预警、监测、防治将有助于减少病虫害对水稻产量的影响。

随着技术的进步,中国农业生产方式正在从传统手工方式向智能化方向转变。传统水稻病虫害识别主要通过人工方式,人工识别存在效率低、工作量大以及难以广泛推广等问题,而且人工识别对识别主体的专业技能要求较高,一般只有该领域的专家才能准确识别出相应的病虫害。而作为农业经营主体的农民不具备该领域的专业知识,很难准确识别出不同种类的病虫害,因此如何借助计算机领域相关技术实现农作物病虫害的快速、准确识别是值得研究的课题。

近年来随着卷积神经网络(Convolutional neural networks, CNN)技术的提出,CNN 在计算机视觉领域的各个子任务中获得了大量的应用,并取得了令人瞩目的性能提升。农业作为该技术的实际应用场景之一,如何在农业中应用 CNN 技术同样获得了大量学者的关注,并提出了不同的技术方案。

赵立新等^[2]将迁移学习方法应用到棉花叶片病虫害识别中,通过改进 AlexNet 模型并在 PlantVillage 数据集上预训练,将预训练好的模型迁移到棉花病虫害数据集,可以达到 97.16% 的平均测试准

确率。刘阳等^[3]通过对 SqueezeNet 网络的改进提出一种小型化和轻量化网络结构,以便部署到移动终端等嵌入设备上。鲍文霞等^[4]提出了一种轻量化残差网络(Light weight residual network, LW-ResNet),用于对水稻害虫进行识别。该模型对 13 类水稻害虫的识别准确率为 92.5%,并且该模型的参数量只有 1.62×10^6 个,浮点运算量只有 0.34×10^9 次。Espejo-Garcia 等^[5]使用迁移学习的方法实现对作物营养缺乏症状的诊断,在初始化模型权重时,采用了基于 ImageNet 数据集的噪声学生模型训练结果,然后使用自建的数据集在 EfficientNetB4 模型上微调。

随着研究的深入,一些学者针对卷积神经网络的不足提出了一些改进方案。比如针对卷积神经网络只能提取识别对象的空间信息而无法获得通道间关系信息,Hu 等^[6]提出的 SENet 模型将通道注意力引入网络。SENet 网络模型获得了 2017 年 ImageNet 大规模图像识别竞赛(ImageNet large scale visual recognition challenge, ILSVRC)图像分类子任务中的第一名。使用通道注意力机制可以显著提高模型的识别准确率。为此,有学者将注意力机制方法引入农业领域。刘阳等^[3]主要从网络规模的小型化以及轻量化角度对 SENet 网络结构提出一系列改进措施并在 PlantVillage 数据集集中的 14 种作物 38 类叶片上进行验证。赵辉等^[7]使用另一种被称之为高效通道注意力(Efficient channel attention, ECA)^[8]的机制,ECA 是通道注意力机制的一种,是对原始通道注意力机制 SENet 的一种改进。赵辉等^[7]将 ECA 通道注意力机制应用到田间杂草识别领域,改进后的模型较改进前的模型平均识别准确率提高了 2.09 个百分点,为智能除草机器人开发奠定了技术基础。孙俊等^[9]以 MobileNet-V2 模型为基础,将 ECA 通道注意力机制引入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。改进后的模型较改进前的模型识别准确率提高了 2.91 个百分点,达到了 92.20% 的识别准确率。甘雨等^[10-11]将坐标注意力(Coordinate attention, CA)机制引入 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,该模型在大规模害虫数据集 IP102^[12]中的识别准确率达到 69.45%,较改进前作物害虫识别准确率提高了 2.32 个百分点。Zhao 等^[13]首先提出结合 Inception 结构和残差结构来构建新的网络结构,然后将改进的块注意力模块(Convolutional block attention module, CBAM)引入网络,实现对玉米、马铃薯和番茄病害叶片的分类识别,3 种作物的整体识别准确率可以达到 99.55%。Zhao 等^[14]使用迁移学习结合注意力机制提出 DTL-SE-ResNet50 模型。该模型以 ResNet50 为基础,将 SENet 方法引入后提出 SE-ResNet50 网络模型。该模型先在 ImageNet 数据集上训练,得到一个初始权重,然后将初始模型在 AI Challenge 2018 公共数据集上进行再次训练获得新的权重,最后将经过上面 2 次训练后的模型在自建的蔬菜病害数据集上微调,形成蔬菜病害 DTL-SE-ResNet50 模型。

可见,注意力机制在农业领域的应用主要使用通道注意力机制方法,并取得了较好的效果。借鉴通道注意力机制的思想,本研究提出多维间注意力机制识别模型 Inter_3DriceNet 模型。本研究提出的模型不仅考虑通道注意力机制信息,同时也考虑图像高度注意力以及宽度注意力机制信息,以获得图像高度之间的关系以及宽度之间的关系。同时,在建立图像高度之间关系的过程中,可以同时建立图像通道和图像宽度之间的关系。同理,在建立图像宽度之间关系的时候,也同时建立图像通道和图像高度之间的关系。通过以上步骤,可以建立图像在通道、高度和宽度之间的多维间三维立体空间关系。

1 数据集

为了真实反映自然环境的实际病害数据,本研究构建了常见水稻病害数据集,所有图片均使用 Canon 相机拍摄于实际田间场景,拍摄图片分辨率从低到高分别为 1 440×1 080、1 423×1 920、3 008×2 000、4 288×2 848、4 928×3 264、6 016×4 000、7 360×4 912。考虑到水稻病害的危害性以及频发性,本研究主要选取影响水稻产量的以下 6 种病害做重点研究:水稻胡麻斑病、水稻白叶枯病、水稻纹

枯病、水稻细菌性条斑病、稻曲病、稻瘟病。6 种病害的部分示例如图 1 所示。

本研究获取的 6 种主要水稻病害数据样本量(6 938 张)如表 1 所示。

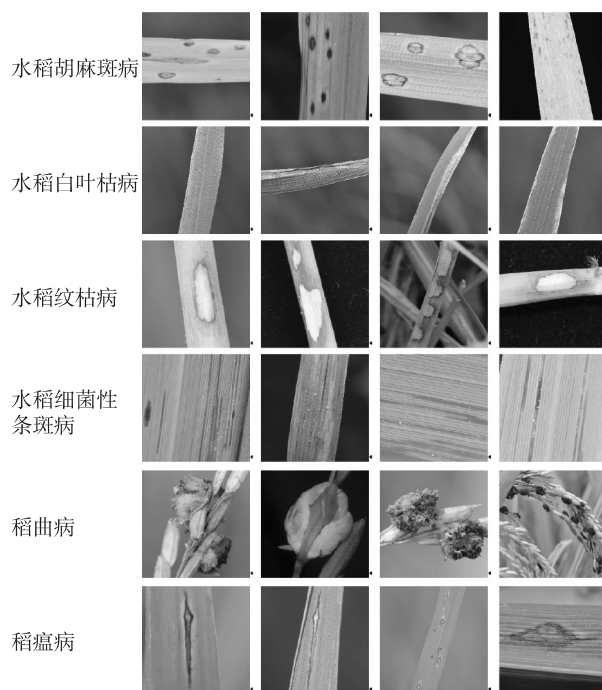


图 1 常见的 6 种水稻病害示例

Fig.1 Samples of six common rice diseases

表 1 6 种水稻病害样本量

Table 1 Sample size of six rice diseases

序号	病害名称	样本量(张)
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

2 多维间立体三维注意力机制

SENet 模型通过建立通道之间的关系获得了 2017 年 ILSVRC 竞赛图像分类子任务冠军,显示出建立通道之间关系的有效性以及优越性。借鉴通道之间关系构建的方法,本研究分别建立了高度(H)和宽度(W)之间的关系以获得更丰富的特征信息。同时,为了简化运算,本研究将以上建立的 3 种关系特征信息进行简单

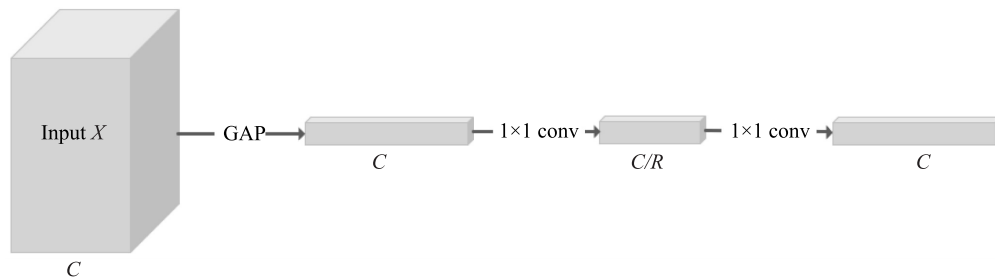
相加并取平均值作为最终的特征信息,最终构建出多维间的三维注意力机制水稻病害识别模型。

2.1 基于通道关系的立体注意力机制构建

2.1.1 一维通道关系的建立 卷积神经网络只能提取物体的空间特性,无法获得不同通道间之间的关系。为解决该问题,Hu 等^[6]提出的 SENet 模型给出了最初的解决方案。为了建立通道之间的关系,SENet 通过压缩和激发 2 个步骤获取通道间的关系。压缩指的是通过一定的方法获得输入特征或者原始输入图像每个通道的全局信息,SENet 模型使用简单的全局平均池化 (Global average pooling, GAP) 操作获得全局信息。激发指的是通过一定的机制获得通道之间的相互关系,SENet 模型通过设置 2 层全连接层实现:首先将输入特征输入第一个全连接层,并将输出维度设置为输入维度的 $1/N$ 倍

(N 是一个超参数),SENet 模型将 N 设置为 16;然后再连接一个全连接层,并且将输出维度升维为和第一层全连接层的输入维度相同。2 层全连接层构成的网络,实际上是含有一个隐藏层的多层感知机 (Multi-layer perceptron, MLP)。通过在全连接层设置不同的维度,强迫网络学习出不同层之间的权重信息,进而获得不同层之间的重要性,从而实现通道之间权重关系的构建。

本研究不使用全连接层,使用 1×1 的卷积层替换全连接层,通过对 1×1 卷积设置不同的输出维度来实现。具体过程为:将输入特征输入 1×1 的卷积层,并且将输出通道数设置为原来的 $1/R$,然后再通过一个 1×1 的卷积,将维度还原为初始的维度,从而构建通道之间的关系。图 2 表明了通道注意力机制的实现方法。



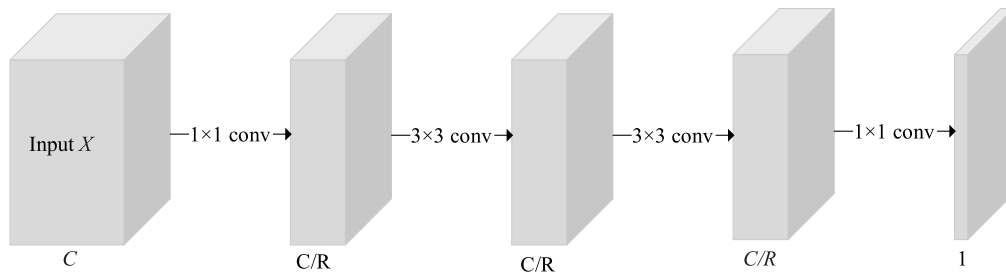
Input X 表示输入的图片;GAP 代表全局平均池化;(1×1 conv)表示卷积核大小是(1×1)的卷积; R 表示缩放因子; C 表示通道数。

图 2 一维通道注意力机制

Fig.2 Attention mechanism of one-dimensional channel

2.1.2 高度和宽度空间关系的构建 2 维空间高度和宽度之间关系的建立是通过堆叠不同的卷积层来实现的。具体实现流程如下:假设输入特征 $X \in R^{C \times H \times W}$,首先将输入特征 X 输入 1×1 的卷积,并且将输出特征降为原来的 $1/R$,然后再次经过 2 层

3×3 卷积,并且保持维度不变,最后输入 1×1 卷积层,并将维度降为 1。通过以上步骤,形成高度和宽度之间的二维注意力关系构建过程。图 3 展示了具体的实现过程。



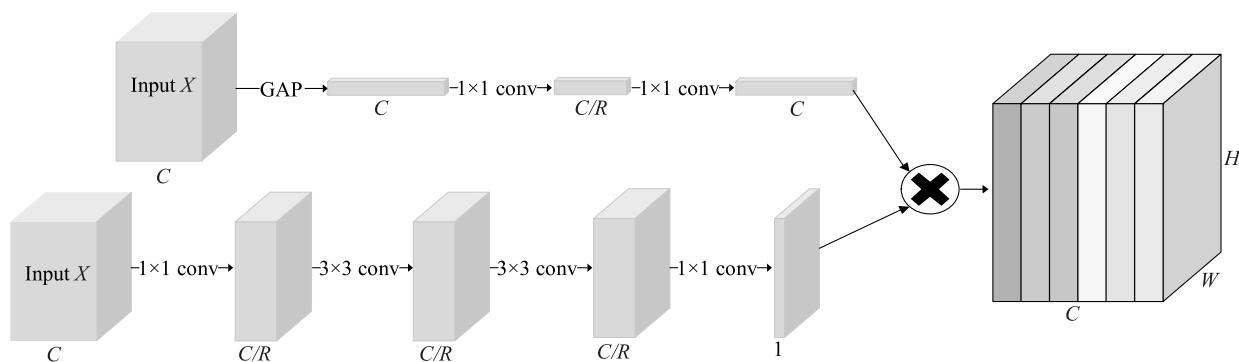
Input X 表示输入的图片;(1×1 conv)表示卷积核大小是(1×1)的卷积;(3×3 conv)表示卷积核大小为(3×3)的卷积; R 表示缩放因子; C 表示通道数。

图 3 图像高度和宽度二维关系

Fig.3 Two-dimensional relationship between image height and width

2.1.3 基于通道的三维注意力机制 基于通道间关系的三维注意力机制构建是将通道注意力乘以空

间注意力获得基于通道的立体注意力。具体的实现过程如图 4 所示。



Input X 表示输入的图片;GAP 代表全局平均池化;(1×1 conv)表示卷积核大小是(1×1)的卷积;(3×3 conv)表示卷积核大小为(3×3)的卷积; R 表示缩放因子; C 表示通道数; H 表示高度维度; W 表示宽度维度。

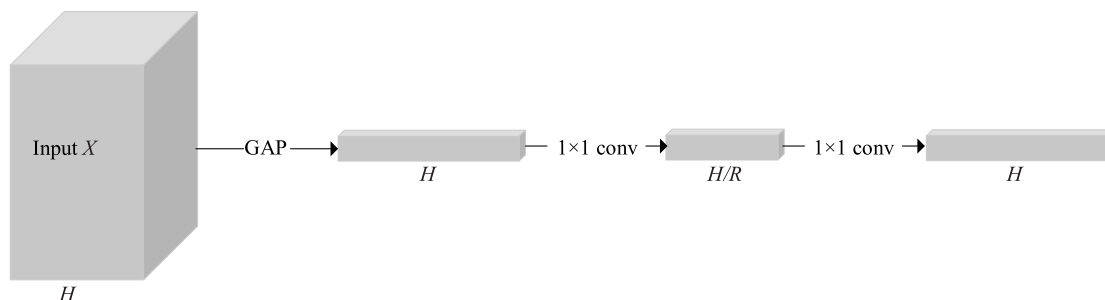
图 4 基于通道的三维注意力机制

Fig.4 Three-dimensional attention mechanism based on channel

2.2 基于高度关系的立体注意力机制构建

2.2.1 一维高度关系的建立 为了获得基于高度的立体注意力,首先需要建立输入图片高度之间的

关系。高度间关系的构建与通道关系的构建类似,将高度通过 2 层 1×1 卷积层,通过设置不同的输出维度来实现。具体实现流程如图 5 所示。



Input X 表示输入的图片;GAP 代表全局平均池化;(1×1 conv)表示卷积核大小是(1×1)的卷积; R 表示缩放因子; H 表示高度维度。

图 5 一维高度注意力机制

Fig.5 One-dimensional height attention mechanism

2.2.2 通道和宽度之间关系的构建 通道和宽度之间的关系构成,也是通过堆叠不同的卷积层来实现的。输入 X 的高度维度经过几层卷积操作可以建立通道和宽度之间的关系。具体的流程如图 6 所示。

2.2.3 基于高度的三维注意力机制 高度间关系的三维注意力机制构建是将高度注意力乘以由通道和宽度之间建立的注意力,以获得基于高度的立体注意力。具体的实现过程如图 7 所示。

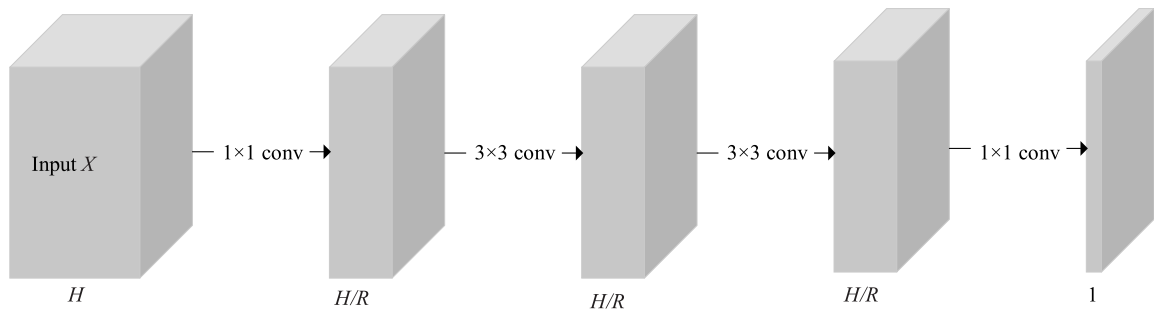
2.3 基于宽度关系的立体注意力机制构建

2.3.1 一维宽度关系的建立 为了获得基于宽度

的立体注意力,首先需要建立宽度之间的关系。宽度关系的构建也是采用和通道类似的方法。具体实现流程如图 8 所示。

2.3.2 通道和高度之间关系的构建 通道和高度之间关系的构建,是将宽度通过不同的卷积层实现,最终建立起通道和高度之间的二维注意力关系。具体的操作流程如图 9 所示。

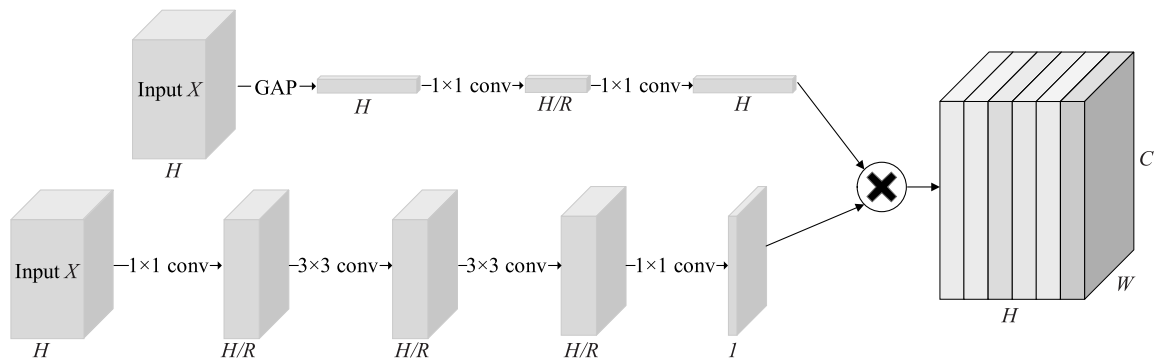
2.3.3 基于宽度的三维注意力机制 为了构建基于宽度关系的三维注意力机制,将宽度注意力乘以通道和高度之间的关系值,获得宽度的立体注意力。具体的实现过程如图 10 所示。



Input X 表示输入的图片; $(1 \times 1 \text{ conv})$ 表示卷积核大小是 (1×1) 的卷积; $(3 \times 3 \text{ conv})$ 表示卷积核大小为 (3×3) 的卷积; R 表示缩放因子; H 表示高度维度。

图 6 图像通道和宽度二维关系

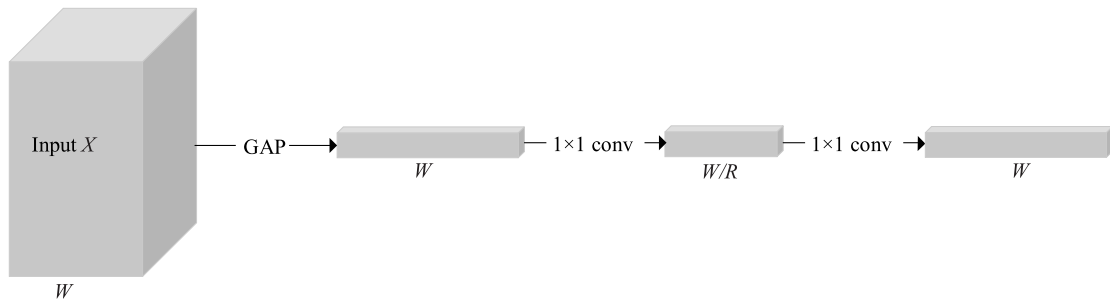
Fig.6 Two-dimensional relationship between image channel and width



Input X 表示输入的图片; GAP 代表全局平均池化; $(1 \times 1 \text{ conv})$ 表示卷积核大小是 (1×1) 的卷积; $(3 \times 3 \text{ conv})$ 表示卷积核大小为 (3×3) 的卷积; R 表示缩放因子; H 表示高度维度; W 表示宽度维度; C 表示通道数。

图 7 基于高度的三维注意力机制

Fig.7 Three-dimensional attention mechanism based on height



Input X 表输入的图片; GAP 代表全局平均池化; $(1 \times 1 \text{ conv})$ 表示卷积核大小是 (1×1) 的卷积; R 表示缩放因子; W 表示宽度维度。

图 8 一维宽度注意力机制

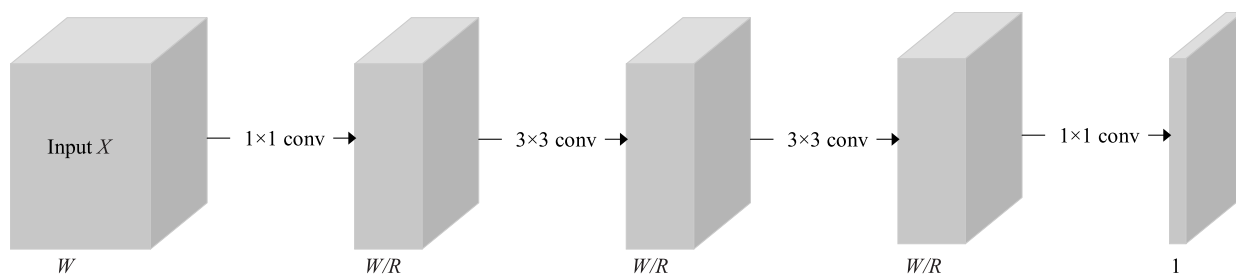
Fig.8 One-dimensional width attention mechanism

2.4 多维间立体注意力机制构建

为了获得以通道、高度以及宽度的多维间立体注意力,本研究将以上 3 种方法获得的注意力值进行平均后作为最终的三维注意力。即,将图 4、图 7 和图 10 得到的特征值相加后再取平均值作为最终的多维间立体注意力机制的值。

3 试验环境与数据处理

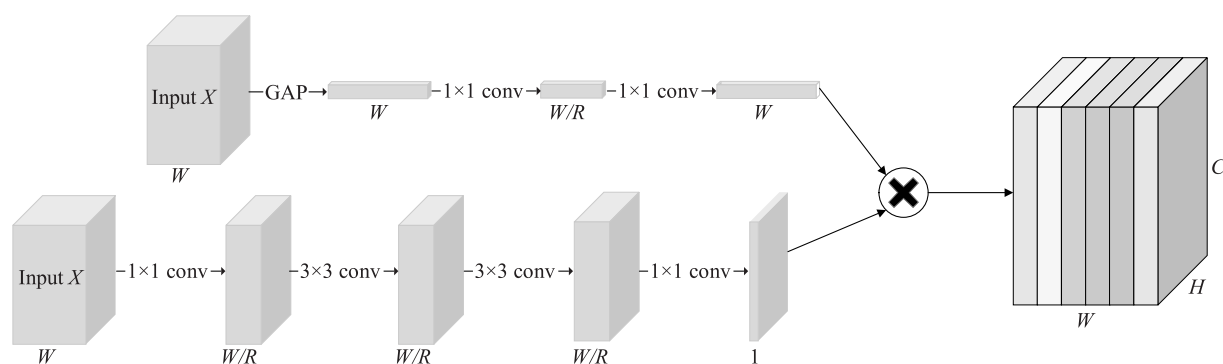
本试验使用 Pytorch 深度学习软件进行仿真验证。电脑配置中,中央处理器 (CPU): Intel i7-9700 @ 3.00 GHz, 图形处理器 (GPU): 英伟达 GeForce RTX2060, 显存大小为 6 GB。



Input X 表输入的图片; $(1 \times 1 \text{ conv})$ 表示卷积核大小是 (1×1) 的卷积; $(3 \times 3 \text{ conv})$ 表示卷积核大小为 (3×3) 的卷积; R 表示缩放因子; W 表示宽度维度。

图 9 图像通道和高度二维关系

Fig.9 Two-dimensional relationship between image channel and height



Input X 表输入的图片; GAP 代表全局平均池化; $(1 \times 1 \text{ conv})$ 表示卷积核大小是 (1×1) 的卷积; $(3 \times 3 \text{ conv})$ 表示卷积核大小为 (3×3) 的卷积; R 表示缩放因子; W 表示宽度维度; H 表示高度维度; C 表示通道数。

图 10 基于宽度的三维注意力机制

Fig.10 Three-dimensional attention mechanism based on width

3.1 试验参数设置

为了训练模型并验证模型的效果,将数据集按照 80% 和 20% 的比例划分为训练集和验证集。由于显存大小只有 6 GB,将批处理大小设置为 16。整个训练集一共迭代训练 100 次,即训练次数设置为 100。本试验使用随机梯度下降 (SGD) 优化器,并设置优化器学习率为 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。缩放因子 (R) 取值为 8。

3.2 数据预处理方法

为了保证训练的稳定性,对图片进行归一化处理。即将图片的 R、G、B 3 个通道值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 $(0.229, 0.224, 0.225)$ 。以上均值和标准差来自对 ImageNet 数据集中所有图片计算后得到的均值和标准差。

因为训练模型要求输入的图片尺寸一致,先将图

片的短边调整为 256 像素大小,然后从图片的中心选取 224×224 像素大小作为训练图片的输入尺寸。

4 结果与分析

为了保证算法比较的公平性,各算法使用相同的试验参数和相同的图像预处理方法。

4.1 不同模型性能对比试验

为了评估本研究 Inter_3DRiceNet 模型的性能,选取当前主流深度模型以及注意力机制网络模型进行对比分析。选取的经典网络模型: ResNet 模型^[15]、MobileNetV2 模型^[16]、DenseNet 模型^[17]、EfficientNet 模型^[18]。选取的注意力机制的模型: SENet 模型和 GCT 模型^[19],其中 SENet 和 GCT 模型是基于 ResNet50 网络结构进行试验,各算法都迭代 100 次,经过 100 次迭代之后在验证集上的识别准确率曲线如图 11 所示。

图 11 表明,表现最差的是 EfficientNet_B0 模

型,该模型不仅准确率最差,而且波动性也是最大的。表现最好的是本研究所提出的模型,算法准确率基本维持在 98% 附近,其次是 DenseNet 模型。而 2 个注意力机制模型 SENet 和 GCT 的准确率仅次于 DenseNet 模型。

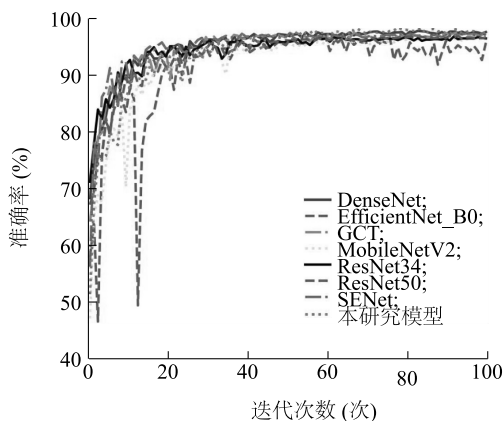


图 11 不同模型准确率对比曲线

Fig.11 Accuracy rate curves of different models

为了进一步分析各模型的准确率,统计了各模型的最高准确率以及对应的轮次。从表 2 可以看出,8 种模型的最高准确率都在 96.00% 以上,本研究提出的模型准确率最高(98.32%),比 ResNet50 模型提高了 1.43 个百分点,比 DenseNet 模型提高了 0.56 个百分点,比 SENet 模型提高了 0.92 个百分点。说明使用多维间三维注意力机制方法可以在一定程度上提高模型的准确率。

表 2 不同模型最高准确率及对应轮次

Table 2 The highest accuracy and epoch of different models

模型	最高准确率(%)	对应轮次
ResNet34	96.78	92
ResNet50	96.89	93
MobileNetV2	96.67	98
DenseNet	97.76	83
EfficientNet_B0	96.67	59
SENet	97.40	96
GCT	97.47	93
本研究(Inter_3DRiceNet)	98.32	77

4.2 6 种常见水稻病害识别结果定性分析

为了更详细地分析每种水稻病害的识别效果,分别计算 6 种水稻病害在测试集上的混淆矩阵值,

计算结果如图 12 所示。

水稻胡麻斑病在测试集上的数目是 291 张图片,其中预测正确的有 283 个,8 张病害图片预测错误,其中错误预测为水稻白叶枯病、水稻纹枯病、水稻细菌性条斑病的个数都是 1 张,有 2 张病害被预测为稻曲病,3 张水稻胡麻斑病的图片被预测为稻瘟病。由混淆矩阵可知,对角线上表示预测正确的数目都远远大于预测错误的数目,证明本研究算法在各类水稻病害上取得了比较好的识别效果。

预测标签	10000	10017	10018	10021	10046	10047
10000	283	1	0	3	1	3
10017	1	153	0	0	0	1
10018	1	2	101	2	1	1
10021	1	1	3	287	1	0
10046	2	3	3	3	140	0
10047	3	2	2	3	0	375
真实标签	10000	10017	10018	10021	10046	10047

图中横坐标数字代表每种水稻病害的识别 ID,不同 ID 对应的水稻病害名称如下,10000:水稻胡麻斑病;10017:水稻白叶枯病;10018:水稻纹枯病;10021:水稻细菌性条斑病;10046:稻曲病;10047:稻瘟病。纵坐标表示预测水稻病害的 ID。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。

图 12 6 种常见水稻病害混淆矩阵

Fig.12 Confusion matrix of six common rice diseases

4.3 6 种常见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,分别从精度、召回率、特异度、F1 得分以及平均准确率这 5 个指标进行定量分析。

4.3.1 定量分析评价指标 TP (True positive) 被称为真正类,表示模型正确地预测为正的样本数; FP (False positive) 被称为假正类,表示模型将样本错误地预测为正的样本数; FN (False negative) 被称为假负类,表示模型将样本错误地预测为负的样本数; TN (True negative) 被称为真负类,表示模型正确地预测为负的样本数。而精度、召回率、特异度等的计算来自于以上 4 个基本概念,为表示它们之间的关系,使用如下公式来说明。

精度(P)的公式为:

$$P = \frac{TP}{TP + FP} \times 100\% \quad (1)$$

精度表示被正确预测的样本数与被预测的样本总数的比率,值越大表示模型的预测能力越好。

特异度(S)的公式为:

$$S = \frac{TN}{FP+TN} \times 100\% \quad (2)$$

特异度表示预测为负样本中预测正确的样本除以真实标签是负样本的数量。

召回率(R)的公式为:

$$R = \frac{TP}{TP+FN} \times 100\% \quad (3)$$

召回率也称之为查全率,表示预测正确的正样本占有所有正样本的比例,值越大表示模型的预测能力越好。

$F1$ 得分定义为:

$$F1 = \frac{2 \times (P \times R)}{P + R} \times 100\% \quad (4)$$

$F1$ 得分是为了综合考虑精度和召回率这 2 个指标之间的关系,可以看作是精度和召回率之间的一种调和平均。因为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。 $F1$ 得分的核心思想在于尽可能地提高精度和召回率的同时,使两者之间的差异尽可能小。

准确率(Accuracy, ACC)的定义为:

$$ACC = \frac{TP+TN}{TP+TN+FP+FN} \times 100\% \quad (5)$$

ACC 表示预测对的样本数占总样本数的比例。但是在正样本、负样本数量不均衡的情况下,准确率指标有一定的缺陷。

4.3.2 定量分析试验结果 根据以上定义,分别计算出每种水稻病害在测试集上的精度、召回率、特异度、 $F1$ 得分以及平均准确率(表 3)。从表 3 可知,精度最低的是稻曲病,只有 92.7%,其次是水稻纹枯病。召回率最低的是水稻纹枯病,只有 92.7%。特异度的值各病害差别不大。 $F1$ 得分最低的是水稻纹枯病,其次是稻曲病。从以上分析可知,该模型对水稻纹枯病和稻曲病的识别效果不太好。

通过分析水稻纹枯病数据集可知,水稻纹枯病数据集中的图片大部分都是横向长条状或者纵向长条状。很多图片的横纵比为 6:1、7:1、8:1,较高的横纵比加大了模型识别的难度。部分横向长条状样本示例图片如图 13 所示。

表 3 6 种常见水稻病害的 5 个指标值

Table 3 Five index values of six common rice diseases

病害名称	精度 (%)	召回率 (%)	特异度 (%)	$F1$ 得分 (%)
水稻胡麻斑病	97.3	97.3	99.3	97.3
水稻白叶枯病	98.7	94.4	99.8	96.5
水稻纹枯病	93.5	92.7	99.5	93.1
水稻细菌性条斑病	98.0	96.3	99.4	97.1
稻曲病	92.7	97.9	99.1	95.2
稻瘟病	97.4	98.7	99.0	98.0

平均准确率为 96.82%。

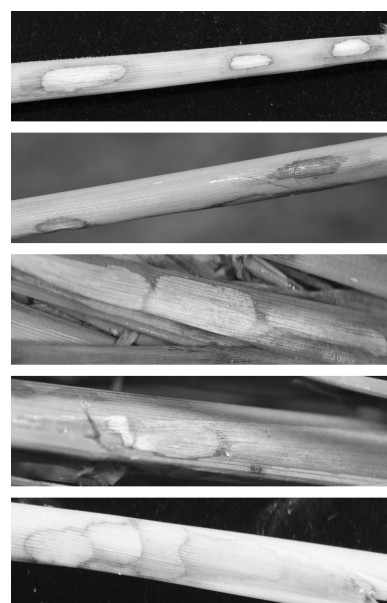


图 13 水稻纹枯病示例图片

Fig.13 Example pictures of rice sheath blight

通过对稻曲病数据集分析可知,拍摄的稻曲病害表现差异较大,病害的颜色有黄色、黑色,病害的范围有整株稻穗或几个稻穗。从图 14 可知,稻曲病图像差异较大,而且该数据集的样本总数只有 716 张,进一步加大了模型提取特征的难度。

4.3.3 类激活图可视化分析 为了直观地展示本研究算法的性能,使用 Grad-CAM 方法^[20]对学习到的特征进行可视化展示。以水稻纹枯病为展示样例,图 15 分别展示了网络模型从层 1 到层 4 学习到的图像特征信息。由图 15 可知,经过网络的逐层学习后,学习出的特征区域范围逐渐增加。比如层 1 主要学习较小区域的边缘信息,层 2 不仅学习到层 1 已有的边缘信息,还学习出其他区域

的边缘信息,并且学习的区域变大了。层 3 可以学习主要区域的边缘信息,即图中中心区域的边缘信息。这种学习方式符合人类观察物体的特点。人类在观察物体时,会重点关注物体主要区域的显著特征。经过层 4 学习后,网络获得了更大区域的显著特征。这样可以在更高的视角理解图像,可以对图像有一个全局的认识,从而避免陷入局部区域的细节之中。图 15 不仅说明模型可以逐层学习更大范围的图像特征,而且说明在模型学习各层特征的过程中,对同一层特征学习并不是平等对待的,而是学习层内重点区域的特征,即将注意力机制的学习也贯穿到各层特征的学习过程之中。比如层 1 重点学习了图像左上和右下区域的边缘特征信息,层 3 重点学习了中间区域的更长范围内的边缘信息。

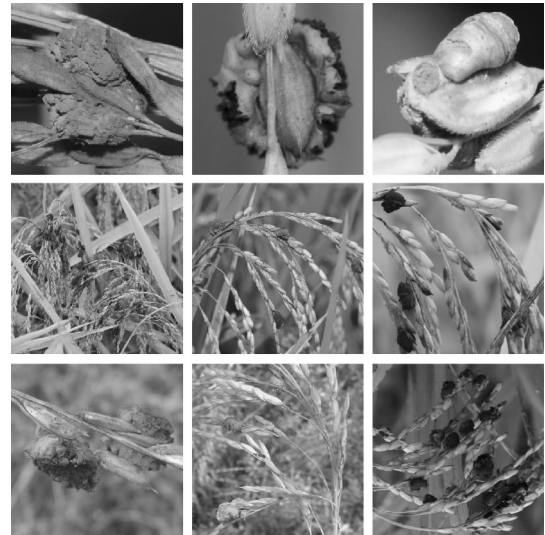
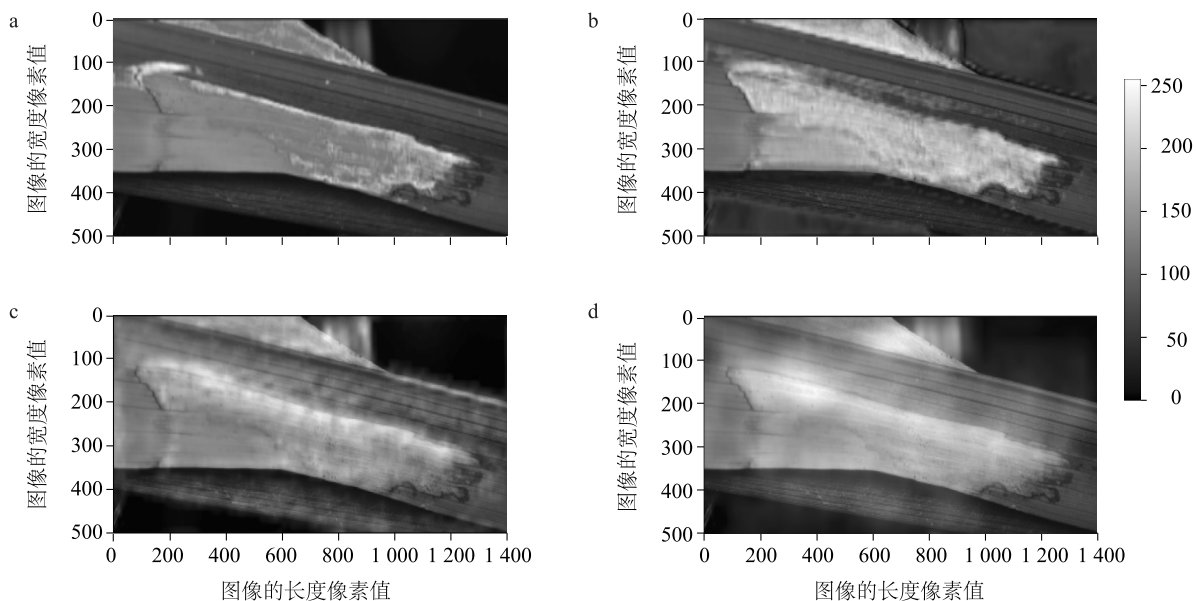


图 14 水稻稻曲病示例图片

Fig.14 Example pictures of rice false smut



a;第1层;b:第2层;c:第3层;d:第4层。横坐标表示图像的长度像素值,纵坐标表示图像的宽度像素值。色度条值的范围是(0,255),表示激活值大小从0到255。数字越大,表示该区域被激活得越显著,对应到激活图上的区域越明亮。

图 15 水稻纹枯病类激活图可视化

Fig.15 Class activation map of rice sheath blight

5 结论

为提高自然环境中常见水稻病害的识别率,从而为精准防治水稻病害提供参考进而提升水稻产量和质量,本研究提出一种多维间关系的三维注意力机制 Inter_3DRiceNet 模型。为了真实反映自然环境情况,本研究收集自然环境下的 6 种常见水稻病

害共计6 938张图片。借鉴通道注意力机制优点,本研究不仅构建了通道注意力机制,同时构建了高度注意力机制以及宽度注意力机制,最终获得多维间的三维注意力机制模型。结果表明,本研究模型在 6 种常见水稻病害数据集上取得了较好的识别准确率,优于 ResNet、MobileNet、DenseNet、EfficientNet 等经典神经网络模型,并优于 SENet 和 GCT 等注意力

机制模型。本研究的模型展示了其在识别常见水稻病害中的有效性。

参考文献:

- [1] 佚名. 2022 年全国农作物重大病虫害发生趋势预报[J]. 中国植保导刊, 2022, 42(4): 107-108.
- [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] 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.
- [4] 鲍文霞, 吴德钊, 胡根生, 等. 基于轻量型残差网络的自然场景水稻害虫识别[J]. 农业工程学报, 2021, 37(16): 145-152.
- [5] 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.
- [6] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]. Piscataway: IEEE, 2018: 7132-7141.
- [7] 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.
- [8] WANG Q, WU B, ZHU P, et al. ECA-Net: efficient channel attention for deep convolutional neural networks[C]. Piscataway: IEEE, 2020: 11531-11539.
- [9] 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.
- [10] 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.
- [11] HOU Q, ZHOU D, FENG J. Coordinate attention for efficient mobile network design[C]. Piscataway: IEEE, 2021.
- [12] WU X, ZHAN C, LAI Y K, et al. IP102: a large-scale benchmark dataset for insect pest recognition[C]. Piscataway: IEEE, 2019.
- [13] ZHAO Y, SUN C, 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.
- [14] ZHAO X, LI K, LI Y, et al. Identification method of vegetable diseases based on transfer learning and attention mechanism[J]. Computers and Electronics in Agriculture, 2022, 193: 106703.
- [15] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]. Piscataway: IEEE, 2016: 770-778.
- [16] SANDLER M, HOWARD A, ZHU M, et al. Mobilenetv2: inverted residuals and linear bottlenecks[C]. Piscataway: IEEE, 2018: 4510-4520.
- [17] HUANG G, LIU Z, VAN DER MAATEN L, et al. Densely connected convolutional networks[C]. Piscataway: IEEE, 2017.
- [18] TAN M, LE Q. Efficientnet: rethinking model scaling for convolutional neural networks[C]. Madison: ACM, 2019.
- [19] YANG Z, ZHU L, WU Y, et al. Gated channel transformation for visual recognition[C]. Piscataway: IEEE, 2020.
- [20] SELVARAJU R R, COGSWELL M, DAS A, et al. Grad-cam: visual explanations from deep networks via gradient-based localization[C]. Piscataway: IEEE, 2017.

(责任编辑:陈海霞)