doi:10.6041/j.issn.1000-1298.2024.05.024

基于选择性注意力神经网络的木薯叶病害检测算法

张家瑜　朱　锐　邱　威　陈坤杰
(南京农业大学工学院, 南京 210031)

摘要: 为了实现在复杂非结构环境下对木薯叶 4 种主要病害的高精度检测,提出一种基于选择性注意力机制的木薯叶病害神经网络检测改进算法 MAISNet (Multiattention IBN Squareplus neural network)。以 V2 - ResNet - 101 为基础网络,先使用多重注意力算法优化加权系数,调整特征通道的语义表达,在特征图中初步构建显著性特征;然后在残差单元之后采用实例批归一化方法来抑制特征表达中的协变量偏移,在特征图中构建出显著性语义特征,实现高质量语义特征表达;最后在残差分支中采用 Squareplus 激活函数替代 ReLU 激活函数,保持语义特征在负数域的数值分布,减少特征拟合过程中的截断误差。对比试验结果显示,经过上述改进后构建出的 MAISNet - 101 神经网络,对 4 种常见木薯叶病害检测的平均准确率达到 95.39%,明显优于目前主流算法 EfficientNet - B5 和 RepVGG - B3g4 等。网络提取特征的可视化分析结果表明,高质量木薯叶病害显著性语义特征,是提高木薯叶病害检测准确率的关键。所提出的 MAISNet 神经网络模型可以完成实际场景下木薯叶病害高精度检测。

关键词: 木薯; 病害检测; 多重注意力算法; 显著性语义特征; Squareplus 激活函数
中图分类号: TP391.4　　**文献标识码**: A　　**文章编号**: 1000-1298(2024)05-0254-09　　**OSID**:

Cassava Leaf Disease Detection Algorithm Based on Selective Attention Neural Network

ZHANG Jiayu　ZHU Rui　QIU Wei　CHEN Kunjie
(*College of Engineering*, *Nanjing Agricultural University*, *Nanjing* 210031, *China*)

Abstract: To achieve high-precision detection of four major cassava leaf diseases in complex unstructured environments, an improved algorithm for cassava leaf disease neural network detection based on the selective attention mechanism, MAISNet, was proposed. Using V2 - ResNet - 101 as the base network, the multiattention algorithm was firstly used to optimize the weighting coefficients, adjust the semantic expression of the feature channels, and the semantic feature saliency expression of cassava leaf disease in the feature map was preliminary constructed; then the instance batch normalization method was used after the residual unit to suppress the covariate offset in the feature expression, highlight the target semantic feature expression in the feature map, and realize the high-quality semantic feature expression. Finally, the Squareplus activation function was used to replace the ReLU activation function in the residual branch to maintain the numerical distribution of semantic features in the negative domain, and reduce the truncation errors in the feature fitting process. The results of the comparison test showed that the MAISNet - 101 neural network constructed after the above improvement achieved an average accuracy of 95.39% for the detection of four common cassava leaf diseases, which was significantly better than the performance of the mainstream algorithms such as EfficientNet - B5 and RepVGG - B3g4. The results of the visualization and analysis of the extracted features of the network showed that high-quality semantic feature saliency representation of cassava leaf diseases was the key to improve the accuracy of cassava leaf disease detection. The proposed MAISNet neural network model can accomplish high-precision detection of cassava leaf diseases in real scenarios, which can provide technical support for precise drug application.

Key words: cassava; disease detection; multiattention algorithm; saliency semantic feature; Squareplus activation function

收稿日期: 2024 - 02 - 18　修回日期: 2024 - 03 - 12
基金项目: 江苏省农业科技自主创新资金项目(CX(20)3172)
作者简介: 张家瑜(1993—),男,博士生,主要从事农业信息检测与控制工程研究,E-mail: 2020212020@stu.njau.edu.cn
通信作者: 陈坤杰(1963—),男,教授,博士生导师,主要从事农业信息检测与控制工程研究,E-mail: kunjiechen@njau.edu.cn

0　引言

木薯是世界上种植最广泛的作物之一，种植区域涵盖非洲（55.5%）、亚洲（30.2%）、美洲（14.3%）以及大洋洲（0.1%）等地[1]，不仅是人类食物的主要来源，还是饲料、淀粉及乙醇燃料的重要生产原料。然而，病害严重威胁了木薯生长和产量，根据 LEGG 等[2]的研究，木薯病害有 30 多种。其中，细菌性枯萎病、木薯棕色条纹、木薯花叶病和木薯绿斑驳病是最主要的病害，是影响木薯品质、导致木薯产量下降的主要原因。因此，开展木薯叶病快速检测及识别技术研究，对木薯叶病防治、提高木薯产量和品质都具有重要的现实意义。

目前，作物病害的检测识别主要有人工检测[3]、光谱检测[4]和图像检测[5]3 种方法。人工检测方法不仅需要专门的知识和经验，而且检查过程耗时耗力，难以适应精准农业对作物精准施药的高效需求。光谱技术虽然有效，但光谱设备成本较高，导致终端设备售价高昂，难以大规模推广使用。相比之下，基于图像的检测算法具有较低的成本和较快的检测速度，成为作物病害检测技术的发展方向。基于图像的机器学习算法[6]在田间复杂非结构环境下，对作物病害检测识别的适应性较差。因此，开发基于深度学习算法[7]的作物病害检测算法，正逐渐引起人们的关注。

在针对木薯叶病害检测的任务中，为了提高卷积神经网络模型的性能，多使用迁移学习对训练结果进行优化[7]。然而，当训练数据中包含的木薯叶病害特征不完整时，迁移学习方法所取得结果的可靠性不高，且受到神经网络模型特征提取和泛化能力的影响较大。因此，为了进一步提高卷积神经网络模型的特征提取能力，DHIVYAA 等[8]使用 Bi - LSTM 算法从高级特征中提取信息，结合低级特征信息建立木薯叶病害特征长期依赖。在卷积神经网络中，使用卷积神经网络的历史特征中的高质量特征，构建出关键特征组件与语义特征之间的特征长期依赖。为了提高特征信息表达能力，MI 等[9]在 DenseNet 中使用 CBAM 算法，基于 DenseNet 特征最大化基数结构逐通道和逐元素调整特征表达，构建高质量特征长期依赖表达，得到用于田间小麦铁锈病检测的 C - DenseNet 神经网络模型。

语义特征的表达质量决定了神经网络性能[10]。为了实现视觉注意力对目标语义区域的搜索和表达，TSOTSOS 等[10]提出赢者通吃（Winner - take - all，WTA）网络作为金字塔神经网络的门控网络，筛选并提取高质量语义信息，在神经网络模型中实现自下而上的注意力表达。ZHANG 等[11]基于文献[10]的研究思路，提出了 Excitation Backprop 反向传播方法，实现概率化 WTA 方法抑制神经元输出，实现自上而下的注意力表达。LIU 等[12]通过修改分类损失函数，优化特征提取算子的特征提取质量，完成对 PDD271 田间植物疾病数据集的弱监督粗粒度标签分类任务[13]。但是，上述研究并没有对复杂非结构环境下，植物疾病语义特征表达方式进行探索。虽然 LEE 等[14]发现在植物分类任务中，相较于叶片的外轮廓形状特征，植物叶脉特征成为卷积神经网络中更具代表性特征，但是，在复杂非结构环境下，样本图像并非聚焦于单一的植物叶片，在图像视场较大的情况下，叶脉特征和叶片外轮廓形状特征很难在卷积神经网络中呈现。

本文提出一种用于木薯叶病害检测的神经网络改进算法 MAISNet。该网络基于 V2 - ResNet - 101 模型，使用多种注意力算法[15]、实例批归一化（Instance batch normalization，IBN）[16]和 Squareplus 激活函数[17]进行改进。通过提高神经网络中语义特征质量，以提高模型对木薯叶病害预测精度，实现对 4 种主要木薯叶病害高精度检测。

1　材料和方法

1.1　木薯叶病害数据集

采用 ZHANG 等[15]使用的木薯叶病害原始数据集进行神经网络模型训练。该原始数据集包含在复杂非结构环境下（光照变化、遮挡、重叠等）拍摄的 21 393 幅木薯叶病害样本图像。在原始的木薯叶病害数据集中，各类别样本图像的分布并不均衡，其中木薯花叶病有 13 158 幅样本图像，而木薯细菌性枯萎病只有 1 086 幅图像。不平衡的样本数据分布导致神经网络在训练时出现不同程度的长尾现象。为了克服特征的长尾障碍，采用数据增强方法来保持训练数据集的类别数据均匀分布。

通过对原始的木薯数据集进行分析，发现 3 个问题：①低质量图像无法清晰呈现木薯叶病害区域，导致神经网络难以在图像中清晰定位疾病病灶。②原始木薯叶病害数据集中不仅包括木薯叶病害样本图像，还包括木薯果肉病害图像、杂志封面和其它不相关内容。③虚焦的木薯叶病害样本图像无法高质量呈现木薯叶病害纹理特征，难以提供高质量病害特征。

针对上述问题，首先对原始的木薯叶病害数据集进行数据清洗，然后使用清洗后的木薯叶病害数据进行数据增强。为了保持类别间的平衡，采用高斯噪声、水平翻转、挖剪和垂直翻转等方法

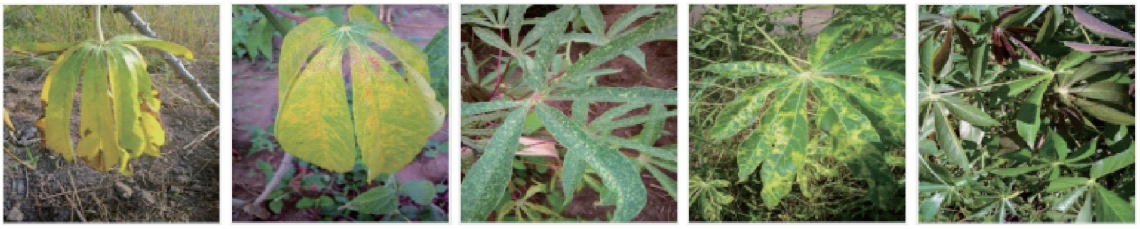
来进行增强。

最终,本文所使用数据集包含木薯细菌性枯萎病(Cassava bacterial blight, CBB)、木薯棕色条纹病(Cassava brown streak disease, CBSD)、木薯花叶病(Cassava mosaic disease, CMD)、木薯绿斑驳病(Cassava green mottle, CGM)和健康木薯叶 5 个类别数据,共 20 000 幅分辨率为 448 像素 ×448 像素 ×3 通道的 RGB 彩色图像,数据集详情如表 1 所示,4 种木薯叶病害示例如图 1 所示。

表 1 木薯叶病害数据集

Tab. 1 Dataset analysis 幅

叶片种类	原始	高斯噪声	水平翻转	垂直翻转	挖剪	总计
CBB	986	985	986	986	57	4 000
CBSD	1 772	0	1 772	456	0	4 000
CMD	1 861	28	1 861	250	0	4 000
CGM	4 000	0	0	0	0	4 000
健康	1 054	838	1 054	1 054	0	4 000



(a) CBB (b) CBSD (c) CGM (d) CMD (e) 健康

图 1 木薯病害示例

Fig. 1 Cassava disease examples

1.2 MAISNet

有关研究显示,经典的神经网络 ResNet 在植物病害检测方面具有优异的性能[18]。从硬件和算法计算成本考虑,ResNet - 101 相较 ResNet - 152 或其它 ResNet 神经网络模型结构更加简洁,实时性更好,更适合作物病害的实时检测。但由于 ResNet - 101 中的 1 ×1 下采样卷积难以取得足够优秀的特征量化结果,因此,本文以采用 3 ×3 下采样卷积的 V1.5 - ResNet - 101[19] 为基础,将 V1.5 - ResNet - 101 中的全连接层替换为全局平均池化层,构建出以 V2 - ResNet - 101 神经网络模型作为基准算法的 MAISNet - 101 神经网络模型。模型结构如图 2 所示,MAISNet - 101 神经网络模型由 1 个头部模块和 4 个 MAIS(Multiattention IBN Squareplus)模块构成。

头部模块主要用于捕捉轮廓信息,主要由 1 个卷积核尺寸为 7 ×7、步长为 2 的二维卷积,1 个批归一化层,1 个 ReLU 激活函数和 1 个最大池化层组成,其结构细节见图 2。

1.3 MAIS 模块

在 ZHANG 等[15]的研究中,针对卷积神经网络在木薯叶病害检测任务中的极弱信号问题进行了探索和修复。但是没有探索和分析神经网络模型性能

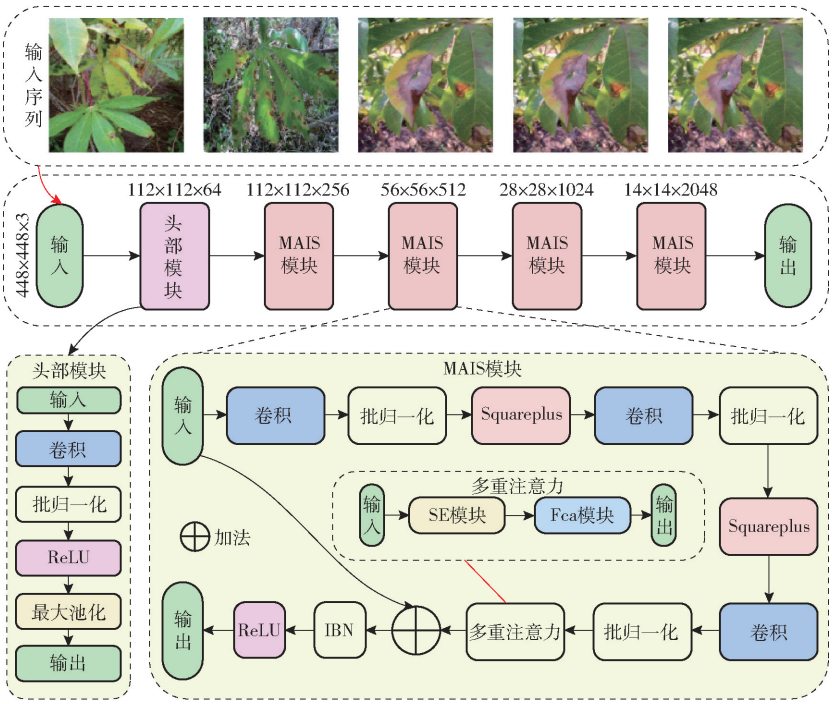


图 2 MAISNet 神经网络模型结构图

Fig. 2 Architecture of MAISNet

与显著性语义特征质量之间的关系。因此，本文提出一种改进的显著性语义特征构建算法，通过多重注意力算法、实例批归一化方法[16]，构建显著性语义特征。然后基于显著性语义特征，使用 Squareplus 激活函数替换残差分支中所使用的 ReLU 激活函数，减少特征拟合过程中的截断误差，进一步提高木薯叶病害检测准确率。最终使用多重注意力算法、实例批归一化方法和 Squareplus 激活函数组成 MAIS 模块。

1.3.1　显著性语义特征

弱监督显著性物体检测方法往往由语义引导，利用类别信息对物体进行粗略聚焦与定位[20]。因此，在弱监督学习的不确切监督学习任务中，高质量的显著性语义特征成为了神经网络模型性能提升的关键。

传统机器学习方法基于不同尺度、不同维度的特征，使用特征加权方法和特征修正方法，构建显著性语义特征[21]。具体步骤为：①特征值加权，对不同尺度、不同维度的多特征线性融合[22]和特征值加权[23-24]，初步构建显著性特征。②特征修正，使用归一化方法以及线性或非线性组合后，调整特征图中的特征值分布，在特征图中构建显著性语义特征。

传统显著性语义特征构建，主要分为特征提取、特征融合和特征修正 3 个步骤[20]。MAIS 模块通过对特征图进行特征值加权和特征修正，构建显著性语义特征，获取木薯叶病害叶片的粗略聚焦与定位结果。

1.3.2　多重注意力算法加权系数优化数学表达

注意力算法主要包括通道注意力算法、空间注意力算法、分支注意力算法、时间注意力算法和时间与空间复合注意力算法[25]。多重注意力算法通过建立映射变换，将输入数据 $\boldsymbol{X}\in\mathbf{R}^{H'\times W'\times C'}$ 映射为输出数据 $\boldsymbol{U}\in\mathbf{R}^{H\times W\times C}$，其中 H、W、C 表示特征图的纵向跨度、横向跨度和特征通道数。如果将 SE 模块[26]输出记作 $\boldsymbol{U}=[u_1,u_2,\cdots,u_c]$，Fca 模块[27]输出记作 $\boldsymbol{K}=[k_1,k_2,\cdots,k_c]$，多重注意力算法的数学表达式为

$$\boldsymbol{U}_c=\boldsymbol{V}_c*\boldsymbol{X}=\sum_{s=1}^{C'}v_c^s*\boldsymbol{X}^s \tag{1}$$

$$\boldsymbol{K}_c=\boldsymbol{T}_c*\boldsymbol{U}=\sum_{s=1}^{C'}\tau_c^s*\boldsymbol{U}^s \tag{2}$$

其中 $\boldsymbol{V}_c=[v_c^1,v_c^2,\cdots,v_c^{C'}]$　$\boldsymbol{T}_c=[t_c^1,t_c^2,\cdots,t_c^{C'}]$

$$\boldsymbol{X}=[x^1,x^2,\cdots,x^{C'}]$$

式中　$*$——卷积运算符号

v_c^s、τ_c^s——二维空间卷积核

$\boldsymbol{V}_c$——单个特征通道，作用于 $\boldsymbol{X}$ 相应通道

$\boldsymbol{T}_c$——单个特征通道，作用于 $\boldsymbol{U}$ 相应通道

$\boldsymbol{X}^s$——多重注意力算法输入特征张量

$\boldsymbol{U}_c$——多重注意力算法中 SE 模块输出特征张量

$\boldsymbol{U}^s$——多重注意力算法中 Fca 模块输入特征张量

$\boldsymbol{K}_c$——多重注意力算法输出特征张量

s——特征通道索引

为简化符号，省略偏差项得

$$\boldsymbol{K}_c=\sum_{s=1}^{C'}v_c^s*\tau_c^s*\boldsymbol{U}^s=\sum_{s=1}^{C'}g_c^s*\boldsymbol{X}^s \tag{3}$$

式中　g_c^s——二维空间卷积核

基于式(3)可知，多重注意力算法中的 Fca 注意力算法所产生的加权系数，调整了 SE 注意力算法所产生的加权系数，最终得到新的加权系数 G_c。通过加权系数 G_c 对特征通道进行精准加权，调整特征图中的特征值表达，优化下一步的特征拟合和特征提取结果。

1.3.3　神经网络特征显著性增强

深度卷积神经网络中神经元(卷积核)对某个模式所响应的区域多呈现“稀疏”特性[28]，导致每幅特征图只包含“语义概念”的局部特征。因此，为了在卷积神经网络中实现局部语义特征修正，使用实例批归一化方法调整特征图中的特征值分布。

归一化方法的平滑计算(Smoothing calculation)具备高效、简洁和稳定的特点，可以降低数字信号的突变。在神经网络中，优秀的归一化方法可以发挥关键作用，凸显相关特征的有效表达。其中，实例批归一化[16](IBN)在消除特征表达协变量偏移问题方面具有出色性能。实例批归一化方法由实例归一化(Instance normalization，IN)和批归一化(Batch normalization，BN)2 个方法组成。其中，实例批归一化方法(IBN)中的实例归一化方法通过逐通道修正特征图中的特征分布，优化特征图中的特征表达，减少了特征表达中的协变量偏移，维持了显著性语义特征低频组件的质量[29]。为此，在 MAISNet-101 神经网络模型中，采用实例归一化方法减少特征表达中的协变量偏移，修正特征图中的局部语义特征，以此来构建特征图中的显著性语义特征。

1.3.4　负数特征信号保持

通过注意力方法，可以对神经网络通道中的显著性特征进行选择。然而，在神经网络通道中，部分有意义的信号被表示为负数，并被 ReLU 激活函数所破坏。此外，ReLU 还存在零梯度和偏置转移(Bias shift)问题。与 ReLU 相比，Softplus 激活函数在$(-\infty,\infty)$范围内皆可微分，且在其定义域范围

内也皆可进行梯度反向传播,使得激活后的函数表达更接近真实的生物学特性。Squareplus 激活函数具有良好的代数表达能力,能够更加准确地呈现神经网络中的一阶导数和二阶导数[17]。为此,选择在残差分支中使用 Squareplus 激活函数来保持神经网络特征通道中有意义的负数特征信号,减少特征拟合过程中的截断误差。

然而,由于 Squareplus 激活函数并不能总是保证神经网络会收敛,尤其是在实例批归一化方法之后使用 Squareplus 激活函数时,神经网络将无法收敛。因此,在 MAIS 模块的输出端采用 ReLU 激活函数。

1.4 训练方法及参数确定

MAISNet-101 神经网络模型在木薯数据集上使用以下设置进行训练:使用 SGD 优化器,初始学习率为 0.2,每个周期衰减 0.96,动量为 0.9;批量归一化动量为 0.9,权重衰减为 1×10^{-5}。在 MAISNet-101 神经网络模型的特征提取算子中,使用 L2 正则化作为优化方法,其中在恒等分支的特征提取算子中 L2 正则化系数设置为 1×10^{-5},在残差分支的特征提取算子中 L2 正则化系数设置为 1×10^{-4}。MAISNet-101 神经网络的批尺寸(Batch size)设置为 12,训练轮次设置为 80。在 SE-block 中,用 Hard-Sigmoid 代替 Sigmoid 为激活函数,以减少计算成本。训练过程中,使用分类交叉熵损失函数作为训练损失函数。模型验证通常选取数据集中 70% 的样本数据作为训练数据。在剩下的 30% 数据集样本数据中,验证数据和测试数据各选取 15% 的样本数据。因此采用 7 折交叉验证,每折交叉验证保留 14.28% 的数据集样本用于数据验证。

由于 TensorFlow 框架无法直接释放交叉验证每折训练后的 GPU 显存资源。因此,使用 Python 编程语言的 multiprocessing.Process 函数调用 Scikit-Learn 框架的 StratifiedKFold 函数实现交叉验证训练时的数据随机抽取,控制交叉验证迭代训练,释放每折训练后的显存空间。

1.5 模型评价指标

神经网络性能指标基于真阳性(True positive, TP)、假阳性(False positive, FP)、真阴性(True negative, TN)、假阴性(False negative, FN)这 4 个细分指标进行计算得出,主要有:准确率(Accuracy)、召回率(Recall)、精确率(Precision)和 F1 值(F1-score)。

2 结果与分析

2.1 不同数据抽取策略对模型性能的影响

机器学习 Scikit-Learn 框架中,K 折交叉验证训练函数 StratifiedKFold 的数据抽取策略通过设置随机状态参数(Random status)决定。随机状态参数设置为 0 时,StratifiedKFold 函数中的“shuffle”(混洗)参数为 False,数据样本不会被随机混洗(Random shuffle),直接对数据集进行分层抽取,抽取后的数据在 Keras 的 Generator 函数中进行混洗。当随机状态参数设置大于 0 时,StratifiedKFold 函数先对样本数据进行混洗,然后对数据样本进行分层抽取。

不同的数据抽取策略所抽取的试验数据呈现出不同的数据域表达。为了选择合适的数据抽取策略训练 V2-ResNet-101 神经网络模型,以清晰且直观地呈现对比试验结果,分别使用随机状态参数为 0 和 834 239 进行 7 折交叉验证训练。其中随机状态参数 834 239 是经过随机选取的较大数值,以确保随机索引排序不会因为参数过于规律而受到影响。采用 2 种随机状态参数进行的检测结果如表 2 所示。

表 2 基于 7 折交叉验证的 V2-ResNet-101 在不同随机状态参数下的性能比较

Tab.2 Comparison of V2-ResNet-101 with different random state parameters using 7-fold cross-validation

%

随机状态参数	准确率	召回率	精确率	F1 值
834 239	77.80	77.61	78.17	77.89
0	86.90	86.80	87.02	86.92

如表 2 所示,V2-ResNet-101 神经网络模型在随机状态参数为 834 239 时,比使用随机状态参数为 0 时的准确率低 9.1 个百分点,说明随机状态参数的选取对模型检测性能有重要影响。因此,以下对比试验中,均将随机状态参数设置为 0,以获取最好的试验效果。

2.2 不同注意力算法对模型性能的影响

为了全面比较多重注意力算法与单通道注意力算法在木薯叶病害检测中的性能,以 V2-ResNet-101[19] 为基础算法,分别采用 SE[26]、Fca[27] 和 Double-SE 3 个常用单通道注意力模块,构建出 SENet、FcaNet 和 Double-SENet 3 个检测模型,与使用多重注意力算法(Multiattention, MA)的 MANet(Multiattention neural network)模型进行对比消融试验,结果如表 3 所示。

由表 3 可知,在 V2-ResNet-101 神经网络中使用单通道注意力方法后,模型性能都有所下降。而在使用多重注意力算法之后,V2-ResNet-101 的准确率从 86.90% 提高到 90.12%,准确率提高

表 3　基于 7 折交叉验证的注意力算法消融试验结果

Tab. 3　Ablation experiment result of different attention mechanisms based on 7 - fold cross-validation　%

模型	准确率	召回率	精确率	F1 值
V2 - ResNet - 101	86. 90	86. 80	87. 02	86. 92
SENet	83. 98	83. 79	84. 16	83. 98
FcaNet	84. 38	84. 20	84. 57	84. 38
Double - SENet	89. 00	88. 93	89. 07	89. 00
MANet	90. 12	90. 00	90. 28	90. 13

3. 22 个百分点，性能显著提高。为了进一步研究多重注意力算法对特征通道的调整结果，对 V2 - ResNet - 101 和 MANet 神经网络模型的第 1 阶段最后 1 个残差单元输出特征图进行特征可视化，特征可视化样本图像如图 3 所示，结果分别如图 4 和图 5 所示。由图 4 和图 5 可知，SE 注意力算法和 Fca 注意力算法所产生的加权系数无法有效构建语义特征在特征图中的显著性表达。而多重注意力方法 MA 和 Double - SE 通过对特征图进行 2 次加权，初步构建了显著性语义特征，呈现了木薯叶病害检测所需要的高质量上下文信息，从而提高了模型的检测性能。



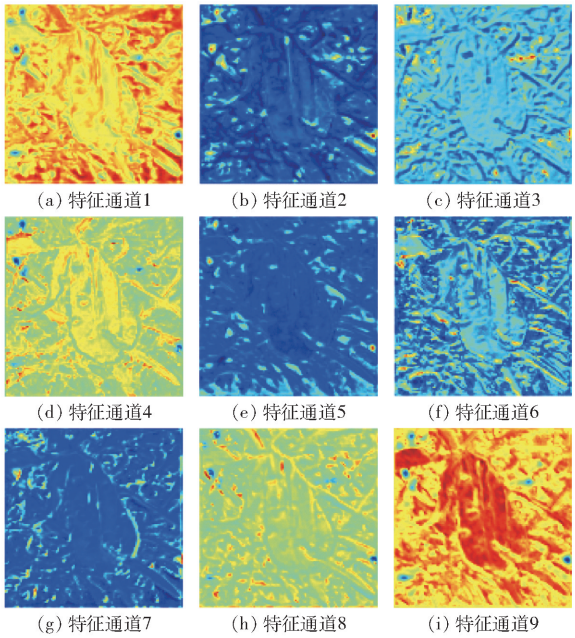
图 3　特征可视化样本图像

Fig. 3　Raw image of feature map visualization

在图 4a、4b、4c、4f、4i 中，V2 - ResNet - 101 神经网络模型虽然提取到了丰富的纹理特征和轮廓特征，然而，过多的无关纹理特征对木薯叶语义特征造成了干扰，无法在特征图中呈现清晰的木薯叶语义特征。如图 4d、4e、4g、4h 所示，过多的纹理特征和轮廓特征造成了对木薯叶语义特征的干扰。

在图 5a、5f、5g、5h、5i 中可见，木薯叶语义特征有着清晰呈现，特征图中无关纹理信息得到了有效抑制，没有对语义特征表达造成干扰。虽然，在图 5b 和图 5d 中存在部分语义特征缺失的情况，然而，特征图中保留的部分语义特征依旧清晰。在图 5c 中，木薯叶语义特征虽然并不显著，但是特征图中所包含的少量无关纹理特征并未对木薯叶语义特征产生干扰。其中图 5e 虽然并不包含语义特征，但是却有效抑制了无关纹理特征，没有引入干扰特征到随后的拟合特征图中。

在图 5 中可见，多重注意力算法在特征图中初步构建出了显著性特征。随着特征图中的语义特征质量的提升，MANet 神经网络模型对木薯叶病害检测的准确率得到了显著提升。



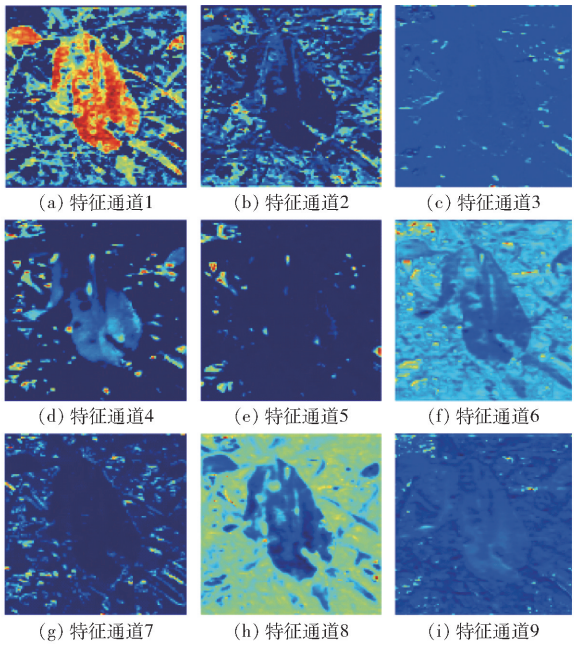
(a) 特征通道1　(b) 特征通道2　(c) 特征通道3

(d) 特征通道4　(e) 特征通道5　(f) 特征通道6

(g) 特征通道7　(h) 特征通道8　(i) 特征通道9

图 4　V2 - ResNet - 101 第 1 阶段最后 1 个残差单元输出特征图

Fig. 4　Part of output feature maps of last residual unit in stage 1 of V2 - ResNet - 101



(a) 特征通道1　(b) 特征通道2　(c) 特征通道3

(d) 特征通道4　(e) 特征通道5　(f) 特征通道6

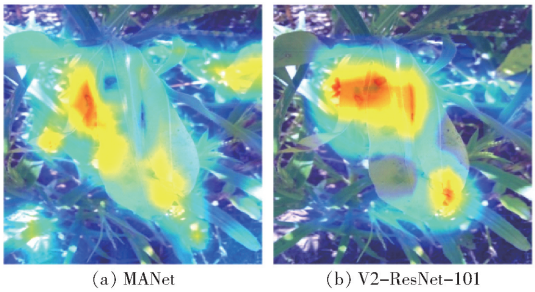
(g) 特征通道7　(h) 特征通道8　(i) 特征通道9

图 5　MANet 第 1 阶段最后 1 个残差单元输出特征图

Fig. 5　Part of output feature maps of last residual unit in stage 1 of MANet

为了清晰地展示神经网络对木薯叶病害的检测效果，使用 Grad - CAM 算法[30] 对 MANet 神经网络模型和 V2 - ResNet - 101 神经网络模型的最后 1 个

残差单元的输出特征进行计算,直观反映分类结果中的目标类别信息,木薯叶病害的 Grad－CAM 类激活映射可视化结果如图 6 所示。在图 6b 中可见,V2－ResNet－101 模型对木薯叶病害的类激活映射区域相对较大,但是引入了较多的非病害区域响应。而在图 6a 中,MANet 的类间激活映射图相较于 V2－ResNet－101,在木薯病害叶片上的覆盖面积更大,且最高响应区域更为集中,同时 MANet 的类激活映射结果呈现出了对背景区域的过多关注。



(a) MANet (b) V2-ResNet-101

图 6 MANet 和 V2－ResNet－101 特征可视化热力图

Fig. 6 Feature visualization heat maps of MANet and V2－ResNet－101

2.3 不同归一化方法对模型性能的影响

HE 等[31]的研究指出,在残差单元后使用批归一化方法影响了卷积神经网络的性能表达。但 CHOI 等[32]的研究显示,通过采用不同的归一化策略和方法,可有效提高神经网络在不同数据域中的泛化能力。为此,在 MANet 神经网络模型的残差单元之后,使用批归一化方法和实例批归一化方法,分别构建 MABNet 和 MAIBNet 神经网络模型,研究归一化方法对模型性能的影响,结果如表 4 所示。实例归一化方法最初被提出用于风格化图像生成任务,并不适用于分类任务,因此实例批归一化算法没有参与表 4 的对比试验。

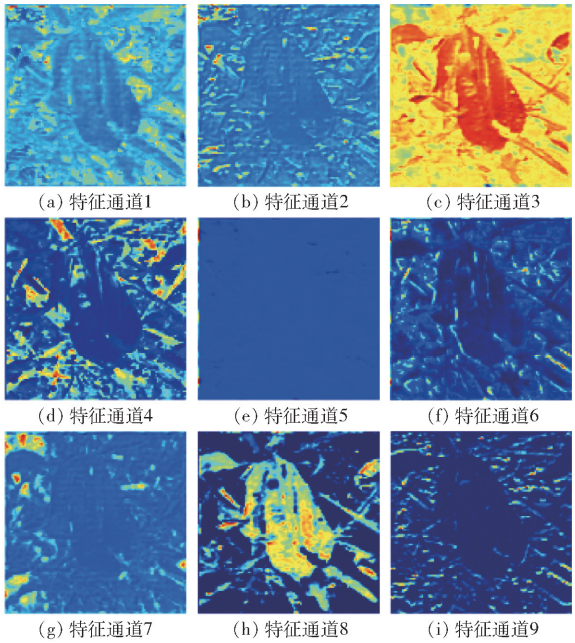
表 4 MANet 和 MAIBNet－101 基于 7 折交叉验证的试验结果对比

Tab. 4 Comparison of MANet and MAIBNet－101 using 7－fold cross-validation %

模型	准确率	召回率	精确率	F1 值
MANet	90.12	90.00	90.28	90.13
MABNet	92.03	91.98	92.13	92.06
MAIBNet－101	94.99	94.96	95.03	95.00

表 4 显示,在 MANet 神经网络模型的残差单元之后使用批归一化方法,模型性能得到明显提高,而采用实例归一化方法后,模型性能得到了显著提升。实例批归一化方法所带来的性能提升,得益于实例批归一化方法中的实例归一化方法对特征图数值表达质量的提升,从而使实例批归一化方法在木薯叶病害检测任务中展现出较强劲的性能。

为了进一步研究实例批归一化算法对特征通道的调整结果,对 MAIBNet－101 神经网络模型的第 1 阶段最后 1 个残差单元输出特征图进行特征可视化,结果分别如图 7 所示。在图 7a、7b、7c、7g、7h 中可以看到清晰的木薯叶语义特征区域。在图 7d、7f、7i 中,木薯叶语义区域仅包含极少的纹理特征,并且较好地呈现出了木薯叶形状特征。图 7 中的图像编号为特征通道索引号。



(a) 特征通道1 (b) 特征通道2 (c) 特征通道3

(d) 特征通道4 (e) 特征通道5 (f) 特征通道6

(g) 特征通道7 (h) 特征通道8 (i) 特征通道9

图 7 MAIBNet－101 第 1 阶段最后 1 个残差单元输出特征图

Fig. 7 Part of output feature maps of last residual unit in stage 1 of MAIBNet－101

由图 7 可见,显著性语义特征的表达有以下 2 个特点:特征图中目标物体的粗略聚焦和定位;减少了无关特征的干扰。实例批归一化方法显著提高了显著性语义特征质量,抑制了无关特征的表达。随着特征图中的语义特征质量的提升,MAIBNet 神经网络模型对木薯叶病害检测的准确率得到了显著提升。

为了清晰展示神经网络对木薯叶病害的检测效果,使用 Grad－CAM 算法[30]对 MAIBNet－101 神经网络模型的最后 1 个残差单元的输出特征进行计算,直观反映分类结果中的目标类别信息,木薯叶病害的 Grad－CAM 类激活映射可视化结果如图 8 所示。在图 8 MAIBNet－101 的类间激活映射中,相较于 MANet 和 V2－ResNet－101 最高响应区域明显更小,对木薯叶病害区域的定位更为精确,有效避免了无关特征的干扰。

2.4 MAISNet－101 算法检测结果及其与几种经典算法比较

为了展示 MAISNet－101 神经网络算法在木薯



图 8　MAIBNet－101 特征可视化热力图

Fig. 8　Feature visualization heat map of MAIBNet－101

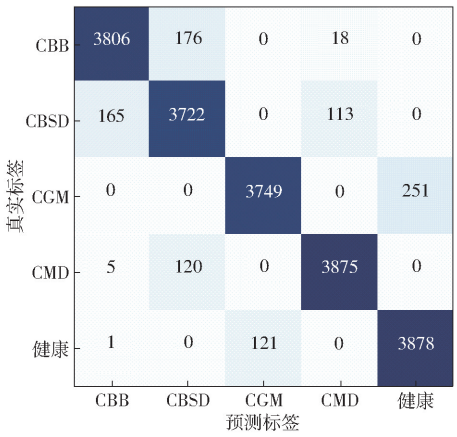


图 9　基于 7 折交叉验证的混淆矩阵

Fig. 9　Confusion matrix based on 7－fold cross-validation

叶病害数据集上的性能，使用 MAISNet－101 算法与几种经典神经网络算法的比较如表 5 所示。在所有性能指标对比中，MAISNet 均取得最佳结果。7 折交叉验证混淆矩阵如图 9 所示。

表 5　基于 7 折交叉验证的木薯叶病害检测经典算法对比结果

Tab. 5　Comparison of proposed method and other methods on cassava dataset using 7－fold cross-validation　%

模型	准确率	召回率	精确率	F1 值
MAISNet－101	95. 39	95. 36	95. 40	95. 38
V2－ResNet－101[19]	77. 84	77. 61	78. 17	77. 89
EfficientNet－B5[33]	92. 43	92. 37	92. 46	92. 42
VGG－19[34]	71. 93	71. 87	72. 02	71. 95
RepVGG－B3g4[35]	93. 08	93. 07	93. 14	93. 11

通过计算每折交叉验证的每类检测准确率，取均值得 CBB、CBSD、CGM、CMD、健康等 5 种木薯叶病害检测准确率分别为 95. 17%、93. 54%、94. 17%、96. 02%、95. 95%。

通过累加每折交叉验证得到最后一次验证结果如图 9 所示。由表 5 可知，模型对各种木薯叶病害检测准确率均较高，说明模型泛化性和鲁棒性较好。

各种算法的准确率、召回率、精确率和 F1 值对比如图 10 所示。观察图 10a 可知，MAISNet－101 网络从第 1 轮迭代训练到第 10 轮迭代训练，均表现出快速的收敛速度。而且在经过 10 轮迭代训练之后，MAISNet－101 的准确率就在一个狭窄的区间内稳定振荡，说明所提出的 MAISNet 算法具有较快的收敛速度和较稳定的特征语义描述能力。另外，通过对比 7 折交叉验证的结果可以发现，尽管卷积层堆叠的深度较大，但 RepVGG－B3g4 曲线的收敛速度明显快于 EfficientNet－B5 神经网络。由此可见，深度卷积神经网络的性能不完全取决于卷积层堆叠的深度。

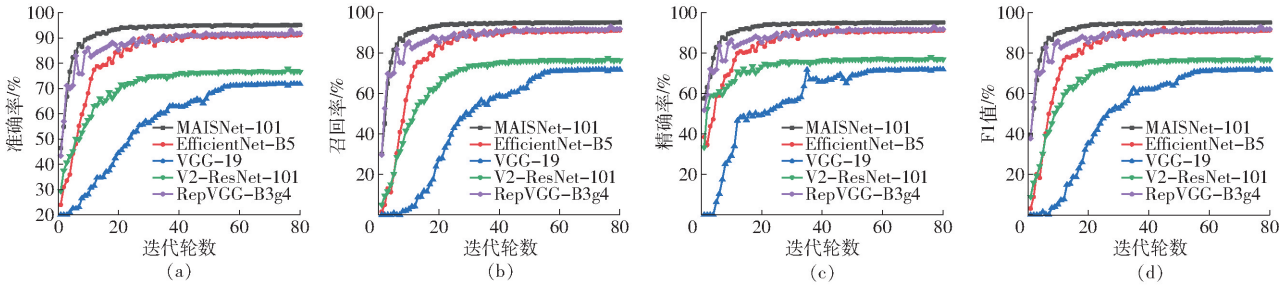


图 10　MAISNet－101 卷积神经网络与经典神经网络模型在木薯数据集上的性能曲线对比

Fig. 10　Performance curves of algorithms on cassava dataset

3　结束语

提出了一种用于木薯叶病害检测的选择性注意力卷积神经网络 MAISNet，通过引入多重注意力机制、实例批归一化方法和 Squareplus 激活函数，相比 V2－ResNet－101 神经网络模型最好结果，在准确率、召回率和精确率上分别提升 8. 49、8. 56、8. 38 个百分点，明显优于几种经典神经网络算法。基于可视化结果可知，显著性语义特征的构建提升了语义特征质量，减少了无关特征干扰，显著提升了神经网络模型在复杂非结构环境下的木薯叶病害检测准确率。所提出的基于选择性注意力神经网络的木薯叶病害检测算法，在木薯叶病害检测方面具有实际应用价值。

参　考　文　献

[1]　ZHANG L, ZHANG J, WEI Y, et al. Microbiome-wide association studies reveal correlations between the structure and

metabolism of the rhizosphere microbiome and disease resistance in cassava[J]. Plant Biotechnology Journal, 2021,19(4): 689 - 701.

[2] LEGG J P, KUMAR P L, MAKESHKUMAR T, et al. Cassava virus diseases: biology, epidemiology, and management[M]// Advances in Virus Research. Elsevier, 2015:85 - 142.

[3] SAMBORSKI S M, TREMBLAY N, FALLON E. Strategies to make use of plant sensors-based diagnostic information for nitrogen recommendations[J]. Agronomy Journal, 2009,101(4):800 - 816.

[4] GOLHANI K, BALASUNDRAM S K, VADAMALAI G, et al. A review of neural networks in plant disease detection using hyperspectral data[J]. Information Processing in Agriculture, 2018,5(3):354 - 371.

[5] MARTINELLI F, SCALENGHE R, DAVINO S, et al. Advanced methods of plant disease detection. A review[J]. Agronomy for Sustainable Development, 2015,35:1 - 25.

[6] BOFARHE O E, BOFARHE J O E, ADEBAYO S, et al. Detection and classification of cassava diseases using machine learning[J]. International Journal of Computer Science and Software Engineering, 2019,7(8):166 - 176.

[7] RAMCHARAN A, BARANOWSKI K, MCCLOSKEY P, et al. Deep learning for image-based cassava disease detection[J]. Frontiers in Plant Science, 2017,8:1852.

[8] DHIVYAA C R, KANDASAMY N, RAJENDRAN S. Integration of dilated convolution with residual dense block network and multi-level feature detection network for cassava plant leaf disease identification[J]. Concurrency and Computation: Practice and Experience, 2022,34:e6879.

[9] MI Z, ZHANG X, SU J, et al. Wheat stripe rust grading by deep learning with attention mechanism and images from mobile devices[J]. Frontiers in Plant Science, 2020,11:558126.

[10] TSOTSOS J K, CULHANE S M, WAI W Y K, et al. Modeling visual attention via selective tuning[J]. Artificial Intelligence, 1995,78(1 - 2):507 - 545.

[11] ZHANG J, BARGAL S A, LIN Z, et al. Top-down neural attention by excitation backprop[J]. International Journal of Computer Vision, 2018,126(10):1084 - 1102.

[12] LIU X, MIN W, MEI S, et al. Plant disease recognition: a large-scale benchmark dataset and a visual region and loss reweighting approach[J]. IEEE Transactions on Image Processing, 2021,30: 2003 - 2015.

[13] ZHOU Z. A brief introduction to weakly supervised learning[J]. National Science Review, 2018,5(1):44 - 53.

[14] LEE S H, CHAN C S, MAYO S J, et al. How deep learning extracts and learns leaf features for plant classification[J]. Pattern Recognition, 2017,71:1 - 13.

[15] ZHANG J, QI C, MECHA P, et al. Pseudo high-frequency boosts the generalization of a convolutional neural network for cassava disease detection[J]. Plant Methods, 2022,18:136.

[16] PAN X, lUO P, SHI J, et al. Two at once: enhancing learning and generalization capacities via ibn-net[C]// Proceedings of the European Conference on Computer Vision (ECCV),2018: 464 - 479.

[17] BARRON J T. Squareplus: a softplus-like algebraic rectifier[J]. arXiv Preprint, arXiv:2112.11687, 2021.

[18] REDDY S R, VARMA G S, DAVULURI R L. Resnet-based modified red deer optimization with DLCNN classifier for plant disease identification and classification[J]. Computers and Electrical Engineering, 2023,105:108492.

[19] NVIDIA. ResNet v1.5 for PyTorch[EB/OL]. [2021 - 09 - 03]. https://ngc.nvidia.com/catalog/model-scripts/nvidia:resnet_50_v1_5_for_pytorch.

[20] 王自全, 张永生, 于英, 等. 深度学习背景下视觉显著性物体检测综述[J]. 中国图象图形学报, 2022,27(7): 2112 - 2128.
WANG Ziquan, ZHANG Yongsheng, YU Ying, et al. Review of deep learning based salient object detection[J]. Journal of Image and Graphics, 2022, 27(7): 2112 - 2128. (in Chinese)

[21] ITTI L, KOCH C, NIEBUR E. A model of saliency-based visual attention for rapid scene analysis[J]. IEEE Transactions on pattern analysis and machine intelligence, 1998,20(11):1254 - 1259.

[22] 许楠,苑迎春,雷浩,等.基于多器官特征融合的枣品种识别方法[J].农业机械学报,2024,55(4):213 - 220,240.
XU Nan, YUAN Yingchun, LEI Hao, et al. Jujube variety recognition method based on multi-organ feature fusion[J]. Transactions of the Chinese Society for Agricultural Machinery,2024,55(4):213 - 220,240. (in Chinese)

[23] 袁培森,欧阳柳江,翟肇裕,等. 基于 MobileNetV3Small - ECA 的水稻病害轻量级识别研究[J]. 农业机械学报,2024,55(1):253 - 262.
YUAN Peisen, OUYANG Liujiang, ZHAI Zhaoyu, et al. Lightweight identification of rice diseases based on improved ECA and MobileNetV3Small[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(1): 253 - 262. (in Chinese)

[24] 袁培森,丁毅飞,徐焕良. 基于深度主动学习与 CBAM 的细粒度菊花表型识别[J]. 农业机械学报,2024,55(2):258 - 267.
YUAN Peisen, DING Yifei, XU Huanliang, et al. Fine-grained chrysanthemum phenotype recognition based on deep active learning and CBAM[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(2): 258 - 267. (in Chinese)

(下转第 272 页)

[28] 安焕炯,翟晨,马倩云,等. 结合 2D - COS 和光谱融合技术的小麦淀粉回生特性定量表征[J]. 光谱学与光谱分析,2023,43(1):162 - 168.
AN Huanjiong, ZHAI Chen, MA Qianyun, et al. Quantitative characterization of wheat starch retrogradation by combining 2D COS and spectral fusion[J]. Spectroscopy and Spectral Analysis, 2023, 43(1): 162 - 168. (in Chinese)
[29] JIANG Hongzhe, RU Yu, CHEN Qing, et al. Near-infrared hyperspectral imaging for detection and visualization of offal adulteration in ground pork[J]. Spectrochim Acta Part A Mol. Biomol. Spectrosc., 2021, 249: 119307.
[30] LIU Zhimin, SHEN Tao, ZHANG Ji, et al. A novel multi-preprocessing integration method for the qualitative and quantitative assessment of wild medicinal plants: gentiana rigescens as an example[J]. Frontiers in Plant Science, 2021, 12: 759248.
[31] HE Kaiming, ZHANG Xiangyu, REN Shaoqing, et al. Deep residual learning for image recognition[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016:770 - 778.
[32] 郝勇,吴文辉,商庆园. 饲料中粗脂肪和粗纤维含量的近红外光谱快速分析[J]. 光谱学与光谱分析,2020,40(1):215 - 220.
HAO Yong, WU Wenhui, SHANG Qingyuan. The research on quantitative analysis of feed crude fat and corase fiber based on near infrared spectroscopy and variables selection methods[J]. Spectroscopy and Spectral Analysis, 2020, 40(1): 215 - 220. (in Chinese)
[33] 黄亦其, 刘祥焕, 黄震宇, 等. 基于高光谱成像的甘蔗叶片早期轮斑病与锈病识别技术[J]. 农业机械学报,2023,54(4):259 - 267.
HANG Yiqi, LIU Xianghuan, HUANG Zhenyu, et al. Identification of early wheel spot and rust on sugarcane leaves based on spectral analysis[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(4): 259 - 267. (in Chinese)

(上接第 262 页)

[25] GUO M, XU T, LIU J, et al. Attention mechanisms in computer vision: a survey[J]. Computational Visual Media, 2022, 8(3):331 - 368.
[26] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 7132 - 7141.
[27] QIN Z, ZHANG P, WU F, et al. Fcanet: frequency channel attention networks[C]// Proceedings of the IEEE/CVF International Conference on Computer Vision,2021: 783 - 792.
[28] 魏秀参. 解析深度学习:卷积神经网络原理与视觉实践[M]. 北京: 电子工业出版社, 2018.
[29] WANG H, WU X, HUANG Z, et al. High-frequency component helps explain the generalization of convolutional neural networks[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 8684 - 8694.
[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.
[31] HE K, ZHANG X, REN S, et al. Identity mappings in deep residual networks[C]// Computer Vision-ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11 - 14, 2016, Proceedings, Part IV 14. Springer International Publishing, 2016: 630 - 645.
[32] CHOI S, KIM T, JEONG M, et al. Meta batch-instance normalization for generalizable person re-identification[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021: 3425 - 3435.
[33] TAN M, LE Q. Efficientnet: rethinking model scaling for convolutional neural networks[C]// International Conference on Machine Learning. PMLR, 2019: 6105 - 6114.
[34] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[C]// 3rd International Conference on Learning Representations (ICLR 2015). Computational and Biological Learning Society, 2015.
[35] DING X, ZHANG X, MA N, et al. Repvgg: making vgg-style convnets great again[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021: 13733 - 13742.