

基于改进的注意力机制残差网络 穴盘幼苗分类算法研究

吴聪, 郭志强, 杨杰*

武汉理工大学信息工程学院, 湖北 武汉 438300

摘要 针对 SENet 的通道注意力机制特征提取单一和分割的幼苗数据集图片存在部分缺失的难点问题, 设计了一种基于双通道注意力机制的残差网络。该网络融合通道注意力机制和空间注意力机制模块, 可同时获得通道和空间维度特征权重, 提升网络的特征学习能力。提出了一种随机擦除方法, 来解决分割样本数据中目标部分缺失的难点问题。在自制的穴盘幼苗 Plant_seed 数据集上的实验结果表明, 在 ResNet34 残差模块和 conv*_x 模块之间均引入注意力机制模块的改进网络 ResNet34+CBAM_basic_conv 的准确率最优, 达到 93.8%, 同时对数据集部分图片进行随机擦除后, 模型分类的错误率下降, 验证了所提方法的优异性能。

关键词 图像处理; 图像分类; 注意力机制; 残差网络; 随机擦除; 穴盘幼苗

中图分类号 TP183

文献标志码 A

DOI: 10.3788/LOP202259.2210002

Algorithm for Plug Seedling Classification Based on Improved Attention Mechanism Residual Network

Wu Cong, Guo Zhiqiang, Yang Jie*

College of Information Engineering, Wuhan University of Technology, Wuhan 438300, Hubei, China

Abstract A residual network based on dual-channel attention mechanism is designed to address the difficulty in extracting single and segmented seedling dataset images from the channel attention mechanism feature of the SENet network, which integrates the channel attention mechanism and spatial attention. The mechanism module can obtain the channel and spatial dimension feature weights simultaneously to enhance the feature learning ability of the network. To address the problem of missing the target in the segmented sample data, a random erasure method is proposed. Experiments on the self-made plug seedling Plant_seed dataset demonstrate that the improved network ResNet34+CBAM_basic_conv, which introduces the attention mechanism module between the ResNet34 network residual module and the conv*_x module, reaches the optimal accuracy of 93.8%. The error rate of the model classification drops after some images in the dataset are randomly erased, demonstrating the excellent performance of the proposed method.

Key words image processing; image classification; attention mechanism; residual network; random erase; plug seedling

1 引言

随着计算机视觉在农业领域的广泛应用, 比如植物生长检测、农产品分级等, 自动获取穴盘幼苗类别并利用智能设备实现穴盘幼苗的自动移栽就显得尤为重要。穴盘苗的识别就是在幼苗生长期识别出穴盘内空穴和弱苗株的坐标位置, 对空穴孔进行补苗, 对弱苗孔进行壮苗替换后重新培育, 以保证每一批穴盘苗的规

格处在同一标准。

对于穴盘苗的识别主要有传统的图像处理识别技术和近年兴起的神经网络分类技术。肖超云^[1]以小麦田间杂草样本为研究对象, 提出了一种改进的基于最大方差准则的阈值分割算法, 并对几种常见的小麦田杂草图像进行了测试, 采用形状特征参数作为模式识别输入的特征向量。实验结果表明, 圆形度是识别阔叶杂草与窄叶杂草最为有效的特征, 但是该算法存在

收稿日期: 2021-08-23; 修回日期: 2021-09-03; 录用日期: 2021-09-24

基金项目: 国家重点研发计划(2019YFD1001900)、国家自然科学基金(51479159)

通信作者: *jieyang@whut.edu.cn

数据量不足的问题。杨镇宇等^[2]利用机器视觉和支持向量机(SVM)技术对花椒外观品质进行检测。何彦虎等^[3]使用K-means聚类算法对采集的穴盘苗图像进行分割,并采用专家系统和SVM算法对10个品种的穴盘苗进行分类,实验结果表明,专家系统的识别率能达到98.3%,不过其需要设计大量的规则,过程繁琐。Lameski等^[4]利用近红外光谱和可见光光谱所产生的归一化植被指数(NDVI)不同的特点,建立烟苗分割绿色区域,并利用机器学习的方法分割烟苗和杂草,实现对不同杂草的区分。刘合琛^[5]提出了一种基于卷积神经网络和灰度共生矩阵^[6](GLCM)进行 Dempster/Shافر 证据理论融合判别的全景植物识别方法,该方法采用二维模式(LBP)提取图像纹理信息,并进行训练和分类,得出了较好的植物分类效果,其难点在于识别时间与识别率之间很难找到最合理的平衡点。

侯晨伟等^[7]采用概率神经网络(PNN)^[8]分类器识别玉米幼苗和杂草,对不同田间环境的玉米幼苗与杂草的识别准确率分别为92.5%和95%,其缺点是当玉米苗和杂草交叉严重时,基于形状特征的杂草识别算法的识别率降低。何艳等^[9]基于AdaBoost算法实现穴盘烟苗分割,然后采用多层感知器和AlexNet^[10]对穴盘烟苗进行识别和分类,两种算法的准确率分别为96.75%和97.58%,但其对选取的特征要求很高,不同的特征对实验结果影响较大。左羽等^[11]将VGG16网络的全连接层改为卷积层并对43类植物图像进行分类,图像的分类准确率达到

97.23%。娄月^[12]将注意力机制模块SENet^[13]引入改进的DenseNet^[14],降低了网络参数规模,适应大尺度图片输入和小样本数据集的训练学习要求,在花卉数据集上有较好的准确率表现,但是其注意力机制提取的特征单一。

本研究在基于神经网络的图像分类算法的基础上,对原有的卷积神经网络结构进行了改进,将通道注意力机制和空间注意力机制引入残差网络ResNet^[15]中,以白茄苗、花菜苗、辣椒苗、茄子苗、芹菜苗、西瓜苗、黄瓜苗等7种穴盘幼苗图像作为数据样本,通过图像矫正、图像分割、数据扩充等预处理操作获得幼苗数据集,然后将随机擦除的数据样本输入改进的残差网络进行训练,实现了穴盘幼苗的多分类。最后,将所提模型与其他主流网络进行了对比分析,以此来验证所提改进模型的有效性。

2 数据预处理

数据预处理工作的主要任务是将采集的穴盘幼苗图像制作成标准的数据集。本研究采集的穴盘图像数据来自武汉如意生态园培育的农作物幼苗,具体要求是农作物的叶片较少遮挡穴盘网格边缘,这样的选择有利于提高穴盘图像分割的准确率。同时选择的幼苗不宜过小,过小的幼苗会增加识别难度。通过筛选,最终所得穴盘图像数据如表1所示,其中BQ、HC、LJ、QZ、QC、XG、HG分别是白茄、花菜、辣椒、茄子、芹菜、西瓜、黄瓜的缩写。

表1 采集的穴盘幼苗图像
Table 1 Images of plug seedlings collected

Plug data	BQ	HC	LJ	QZ	QC	XG	HG
Quantity /disc	10	29	10	29	15	9	11
Plug specification	12×6	12×6	12×6	12×6	12×6	16×8	12×6

2.1 穴盘矫正

穴盘图像在采集过程中,容易受到环境因素的干扰,如光照强度、人为因素等。因此,首先要对图像进行预处理,包括直方图均衡化、图像归一化、中值滤波、穴盘矫正操作。但是研究中发现,除穴盘矫正对图像分类结果影响较大外其余图像预处理操作对实验结果影响可忽略不计,故仅对穴盘幼苗图像进行矫正处理。首先创建图像矫正的算法模型,然后从失真的图像信号中提取所需要的信息,最后沿着图像失真的逆向过程恢复原始图像。在几何校正中,图像空间几何坐标变换分为两种:1)三角形线性法;2)二元多项式法。由于二元多项式算法精度较高,故采用二元多项式法,其表达式为

$$\begin{cases} x = \sum_{i=0}^n \sum_{j=0}^{n-i} a_{ij} u^i v^j \\ y = \sum_{i=0}^n \sum_{j=0}^{n-i} b_{ij} u^i v^j \end{cases}, \quad (1)$$

式中: (x, y) 为矫正之前的坐标; (u, v) 为矫正后的图像坐标; a_{ij}, b_{ij} 为多项式系数; $n = 1, 2, 3, \dots$ 。某穴盘矫正之后的图像如图1所示。

2.2 穴盘图像分割

穴盘图像分割的质量直接影响神经网络训练的质量,好的图像分割算法对制作数据集、训练分类网络尤为重要,AdaBoost检测算法是目前运用最成功、最稳定的检测算法。其检测过程可分为3部分:Haar特征提取、基于AdaBoost的训练和分类器级联。工作原理为针对同一个训练集训练不同的分类器,然后把把这些弱分类器进行分析、归类,进而形成一个更强的最终分类器。给定训练集 $(x_1, y_1), \dots, (x_N, y_N)$,其中 $y_i \in \{1, -1\}$ 表示 x_i 的正确类别标签, $i = 1, \dots, N$ 。AdaBoost算法的具体步骤如下:

1) 计算出训练集上样本的初始分布,

$$D_1(i) = \frac{1}{N}. \quad (2)$$

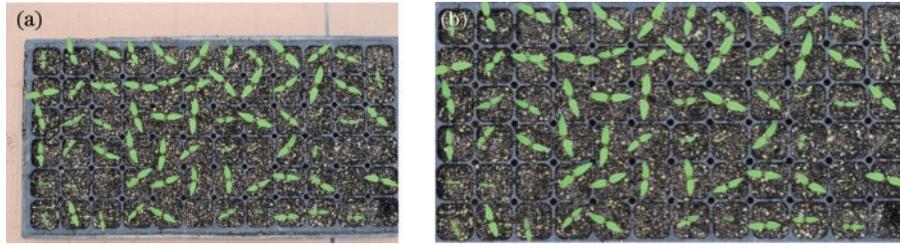


图1 矫正穴盘图像示例。(a)原始穴盘图像;(b)矫正的穴盘图像

Fig. 1 Example of corrected disk image. (a) Original disk image; (b) corrected disk image

2) 对 $t=1, \dots, T$ 分布 D_t 上找出具有最小错误率的弱分类器 $h_t: X \rightarrow \{-1, 1\}$, 其中某弱分类器在分布 D_t 上的错误率的表达式为

$$\epsilon_t = P_{D_t}[h_t(x_i) \neq y_i]. \quad (3)$$

3) 计算该弱分类器的权重系数,

$$a_t = \frac{1}{2} \ln \left(\frac{1 - \epsilon_t}{\epsilon_t} \right). \quad (4)$$

4) 更新训练样本的分布,

$$D_{t+1}(i) = \frac{D_t(i) \exp[-a_t y_i h_t(x_i)]}{Z_t}, \quad (5)$$

式中: Z_t 为归一化常数。

5) 得出最后的强分类器,

$$H_{\text{final}}(x) = \text{sign} \left[\sum_{t=1}^T a_t h_t(x) \right]. \quad (6)$$

通过 AdaBoost 算法分类器得到的穴盘图像分割结果如图 2 所示。

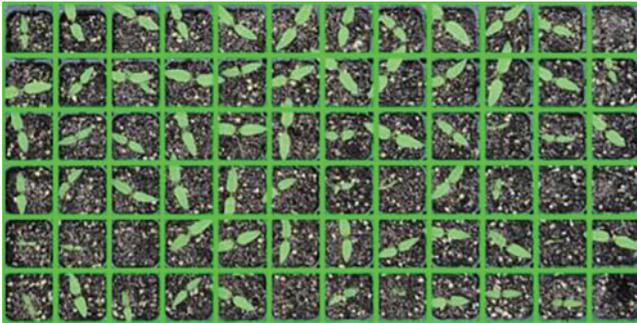


图2 穴盘图像分割结果

Fig. 2 Segmentation result of plug image

实验共采集了 113 张穴盘幼苗图像, 通过 AdaBoost 分割算法获得了 8640 张穴盘幼苗样本, 由于样本的不平衡, 不同种类的幼苗图片的数量差别明显。为了增加网络的鲁棒性并减少过拟合, 采用旋转的方法对穴盘样本图像进行扩充, 最后与项目合作的专家共同定制壮苗、弱苗、空穴分类标准, 并将扩充的数据样本制作成标准的数据集。在制作数据集的过程中, 将所有空穴(K_X)归为同一类, 最终网络需要对 15 ($7 \times 2 + 1$) 种不同类别穴盘幼苗进行训练分类。将制作的数据集命名为 Plant_seed, 其分布如表 2 所示。

表 2 Plant_seed 数据集构成

Table 2 Plant_seed dataset composition

Plant_seed	BQ	HC	LJ	QZ	QC	XG	HG
Strong seedling	645	1888	400	2051	948	3395	503
Weak seedling	75	420	850	469	610	1165	550
K_X				3249			
Total				17219			

2.3 穴盘图像分割难点与解决方案

本研究关于穴盘幼苗分类的难点主要是分割的图片存在部分缺失。穴盘幼苗的种类不同、生长的周期不一致, 导致穴盘幼苗存在分布不规则, 最终造成算法分割的幼苗图片部分出现缺失, 如果直接将存在缺失的图片放入模型训练, 会增加模型分类的错误率, 模型也难以学习到缺失图片的这部分特征。存在缺失部分的图片如图 3(a)~(d) 所示。

从 [图 3(a)~(d)] 可以看出, 由于幼苗生长偏移穴盘位置, 分割的图片只有部分叶片。为了提高模型对缺失图像的分类准确率, 学习到含有缺失部分的幼苗图像特征, 采用一种矩形框随机擦除的方法对训练样本进行遮挡处理, 制造穴盘图片被部分遮挡的数据样本。随机选择图像中的一个矩形区域 I_e , 并用随机值擦除其像素。假设训练图像面积 $S = W \times H$ 。随机初始化“擦除矩形”大小为 S_e , “擦除矩形”的纵横比 $r_e \in (r_1, r_2)$, 矩形 I_e 的高 H_e 和宽 W_e 分别 $\sqrt{S_e \times r_e}$ 和 $\sqrt{S_e / r_e}$ 。在图像中随机取一个点 $P(x_e, y_e)$, 矩形区域以这个点展开 $I_e = (x_e, y_e, x_e + W_e, y_e + H_e)$ 。在选定的擦除区域内, 每个像素都被分别随机设定在 $[0, 255]$ 范围内。研究发现 $p=0.2$ 、 $s_1=0.3$ 、 $s_b=0.4$ 、 $r_1=0.3$ 、 $r_2=3$ 时的效果最好, 即对数据样本量的 20% 进行擦除处理, 擦除的面积占比在 0.3~0.4 之间, 擦除矩形的长宽比在 0.3~3 之间。部分擦除的数据如图 3(e)~(h) 所示, 本研究采用的随机擦除方法可以模拟实际场景中存在的分割缺失问题, 这一方法加大了网络训练难度, 一定程度提高了网络泛化能力, 减小了网络分类的错误率。

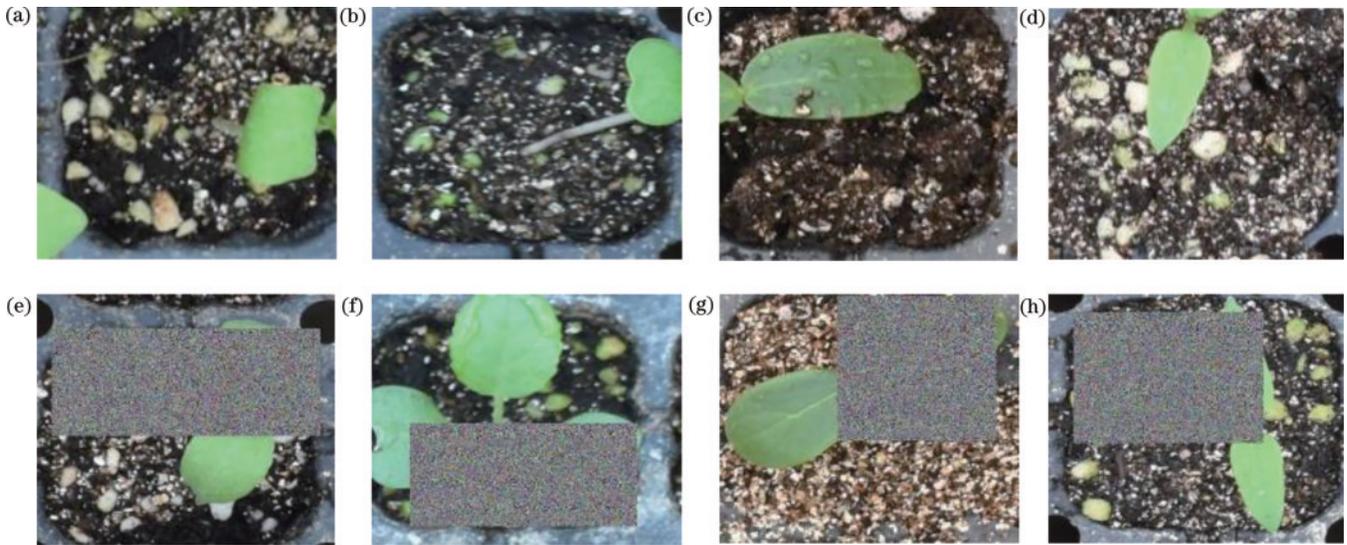


图 3 穴盘图片分割缺失与正常图片擦除处理。(a)分割缺失的白茄壮苗;(b)分割缺失的花菜壮苗;(c)分割缺失的黄瓜壮苗;(d)分割缺失的辣椒壮苗;(e)擦除处理后的白茄壮苗;(f)擦除处理后的花菜壮苗;(g)擦除处理后的黄瓜壮苗;(h)擦除处理后的辣椒壮苗

Fig. 3 Loss of plug picture segmentation and normal picture erasure processing. (a) Strong seedling of white eggplant with segmentation loss; (b) strong seedling of cauliflower with segmentation loss; (c) strong seedling of cucumber with segmentation loss; (d) strong seedling of capsicum with segmentation loss; (e) strong white eggplant seedling after erasing processing; (f) strong cauliflower seedling after erasing processing; (g) strong cucumber seedling after erasing processing; (h) strong capsicum seedling after erasing processing

3 分类模型

深度神经网络通常比较难训练,这是由于随着网络层数的增加,会产生显著的梯度消失或者梯度爆炸的问题。同时在网络能够收敛的前提下,随着深度的增加,正确率开始饱和甚至下降,导致网络的退化,而残差网络 ResNet 正是为了解决这一系列问题而提出的。与此同时,注意力机制与神经网络的结合也成为目前研究的热点。研究发现,在 Plant_seed 数据集中,图像的目标都是单一的,同时数据集采用旋转法进行扩充,这对关键区域识别的要求更高,而空间注意力机制可以很好解决这一问题。这是因为空间注意力机制使用一种空间转换器将图片的空间域信息进行对应的空间变换,该变换具有旋转、缩放变换不变性。本研究将残差网络和空间通道注意力机制融合提出了一种融合注意力机制的残差网络,并对残差网络进行了改进,

以提高网络的分类性能。

3.1 注意力机制

3.1.1 通道注意力机制(CAM)

通道注意力机制利用特征通道间的关系来产生通道权重图,为了得到通道维度的特征权重,首先在空间维度进行维度压缩操作,使空间的特征降维为一维权重向量,这一过程一般采用全局平均池化。而本研究在使用全局平均池化(global average pooling)降维的基础上,融合最大池化(global max pooling)操作进行通道特征降维。由两个池化函数可以得到两个一维矢量,其特点在于全局平均池化对特征图上的每一个像素点都有反馈,而最大池化在进行梯度反向传播计算时只有特征图中响应最大的地方有梯度反馈。图 4 展示了通道注意力机制的架构。通道注意力机制算法与 SENet 区别在于增加了最大池化操作提取通道特征权重。首先将特征图在

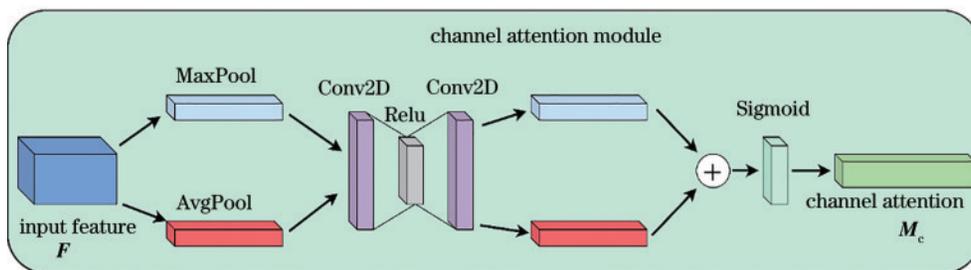


图 4 通道注意力机制

Fig. 4 Channel attention mechanism

空间维度上进行压缩,其中 F 表示输入特征图, F_{avg}^c 和 F_{max}^c 分别代表经过全局平均池化和最大池化计算

之后的特征向量, W_1 和 W_0 代表的是两层卷积 (Conv2D) 的参数,可描述为

$$\begin{cases} M_c(F) = \sigma \{ \text{Conv2D}[\text{AvgPool}(F)] + \text{Conv2D}[\text{MaxPool}(F)] \} \\ \sigma = W_1[W_0(F_{avg}^c)] + W_1[W_0(F_{max}^c)] \end{cases}, \quad (7)$$

式中: W_1 和 W_0 之间的特征需要使用 Relu 激活函数来处理; σ 表示 Sigmoid 激活函数。

3.1.2 空间注意力机制 (SAM)

空间注意力模块利用特征间的空间关系生成空间注意力图,不同于通道注意力机制,空间注意力模块注意力更加关注关键信息的位置特性,这是一个完整的通道关联。空间注意力模块算法的流程如图 5 所示。

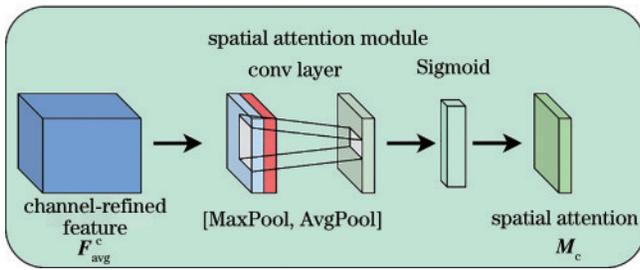


图 5 空间注意力模块

Fig. 5 Spatial attention module

空间注意力模块使用平均池化和最大池化对输入特征图进行压缩操作,但是这里的压缩变成了通道层面上的压缩,对输入特征分别在通道维度上进行了均值和最大值采样操作,得到了两个二维的特征,将其按通道维度拼接在一起得到一个通道数为 2 的特征图,之后使用一个包含单个卷积核的隐藏层对其进行卷积操作,最后通过 Sigmoid 激活函数得到在空间维度上与输入的特征图一致的包含空间信息的权重参数层。记经过平均池化和最大池化操作后的特征图为 $F_{avg}^s \in \mathbf{R}^{1 \times H \times W}$ 和 $F_{max}^s \in \mathbf{R}^{1 \times H \times W}$,该部分的数学处理可以描述为

$$M_s(F) = \sigma \{ f^{7 \times 7} \{ [\text{AvgPool}(F); \text{MaxPool}(F)] \} \} = \sigma \{ f^{7 \times 7} [F_{avg}^s; F_{max}^s] \}. \quad (8)$$

3.2 注意力机制融合残差模块

在对残差注意力机制融合模块进行设计时,首先要考虑的应该是两种注意力机制模块的自融合方法设计。通道注意力机制和空间注意力机制这两个模块能够以并行或者串行顺序的方式组合在一起,总共有 channel+spatial、spatial+channel、spatial 并行 channel 等 3 种组合方式,文献[16]指出,两个 attention 子模块的连接顺序里面 channel+spatial 的方式更好,故本研究同样选择这一方式搭建注意力机制残差模块。相对于直接将通道注意力机制插入残差模块最后一层卷积层,本研究提出了新的融合方式,图 6 是将串联的通道注意力机制和空间注意力机制组合为 CBAM 模块插入 ResNet34 残差模块的 2 个 weight layer 之间的插入方式。对于 ResNet50 选择在 Bottleneck 残差模块的 3 个 weight layer 之间插入两个 CBAM 注意力机制模块。同时,还设计了在 ResNet 卷积块之间插入 CBAM 模块的方式,如图 7 所示。在 conv2_x、conv3_x、conv4_x、conv5_x 之间插入 3 个 CBAM 模块,在卷积块之间引入 CBAM 注意力机制模块。最终改进的 ResNet 架构如表 3 所示,选择 ResNet-34 和 ResNet-50 作为基准网络进行基于注意力机制的改进,其中对勾表示在卷积块之间插入 CBAM 结构,同时将网络的输出改为 15 个类别的分类器。最后对改进后的网络进行基于 Plant_seed 数据集的一系列实验来验证注意力机制融合残差网络的有效性。

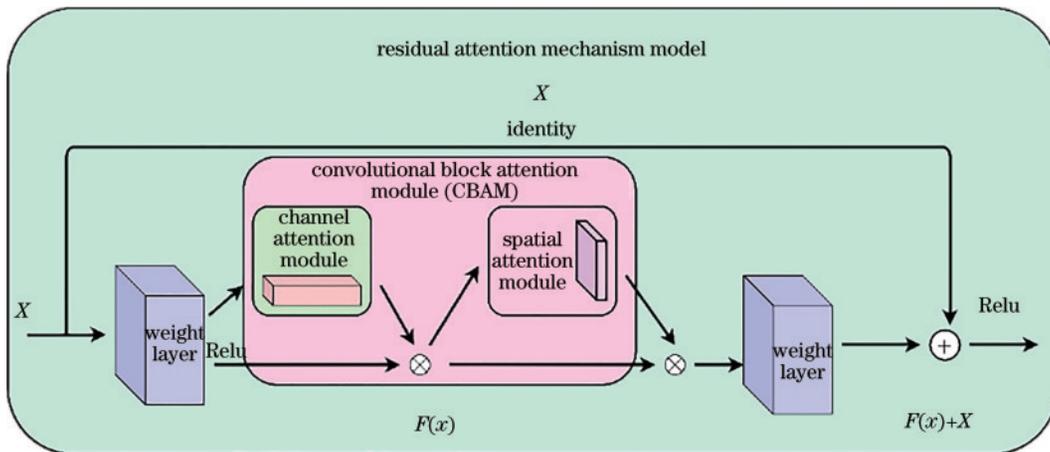


图 6 注意力机制融合残差模块

Fig. 6 Attention mechanism fusion residual module

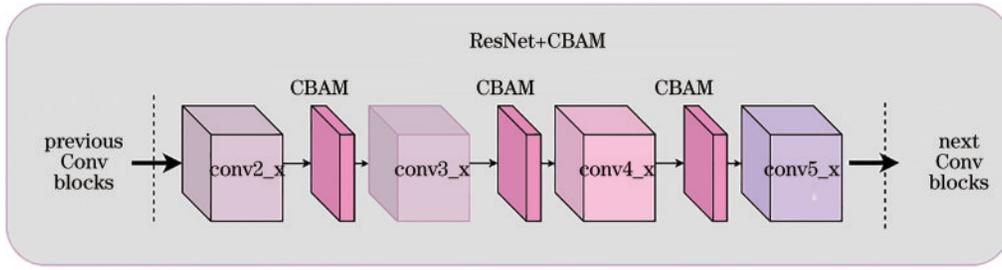


图 7 ResNet 卷积块间插入 CBAM 模块

Fig. 7 Inserting CBAM module between ResNet convolution blocks

表 3 改进的 ResNet 参数

Table 3 Improved ResNet parameters

Layer name	Output size	34-layer	50-layer	CBAM
conv1	112 × 112	7 × 7, 64, stride 2 3 × 3 maxpool, stride 2		
conv2_x	56 × 56	$\begin{bmatrix} 3 \times 3, 64 \\ \text{CBAM} \\ 3 \times 3, 64 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ \text{CBAM} \\ 3 \times 3, 64 \\ \text{CMAB} \\ 1 \times 1, 256 \end{bmatrix} \times 3$	✓
conv3_x	28 × 28	$\begin{bmatrix} 3 \times 3, 128 \\ \text{CBAM} \\ 3 \times 3, 128 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 128 \\ \text{CBAM} \\ 3 \times 3, 128 \\ \text{CMAB} \\ 1 \times 1, 512 \end{bmatrix} \times 4$	✓
conv4_x	14 × 14	$\begin{bmatrix} 3 \times 3, 256 \\ \text{CBAM} \\ 3 \times 3, 256 \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1, 256 \\ \text{CBAM} \\ 3 \times 3, 256 \\ \text{CMAB} \\ 1 \times 1, 1024 \end{bmatrix} \times 6$	✓
conv5_x	7 × 7	$\begin{bmatrix} 3 \times 3, 512 \\ \text{CBAM} \\ 3 \times 3, 512 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 512 \\ \text{CBAM} \\ 3 \times 3, 512 \\ \text{CMAB} \\ 1 \times 1, 2048 \end{bmatrix} \times 3$	
	1 × 1	Average pool, 15-d fc, softmax		
Gflops / MB		21.97	26.05	

4 实验结果分析

4.1 实验参数

实验在 Inter Xeon CPU E5-2620 v4@2.1 GHz、64 GB 内存、RTX 2080 Ti GPU、64 位 Windows 10 操作系统、Python3.6 和 Pytorch 框架下进行。使用 Plant_seed 数据集对模型进行训练和验证,其中训练集和验证集的划分比例为 7:3,同时对训练集图片的 20% 进行随机擦除处理。选择的 epoch 都为 50,学习率为 0.0001, batch size 为 32,使用 Adam 优化器进行优化计算,对网络模型设计不同的实验进行分析和验证。

4.2 CBAM 模块插入方式实验

本小节验证不同的 CBAM 模块插入方式对 ResNet34 的性能影响,首先单独以图 6 和图 7 的模块插入方式进行训练,接下来将图 6 与图 7 结合在一起组成新的网络进行训练,最后将 SENet 与图 6 的插入方式结合训练。训练的结果如表 4 所示。从表 4 的训练结果可以看出:不同的 CBAM 模块插入方式对原 ResNet 的参数量 (Params) 和浮点数计算量 (Gflops) 的改变是很小的,基本上可以忽略不计;但是从准确率上看,在 ResNet34 中插入 CBAM 注意力机制模块后对网络的性能均有所提升,相对于 ResNet34,在其残差模块和 conv*_x 模块之间均引入的网络 ResNet34+CBAM_basic_conv 的准确率提升了 3.11 个百分点,与此同时,相较于 ResNet34+SE 架构来说有 1.64 个百分点的准确率提升。这说明基于通道注意力机制和空间注意力机制网络相较于 SENet 来说是有效的。

表 4 CBAM 模块插入方式对比

Table 4 Comparison of CBAM module insertion methods

Architecture	Params / MB	Gflops	Accuracy / %
ResNet34	21.79	11.75	90.69 ± 1.61
ResNet34+SE	21.95	11.75	92.04 ± 1.36
ResNet34+SE+CBAM_basic	22.11	11.76	93.68 ± 0.62
ResNet34+CBAM_basic	21.96	11.76	93.46 ± 0.84
ResNet34+CBAM_conv	21.81	11.75	91.28 ± 0.92
ResNet34+CBAM_basic_conv	21.97	11.76	93.80 ± 0.80
ResNet50	25.56	13.16	91.58 ± 1.82
ResNet50+SE	28.07	13.18	93.19 ± 0.81
ResNet50+CBAM_bottle	25.87	13.18	92.32 ± 0.98
ResNet50+CBAM_conv	25.73	13.16	92.23 ± 0.97
ResNet50+CBAM_bottle_conv	26.05	13.19	94.89 ± 0.61
ResNet50+SE+CBAM_bottle	28.39	13.20	95.42 ± 0.92

同时还绘制出基于 ResNet34 和 ResNet50 的每个网络在每个 epoch 下的准确率 (accuracy) 和损失曲线,结果如图 8 所示,在加入 CBAM 注意力模块后,ResNet34 不仅能够达到更高的准确率,同时训练起始的准确率也更高,这说明网络能够更加高效学习到目标特征,加快了网络的训练速度。这一特点也可以从损失曲线图

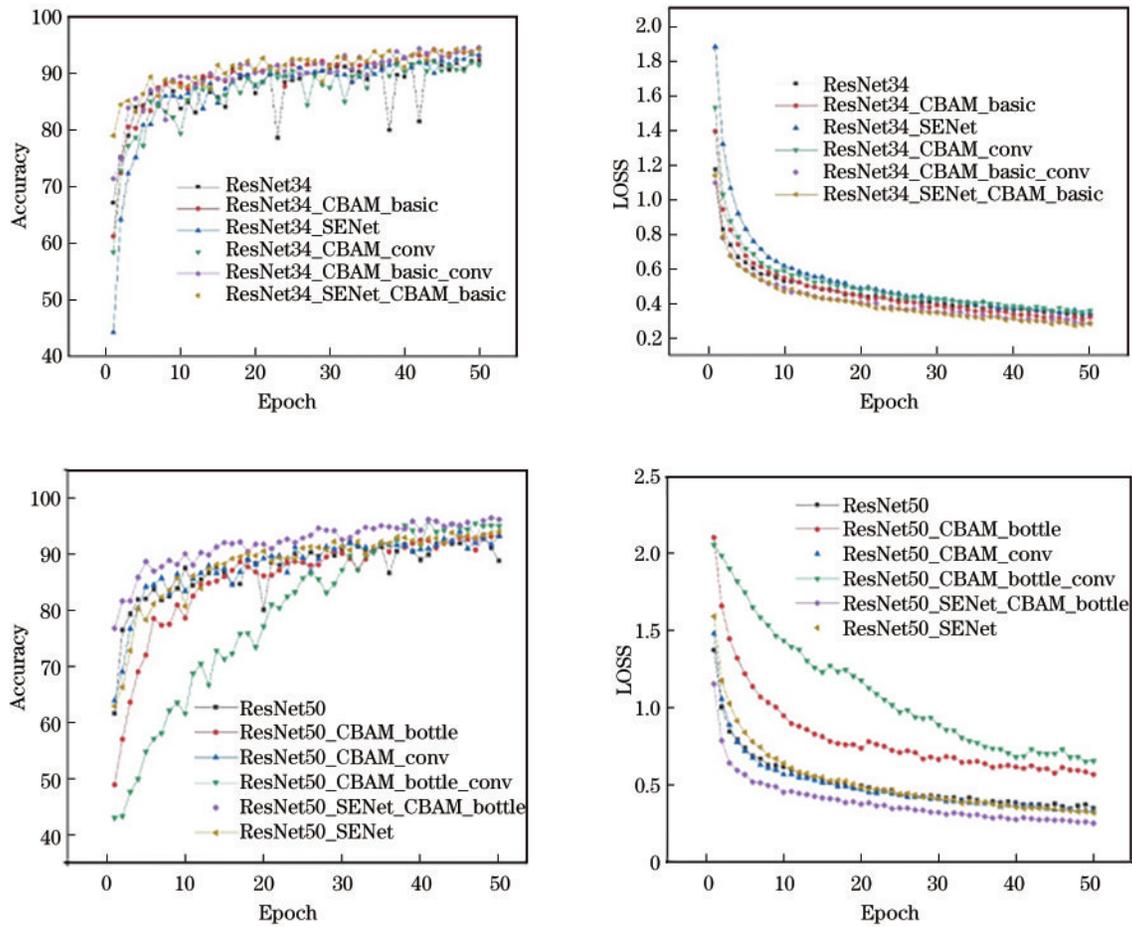


图 8 ResNet 不同注意力机制准确率与损失曲线

Fig. 8 Accuracy and loss curves of different attention mechanisms in ResNet

上看出。对于损失曲线, ResNet34_SENet_basic_conv 的损失相对于其他网络来说是最小的, 但是对于 ResNet50 来说准确率提升不如 ResNet34。

为了更加准确衡量不同注意力机制网络模型的对每一类幼苗的分类准确率, 选择分类准确率(P)、召回

率(R)、整体的平均精准率(AP)和平均召回率(AR)指标进行评价。各个模型的实验结果如表 5 所示。由于 ResNet34 的不同注意力机制模块可以更好展现不同网络的特性, 故只选择 ResNet34 为基准网络进行精确率、召回率计算。

表 5 各个模型在 Plant_seed 的精准率和召回率结果

Table 5 Accuracy and recall results of each model in Plant_seed

Architecture	BQW		BQS		HCW		HCS	
	P	R	P	R	P	R	P	R
ResNet34	0.655	0.864	0.974	0.959	0.959	0.746	0.993	0.940
ResNet34+SE	0.905	0.864	0.975	1.000	0.814	0.833	0.974	0.935
ResNet34+CBAM_basic	0.947	0.818	0.975	0.995	0.852	0.913	0.980	0.963
ResNet34+CBAM_conv	0.889	0.727	0.974	0.984	0.869	0.738	0.973	0.943
ResNet34+CBAM_basic_conv	0.905	0.864	0.990	1.000	0.753	0.897	0.981	0.924
ResNet34+SE+CBAM_basic	1.000	0.864	0.985	0.990	0.915	0.770	0.93	0.979

Architecture	HGW		HGS		K_X		LJW		LJS		QZW	
	P	R	P	R	P	R	P	R	P	R	P	R
ResNet34	0.934	0.945	0.993	0.940	0.914	0.943	0.915	0.973	0.980	0.808	0.933	0.700
ResNet34+SE	0.987	0.909	0.936	0.980	0.924	0.959	0.929	0.976	0.990	0.858	0.925	0.879
ResNet34+CBAM_basic	0.957	0.933	0.924	0.967	0.942	0.964	0.943	0.973	0.930	0.892	0.966	0.800
ResNet34+CBAM_conv	0.938	0.915	0.936	0.973	0.927	0.939	0.906	0.984	0.951	0.817	0.963	0.750

表5 (续)

Architecture	HGW		HGS		K_X		LJW		LJS		QZW	
	P	R	P	R	P	R	P	R	P	R	P	R
ResNet34+CBAM_basic_conv	0.952	0.97	0.966	0.953	0.956	0.962	0.955	0.988	0.991	0.900	0.873	0.979
ResNet34+SE+CBAM_basic	0.93	0.964	0.979	0.913	0.953	0.957	0.898	1.000	1.000	0.792	0.952	0.850

Architecture	QZS		QCW		QCS		XGW		XGS		AP	AR
	P	R	P	R	P	R	P	R	P	R		
ResNet34	0.920	0.995	0.847	0.934	0.951	0.894	0.821	0.774	0.947	0.930	0.911	0.892
ResNet34+SE	0.965	0.977	0.854	0.896	0.924	0.894	0.918	0.736	0.923	0.973	0.929	0.911
ResNet34+CBAM_basic	0.956	0.992	0.896	0.945	0.960	0.930	0.895	0.734	0.934	0.966	0.937	0.919
ResNet34+CBAM_conv	0.940	0.995	0.882	0.896	0.927	0.937	0.784	0.736	0.919	0.944	0.919	0.885
ResNet34+CBAM_basic_conv	0.987	0.972	0.805	0.973	0.979	0.838	0.907	0.808	0.947	0.980	0.930	0.933
ResNet34+SE+CBAM_basic	0.963	0.985	0.909	0.874	0.918	0.940	0.877	0.857	0.964	0.961	0.945	0.913

将所提网络与不同的神经网络进行了横向对比,选择经典的分类网络和最近三年提出的网络在 Plant_seed 数据集下训练分类,对比的网络训练环境与本实验训练的模型保持一致,保证对比的公正性,即训练的软硬件配置相同,epoch 都为 50,学习率为 0.0001, batch size 为 32,使用 Adam 优化器进行优化计算。训练结果如表 6 所示。

表 6 Plant_seed 数据集不同网络训练结果

Table 6 Different network training results of Plant_seed dataset

Architecture	Param / MB	Gflops	Accuracy / %
AlexNet ^[17]	14.6	3.96	88.76±0.84
GoogleNet ^[18]	5.99	10.14	87.10±1.30
RegNet_400mf ^[19]	4.30	0.40	94.63±0.97
EfficientNet_B0 ^[20]	5.30	0.40	86.57±1.13
EfficientNetV2_S ^[21]	24.00	8.80	94.85±0.82
ResNet34	21.79	11.75	90.69±1.61
ResNet34+CBAM_basic_conv	21.97	11.76	93.80±0.80
ResNet50	25.56	13.16	92.96±0.64
ResNet50+CBAM_bottle_conv	26.05	13.19	94.89±0.61

4.3 随机擦除法探究实验

最后,探讨了在穴盘图片进行随机擦除的处理后,网络模型的泛化能力如何变化。为了更好对比网络性能,设计两种方法进行探讨,首先对比了穴盘图片擦除前后网络模型分类的错误率,结果如表 7 所示。

表 7 模型分类错误率对比

Table 7 Comparison of model classification error rates

Architecture	Error / %
ResNet34	7.67±0.54
ResNet50	6.55±0.36
Proposed(ResNet34)	5.96±0.49
Proposed(ResNet50)	5.04±0.37
Proposed(ResNet34+erasing)	4.23±0.21
Proposed(ResNet50+erasing)	4.14±0.13

其次对分类输出图片的特征权重绘制了热力图,研究图片中对模型分类决策起决定性的区域,探索注意力机制对模型性能提升的本质原因,得到的结果如图 9 所示,图中红色区域代表分类决策权重分布范围。

4.4 结果与分析

从表 4 和图 8 可以看出,在 Plant_seed 数据集上,在 ResNet34 残差模块和 conv*_x 模块之间均引入的网络 ResNet34+CBAM_basic_conv 模型准确率提升均优于其他模型,相比于单独在残差块和 conv*_x 模块引入 CBAM 模块准确率分别提升 0.34 个百分点和 2.52 个百分点,这是由于残差块里面引入注意力机制能够更高效学习到图像样本的特征。同时 SENet 与 CBAM 结构融合于 ResNet34 后进行训练,相比于 ResNet34+SE,融合后的网络在准确率上也有 1.64 个百分点的准确率提升。这说明 SENet 与 CBAM 模块的融合也是提升网络性能的不错选择。但是对于 ResNet50 来说,注意力机制模块对网络的准确率提升是不如 ResNet34 的,这是因为注意力机制更容易在浅层网络学习到有效的特征,而对于深层网络的特征不敏感。同时在引入注意力模块之后,起始点 epoch 的准确率最大相差 10.8 个百分点,这说明所提网络的图像特征提取能力更高效。

从表 5 可以看出,不同的网络模型对不同种类的穴盘幼苗的分类准确率差异较大,但是从整体上看,在引入注意力模块后,均可提升大部分幼苗的分类准确率,如花菜弱苗和空穴的识别率均提升较大。ResNet34+CBAM_basic_conv 网络的平均召回率能够达到 93.3%,说明该网络对数据集的正样本预测准确能力较高。而 ResNet34+CBAM_basic_conv 网络的平均精确率表现更好,说明该网络预测正样本的能力更强。从表 6 可以看出:与传统的图像分类网络 AlexNet、GoogleNet 相比,所提注意力残差融合网络在准确率上提升较大;与近三年提出的新型分类网络对比,注意力残差融合网络模型均优于其他网络,这说明了基于注意力机制的改进网络模型的有效性。

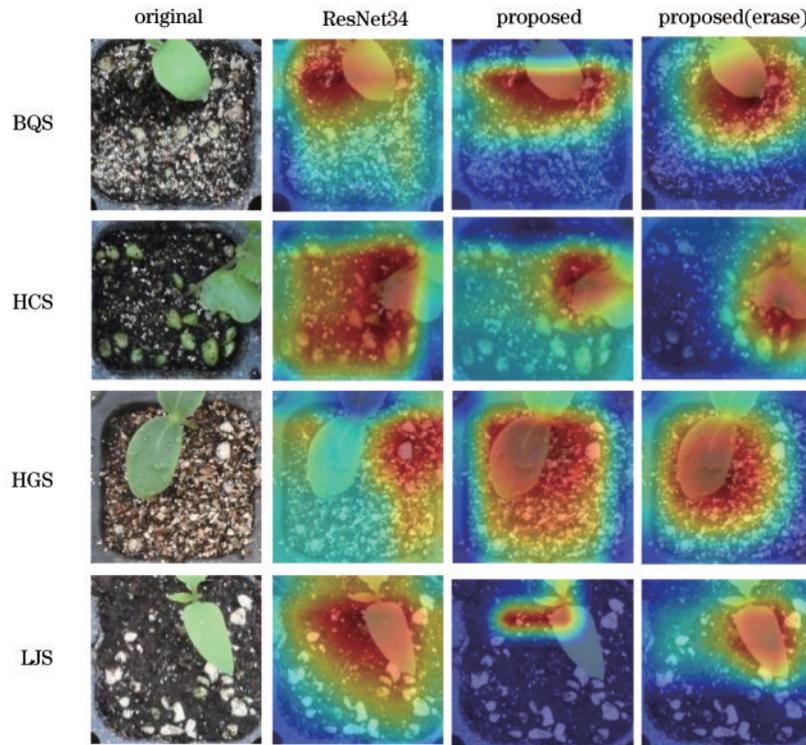


图 9 分类输出图片的特征权重热力图

Fig. 9 Feature weight heat map of classified output image

从表 7 可以看出,所提模型在对数据集图片进行随机擦除后可以有效降低分类错误率,即该网络可以提高缺失部分图片分类的准确性。同时从图 9 可以看出,所提注意力机制网络对穴盘图像识别的权重区域更加集中,这与注意力机制中更加关注分类特征区域的原理是一致的。

5 结 论

本研究基于图像分类技术,对农作物育苗过程中的幼苗分类进行了实际考察和研究,在现有的幼苗分类方法进行总结和改进,将注意力机制模块和残差模块进行融合设计了新的网络模型,同时针对穴盘图像分割后存在部分幼苗形态缺失的问题,提出的随机擦除的方法在穴盘幼苗分类任务上取得了较好的分类效果。实验结果表明:在残差模块引入注意力机制模块可以提高分类准确率和初始分类准确率,对特征的提取能力更强;注意力机制模块对越深的网络性能的提升效果越不明显,这是由于更深的网络特征更加抽象,注意力机制难以学到有效特征;对数据集使用随机擦除的方法可以有效解决穴盘图像部分缺失的问题,提高网络的泛化能力。

参 考 文 献

- [1] 肖超云. 基于机器视觉的杂草识别研究[D]. 镇江: 江苏大学, 2007.
Xiao C Y. The research on identification of weeds based on machine vision[D]. Zhenjiang: Jiangsu University, 2007.
- [2] 杨镇宇, 祝诗平. 基于机器视觉和 SVM 的花椒外观品

质检测[C]//全国虚拟仪器学术交流大会. 北京: 中国仪器仪表学会, 2009: 77-80.

Yang Z Y, Zhu S P. Appearance quality detection of *Zanthoxylum bungeanum* based on machine vision and SVM[C]//National Virtual Instrument Academic Exchange Conference. Beijing: China Instrument and Control Society, 2009: 77-80.

- [3] 何彦虎, 武传宇, 童俊华, 等. 基于专家系统的穴盘苗品种识别算法设计与试验[J]. 江苏农业科学, 2019, 47(4): 176-181.
He Y H, Wu C Y, Tong J H, et al. Design and experiment of identification algorithm of plug seedling based on expert system[J]. Jiangsu Agricultural Sciences, 2019, 47(4): 176-181.
- [4] Lameski P, Zdravevski E, Kulakov A. Weed segmentation from grayscale tobacco seedling images[M]//Rodić A, Borangiu T. Advances in robot design and intelligent control. Advances in intelligent systems and computing. Cham: Springer, 2016, 540: 252-258.
- [5] 刘合琛. 基于 CNN 和 GLCM 融合判别的全景植物识别方法[D]. 武汉: 湖北工业大学, 2020.
Liu H C. Panoramic plant recognition method based on CNN and GLCM fusion discrimination[D]. Wuhan: Hubei University of Technology, 2020.
- [6] 李金凤, 赵雨童, 黄纬然, 等. 基于灰度共生矩阵的多尺度分块压缩感知算法[J]. 激光与光电子学进展, 2021, 58(4): 0410002.
Li J F, Zhao Y T, Huang W R, et al. Multi-scale block compressed sensing algorithm based on gray-level co-occurrence matrix[J]. Laser & Optoelectronics Progress, 2021, 58(4): 0410002.
- [7] 侯晨伟, 陈丽. 基于概率神经网络的玉米苗期杂草识别

- 方法的研究[J]. 农机化研究, 2010, 32(11): 41-43, 47.
- Hou C W, Chen L. Weed identification method based on probabilistic neural network in the corn seedlings field[J]. Journal of Agricultural Mechanization Research, 2010, 32(11): 41-43, 47.
- [8] 张翠军, 赵娜. 基于概率神经网络改进的GrabCut算法[J]. 激光与光电子学进展, 2021, 58(2): 0210024.
- Zhang C J, Zhao N. Improved GrabCut algorithm based on probabilistic neural network[J]. Laser & Optoelectronics Progress, 2021, 58(2): 0210024.
- [9] 何艳, 夏志林, 祝诗平, 等. 基于机器视觉的穴盘烟苗自动间苗算法研究[J]. 云南农业大学学报(自然科学), 2019, 34(1): 50-57.
- He Y, Xia Z L, Zhu S P, et al. Automatic thinning algorithm of plug seedling based on machine vision[J]. Journal of Yunnan Agricultural University (Natural Science), 2019, 34(1): 50-57.
- [10] 杨旭, 尚振宏. 基于改进 AlexNet 的人脸表情识别[J]. 激光与光电子学进展, 2020, 57(14): 141026.
- Yang X, Shang Z H. Facial expression recognition based on improved AlexNet[J]. Laser & Optoelectronics Progress, 2020, 57(14): 141026.
- [11] 左羽, 陶情, 吴恋, 等. 基于卷积神经网络的植物图像分类方法研究[J]. 物联网技术, 2020, 10(3): 72-75.
- Zuo Y, Tao Q, Wu L, et al. Research on plant image classification method based on convolution neural network[J]. Internet of Things Technologies, 2020, 10(3): 72-75.
- [12] 娄月. 基于改进胶囊神经网络的植物识别方法研究[D]. 长春: 吉林农业大学, 2020.
- Lou Y. Research of plant recognition method based on improved capsule neural network[D]. Changchun: Jilin Agricultural University, 2020.
- [13] Hu J, Shen L, Albanie S, et al. Squeeze-and-excitation networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42(8): 2011-2023.
- [14] Huang G, Liu Z, Laurens V, et al. Densely connected convolutional networks[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition, July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 2261-2269.
- [15] 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, June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 770-778.
- [16] Woo S, Park J, Lee J Y, et al. CBAM: convolutional block attention module[M]//Ferrari V, Hebert M, Sminchisescu C, et al. Computer vision-ECCV 2018. Lecture notes in computer science. Cham: Springer, 2018, 11211: 3-19.
- [17] Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks[J]. Communications of the ACM, 2017, 60(6): 84-90.
- [18] Chollet F. Xception: deep learning with depthwise separable convolutions[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition, July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 1800-1807.
- [19] Radosavovic I, Kosaraju R P, Girshick R, et al. Designing network design spaces[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 13-19, 2020, Seattle, WA, USA. New York: IEEE Press, 2020: 10425-10433.
- [20] Tan M X, Le Q V. EfficientNet: rethinking model scaling for convolutional neural networks[C]//Proceedings of the 36th International Conference on Machine Learning, {ICML} 2019, June 9-15, 2019, Long Beach, California, USA. Cambridge: PMLR, 2019, 97: 6105-6114.
- [21] Tan M X, Le Q V. EfficientNetV2: smaller models and faster training[C]//Proceedings of the 38th International Conference on Machine Learning, {ICML} 2021, July 18-24, 2021, Virtual Event. Cambridge: PMLR, 2021, 139: 10096-10106.