

基于深度主动学习与CBAM的细粒度菊花表型识别

袁培森 丁毅飞 徐焕良

(南京农业大学人工智能学院, 南京 210095)

摘要: 针对菊花种类繁多,花型差别细微,准确标注比较困难的问题,基于深度主动学习与混合注意力机制模块(Convolutional block attention module,CBAM),提出了一种标号数据不足情况下的菊花表型智能识别方法和框架。首先,通过主动学习策略基于最优标号和次优标号法(Best vs second-best,BvSB)在未标记菊花样本中选取信息量较大的样本进行标记,并将标记后的样本放入训练样本中;其次,使用深度卷积神经网络 ResNet50 作为本文的主干网络训练标记样本,引入混合注意力机制模块 CBAM,使模型能够更为准确地提取细粒度图像中的高层语义信息;最后,用更新后的训练样本继续训练分类模型,直到模型达到迭代次数后停止。实验结果表明,该方法在少量菊花标记样本下,精确率、召回率和 F1 值分别达到 93.66%、93.15% 和 93.41%。本文方法可为标号数据不足情况下的菊花等花卉智能化识别提供技术支撑。

关键词: 菊花表型; 细粒度图像识别; 主动学习; ResNet50; 注意力机制模块

中图分类号: S4

文献标识码: A

文章编号: 1000-1298(2024)02-0258-10

OSID:



Fine-grained Chrysanthemum Phenotype Recognition Based on Deep Active Learning and CBAM

YUAN Peisen DING Yifei XU Huanliang

(College of Artificial Intelligence, Nanjing Agricultural University, Nanjing 210095, China)

Abstract: Chrysanthemums have a wide variety of flower types with subtle differences in flower phenotypes, which are difficult to label accurately, and this poses a great challenge for intelligent classification and recognition of chrysanthemums. Based on deep active learning and hybrid attention mechanism module, i. e. convolutional block attention module (CBAM), a method and framework for intelligent recognition of chrysanthemum phenotypes under insufficient labeling data was proposed. Firstly, the more informative samples among the unlabeled chrysanthemum samples were selected for labeling by an active learning strategy based on the optimal labeling and second-optimal labeling method BvSB (Best vs second-best), and the labeled samples were put into the training samples; secondly, a deep convolutional neural network ResNet50 was used as the backbone network to train the labeled samples, and the hybrid attention mechanism module CBAM was introduced, so that the model can more accurately extract the high-level semantic information in fine-grained images; finally, the classification model continued to be trained with the updated training samples until the model reached the number of iterations and then stopped. The experimental results showed that the method can achieve 93.66%, 93.15% and 93.41% of precision, recall and F1 value respectively with a small number of chrysanthemum labeled samples. The method can provide technical support for intelligent identification of chrysanthemums and other flowers under the situation of insufficient labeling data.

Key words: chrysanthemum phenotype; fine-grained image recognition; active learning; ResNet50; attention mechanism module

0 引言

植物表型是由基因和环境条件决定的,表示

植物的结构及组成、生长过程及结果的全部特征和性状^[1]。菊花是菊科菊属多年生草本植物,更是中国十大传统名花之一。菊花种类繁多,目前

收稿日期: 2023-07-12 修回日期: 2023-09-10

基金项目: 国家自然科学基金项目(61502236)、国家级创新训练专项项目(202310307095Z)和江苏省研究生实践创新计划项目(SJCX23_0203)

作者简介: 袁培森(1980—),男,副教授,博士,主要从事智能数据处理与服务研究,E-mail: peiseny@njau.edu.cn

已知的菊花品种超过 7 000 种^[2]。根据菊花表型特征识别其品种属于细粒度图像识别任务^[3],细粒度图像识别任务与一般图像识别任务相比对象的粒度更小,只能根据非常细微的差别才能进行分类,如动物或植物的种类识别任务等。菊花表型的复杂性使得某些菊花类内差异大,但类间相似性高,因此菊花图像表型识别是一个巨大的挑战,需要多年的菊花培育经验和知识积累,且识别过程耗时耗力,所以基于经验的菊花表型识别方法并不适用于规模庞大的识别任务^[4]。

随着深度学习^[5]的兴起,不少学者使用深度神经网络实现花卉表型智能识别,利用搭建好的大规模处理层的人工神经网络自主学习花卉样本中的表型特征与内在规律。WANG 等^[6]建立了一个包含 213 个品种的菊花数据集,使用 ResNet18 作为骨干网络,实现对多品种菊花的准确识别。QI 等^[7]提出一种轻量级茶菊花检测深度学习模型 TC - YOLO,完成开花期茶菊花的自动检测。徐光柱等^[8]提出一种基于轻量级神经网络的花卉图像分类方法,利用重量级网络辅助训练轻量级网络,实现对花卉图像的高效分类。齐振岭等^[9]构建了牡丹花卉图像数据集,应用纯卷积神经网络 ConvNeXt - Tiny 模型算法实现多品种牡丹花卉的分类与识别。

常规的深度学习模型虽然能够取得较好的识别结果,但不能快速聚焦图像中重要和敏感的区域。注意力机制模块^[10]通过给图像中每个通道赋予权值的方法,关注图像中更为重要的特征信息,忽略无关特征信息。项和雨等^[11]提出一种新的残差注意力神经网络 RAN - 11,该网络以残差注意力网络 Attention - 56 和 Attention - 92 为基础构建,完成对太湖中浮游植物的准确识别。尚远航等^[12]提出一种基于混合注意力机制深度残差网络的植物病害识别方法,实现对多种植物病害的分类与识别。

然而,上述研究都是在标签充裕的情况下进行的,当数据集较为庞大但研究人员不具备充足的标签成本时,会面临标签数据集不够导致模型过拟合的情况。唐英鹏等^[13]提出使用主动学习的方法解决上述问题,该学习方法在标签成本不足时,选择样本集中更具备标注价值的样本进行专家标注,从而更加高效的完成识别任务^[14]。主动学习通过选择对模型信息量最大的标注样本来帮助降低标注成本,从而以较少的标注提高模型性能。赵小康等^[15]提出一种基于无监督主动学习方法的训练样本选择策略。利用 2 个年份 3 个品种苹果的近红外光谱数据,实现更为合理与高效的苹果样本选择。YANG 等^[16]提出一种基于类间相似度分析的图像信息贡

献度评价方法,该方法与主动学习图像选择策略相结合,可以为植物病害数据集的智能采集和标注提供指导。

上述研究结果虽然较为高效,但研究对象大多采用公开数据集,且数据集中类别差异较大。针对本文细粒度菊花数据集种类较多且类别之间差异性较小、标注难度较大等特点,使用网络层次更深的残差块网络 ResNet50 作为主干网络,引入混合注意力机制模块(CBAM)提取细粒度菊花图像高层语义信息,采用主动学习策略(BvSB)选择更具标注价值的样本以节省标注成本,实现对菊花表型的智能识别。

1 实验材料

菊花图像数据集采集自江苏省南京市湖熟菊花基地,北纬 31. 96°,东经 118. 84°。采集时间为 2019 年 11 月。共 20 种菊花品种,其名称分别为虎须、天雨流星、孩儿面、太平的静观、巧堆腊雪、仙姿、红露仙衣、鹏程万里、南农金鹂、巴卡红、淡黄、白蜂窝、光玉、秦淮白茶、白皇后、杏花春雨、墨宝、晴光、桔丹特、溧水金波。数据集见表 1。

表 1 菊花图像数据集			
Tab. 1	Chrysanthemum image data set		幅
名称	训练数量	验证数量	测试数量
虎须	81	10	10
天雨流星	76	9	9
孩儿面	84	10	10
太平的静观	80	9	9
巧堆腊雪	80	9	9
仙姿	80	9	9
红露仙衣	85	10	10
鹏程万里	81	10	10
南农金鹂	76	9	9
巴卡红	100	12	12
淡黄	102	12	12
白蜂窝	72	9	9
光玉	80	9	9
秦淮白茶	86	10	10
白皇后	100	12	12
杏花春雨	87	10	10
墨宝	80	9	9
晴光	80	9	9
桔丹特	96	11	11
溧水金波	81	10	10

在实验中将每类菊花表型数据集按照比例 8: 1: 1划分为训练集、验证集和测试集,如表 1 所示。共 20 种不同类型的菊花,总计 1 985 幅菊花图像,不同种类菊花图像如图 1 所示。



图 1 菊花图像

Fig. 1 Chrysanthemum images

2 网络框架

本文提出的基于深度主动学习与注意力机制的菊花表型识别方法,充分利用深度学习的强大分类能力以及主动学习的标记效率。本文方法网络框架图如图 2 所示。

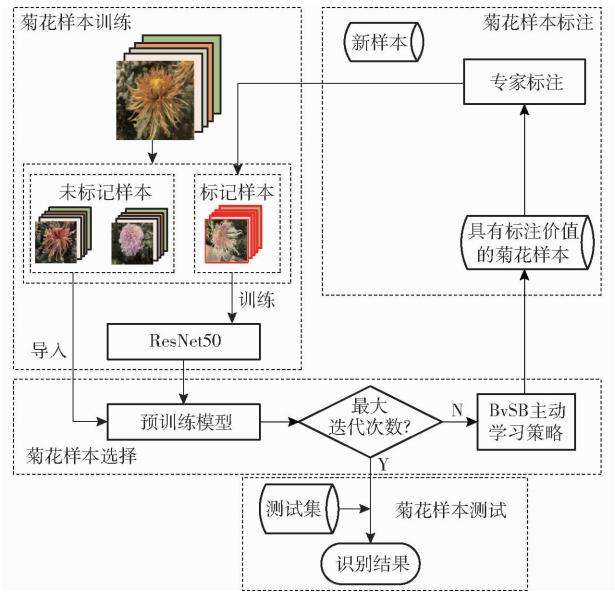


图 2 本文方法网络框架图

Fig. 2 Network frame diagram of the proposed method

网络框架共分为 4 个模块,相较于传统的深度学习训练模型主要改进在于第 2 和第 3 模块。改进

如下:①菊花样本训练模块,将预处理后的菊花图像作为网络的输入,其中有标记的样本作为训练集,用来训练 ResNet50 模型,而未标记的样本则作为主动学习的候选池。②菊花样本选择模块,将候选池中的未标记样本全部输入到训练好的 ResNet50 模型中,基于模型最后一层的 Softmax 函数输出得到候选池中每个样本的类别概率,使用主动学习策略 BvSB 选择其中更具标注价值的样本数据。③菊花样本标注模块,对上一模块选择的具有标注价值的菊花样本进行专家标注,并将新标记的样本加入到训练集中,对神经网络进行新一轮的迭代训练,重复以上步骤,直至达到最大迭代次数为止。④菊花样本测试模块,将菊花测试集输入到最终训练好的 ResNet50 模型中得到测试结果。

2.1 注意力机制模块

菊花图像识别属于细粒度识别^[17]问题,引入注意力模块^[10]可以让模型快速聚焦于菊花图像中的重要与敏感区域,极大程度加快特征的提取并简化分类过程。

在深度学习中,注意力机制是一种机器视觉机制,该机制通过模仿人脑处理视觉信息的思维方式,可以将注意焦点迅速聚焦于图像中的重要区域,而忽略其它无关背景区域,从而更加高效地处理视觉信息^[18]。注意力机制主要分为通道注意力机制^[19]

与空间注意力机制^[20]。前者为每个通道都赋予权值,从而提升重要的通道特征,忽略无关通道信息。后者重视图像空间中的信息交互,突显关键区域的特征表示。为了综合二者的优势,本文使用一种混合注意力机制 (Convolutional block attention module, CBAM)^[21],该机制融合了空间注意力机制与通道注意力机制各自的优势,形成一种综合的注意力机制方法,如图 3 所示。

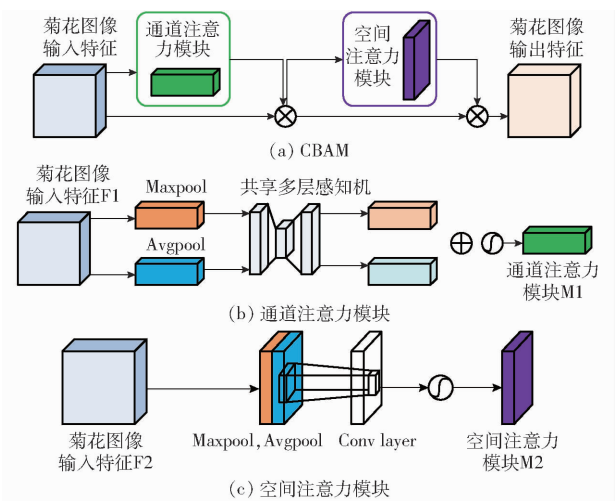


图 3 CBAM 结构图

Fig.3 Structure diagram of CBAM

CBAM 整体结构如图 3a 所示,CBAM 在原通道注意力模块的基础上连结了一个空间注意力模块。其处理流程如下:①CBAM 内的通道注意力模块如图 3b 所示,在通道注意力模块中,输入的菊花特征图先分别经过一个最大池化层和一个平均池化层,得到 2 个一通道的特征图;然后在共享多层感知机中,把输出的特征图做加和操作;最后使用 Sigmoid 函数做激活操作,得到最终输出的菊花图像通道注意力特征。将该特征和初始输入特征图做乘法操作,得到空间注意力模块所需要的菊花图像输入特征。②CBAM 内的空间注意力模块如图 3c 所示,在空间注意力模块中,输入的菊花特征图先后经过一个最大池化层和一个平均池化层,得到 2 个一通道的特征图;将所得特征图做通道拼接后,经过一个卷积层降维成一通道特征图;最后使用 Sigmoid 函数激活操作生成最终的菊花图像空间注意力特征。③完成上述菊花图像特征提取操作后,将菊花图像的空间注意力特征与空间注意力模块的菊花图像输入特征做乘法,得到最后需要的菊花图像输出特征。

2.2 融合 CBAM 的 ResNet50

本文方法的主干网络使用融合了混合注意力机制模块 CBAM 的 ResNet50 模型^[22]。ResNet50 网络结构如表 2 所示。

表 2 ResNet50 结构
Tab.2 Structure of ResNet50

网络层	卷积核
Conv1	7 × 7,64 卷积层
	3 × 3 最大池化层
Conv2	$\begin{pmatrix} 1 \times 1,64 \\ 3 \times 3,64 \times 3 \\ 1 \times 1,256 \end{pmatrix} \times 3$
Conv3	$\begin{pmatrix} 1 \times 1,128 \\ 3 \times 3,128 \times 4 \\ 1 \times 1,512 \end{pmatrix} \times 4$
Conv4	$\begin{pmatrix} 1 \times 1,256 \\ 3 \times 3,256 \times 6 \\ 1 \times 1,1\,024 \end{pmatrix} \times 6$
Conv5	$\begin{pmatrix} 1 \times 1,512 \\ 3 \times 3,512 \times 3 \\ 1 \times 1,2\,048 \end{pmatrix} \times 3$
平均池化层	
Softmax	

ResNet50 神经网络由 5 个 Conv 层、2 个池化层和 1 个 Softmax 激活层组成,Conv 层由 Bottleneck 组成。而本文使用的模型在每一个卷积块后面添加了一个混合注意力机制模块 CBAM,本文模型结构图如图 4 所示。

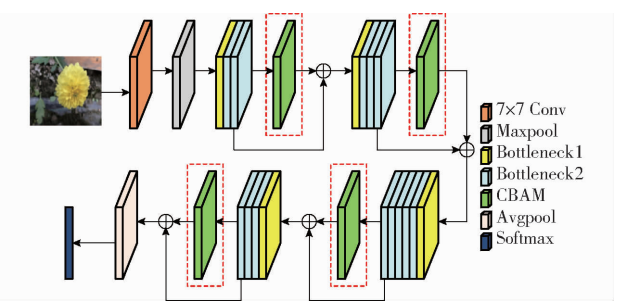


图 4 融合 CBAM 的 ResNet50 结构图

Fig.4 Structure diagram of ResNet50 integrated CBAM

ResNet50 模型共分为 5 个阶段,第 1 阶段特征图经过一个卷积核大小为 7 × 7 的卷积层后,输入到一个全局最大池化层中,得到输出特征;第 2 至第 5 阶段特征图分别经过数量为 3、4、6、3 的 Bottleneck 后得到输出特征,其中单个 Bottleneck 的结构如图 5 所示。

每一个 Bottleneck 中存在 3 个卷积核大小分别为 1 × 1、3 × 3、1 × 1 的卷积层,通过加入一条残差边的方式可以跳过训练效果不好的网络层以及避免梯度爆炸问题,通过卷积计算得到每一层卷积层输出的特征图。特征图计算公式为

$$f_i^l = \sigma(W^l f_i^{l-1} + b^l) \tag{1}$$

式中 i ——菊花类别序号

l ——卷积层序号

f_i^l ——第 l 层卷积层输出的第 i 个类别的菊

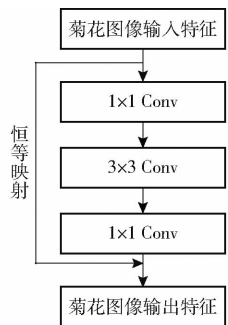


图5 Bottleneck 结构图
Fig.5 Structure diagram of bottleneck

花特征图

σ ——ReLU 激活函数

W^l ——第 l 层卷积层权重

b^l ——第 l 层卷积层偏置

使用的 4 个 CBAM 分别放置于网络中第 2 至第 5 阶段中每个阶段最后一个 Bottleneck 的后方,其作用为再次提取菊花图像特征,以聚焦图像中较为重要的特征信息,将其输出的特征与原特征做加和操作,从而获得每一阶段的最终输出特征。

平均池化层负责做相关池化操作,减小特征图尺寸,提取图中更多通道数,即把图中更多的信息提取出来。全连接层负责整合池化层和卷积层中能够区分菊花类别的相关信息,通过 ReLU 激活函数激活,最后利用 Softmax 函数对菊花图像分类,得到每幅菊花图像最终所属类别。类别计算公式为

$$\text{Softmax}(p_i) = \frac{e^{p_i}}{\sum_{i=1}^C e^{p_i}} \tag{2}$$

式中 C ——菊花类别总数

p_i ——第 i 个菊花类别的输出概率

使用交叉熵损失函数优化网络参数,其计算公式为

$$Loss = -\frac{1}{M} \sum_{c=1}^M \sum_{c=1}^C y_c \lg \hat{y}_c \tag{3}$$

式中 $Loss$ ——交叉熵损失函数

c ——菊花类别

$y_c, \hat{y}_c$ ——实际类别和预测类别

M ——训练的菊花样本总数

2.3 主动学习策略

主动学习算法中最重要的部分就是主动学习策略的选择,选择合适的查询策略可以挑选出数据集中最具标记价值的样本,从而提高模型分类准确性。而不确定性采样法是主动学习策略中使用较多的方法,本文将基于最小置信度 (Minimum confidence, MC)^[23]、基于最优标号和次优标号法 (Best vs second-best, BvSB)^[24] 和基于熵采样 (Entropy

sampling, ES)^[25] 3 种基于不确定性采样的主动学习策略与随机采样法 (Random sampling, RS)^[26] 进行对比实验。

2.4 BvSB 采样

BvSB 采样法适用于一些容易分成两类的样本。该方法选择模型最后一层 Softmax 函数输出的类别最大概率和次大概率之差最小的样本进行标注,因为这些样本在最大概率和次大概率的两个类别之中极易分类错误,所以具有很高的标注价值。基于 BvSB 的采样法公式为

$$P_{\text{BvSB}} = \min(P_{\text{max}} - P_{\text{second-max}}) \tag{4}$$

式中 P_{BvSB} ——BvSB 采样法

P_{max} ——最大概率

$P_{\text{second-max}}$ ——次大概率

本文方法流程图如图 6 所示。

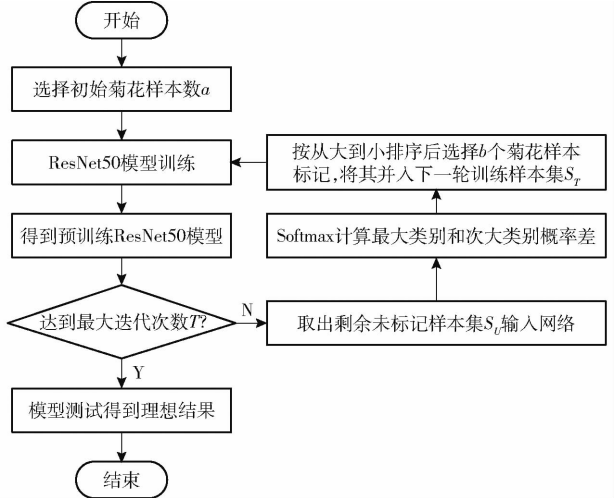


图6 BvSB 采样法流程图
Fig.6 BvSB sampling flowchart

首先选择初始样本数将其标记,并定好迭代次数,每次标记样本数量等参数,然后将初始样本放入本文模型中进行训练。若未达到最大迭代次数,则取出剩余的未标记样本输入预训练的模型,通过 Softmax 函数计算得到每个样本最大类别和次大类别概率差,将所有未标记样本按该值从大到小排列,根据每轮的标记样本数标记下一轮需要训练的样本;若达到最大迭代次数,则把准备好的测试集放入训练好的 ResNet50 模型中进行测试,得到最终结果。BvSB 算法流程如下:

输入:训练样本集 S_T , 未标记样本集 S_U , 标记轮次 T , 首次随机选取样本数 a , 每次主动选取样本数 b ;
初始化: $t = 0$

- (1) 随机选择 x 个标记样本, 构造初始训练样本集 S_T ;
- (2) While $t < T$ do;
- (3) 使用 S_T 训练本文模型;

(4)使用训练好的模型计算 S_U 中的每个样本的类别概率;

(5)主动选择 b 个样本进行标记;

(6) $S_U = S_U - b$;

(7)EndWhile;

输出:菊花类别。

3 实验与结果分析

3.1 实验环境

实验平台为 Windows 11,64 位操作系统,CPU 为 i7-12700H,2.70 GHz,内存为 16 GB。GPU 为 NVIDIA Geforce GTX 3060,内存为 8 GB。

3.2 数据预处理

为充分利用菊花图像中庞大的语义信息,需先对其进行预处理操作。本文图像预处理步骤为:①所有菊花图像统一转换为 3 通道的 RGB 图像。②所有菊花图像尺寸调整为 224 像素×224 像素。③调整后菊花图像以中心为原点,进行随机平移、旋转、剪裁等数据增强操作。④对所有菊花图像进行正则化,从而预防模型过拟合。

3.3 实验参数与评价标准

菊花样本类别为 20,采用主动学习策略每次标记 500 份样本,总共标记 4 轮,批量大小设置为 4,每标记一轮样本,网络训练迭代 100 轮,初始学习率设置为 10^{-4} ,优化器使用 Adam。

采用精确率、召回率和 F1 值作为评价指标来衡量模型识别效果。

3.4 实验结果分析

3.4.1 不同查询策略对识别效果的影响

为验证不同主动学习查询策略对模型识别效果的影响,使用 MC 采样^[23]、BvSB 采样^[24]、ES^[25]和 RS^[26]4 种主动查询策略,对比它们在本文菊花数据集上的识别效果,实验结果如表 3 所示。

表 3 不同主动学习查询策略的实验结果

Tab.3 Results of experiments on different active learning query strategies

查询策略	精确率	召回率	F1 值
BvSB	93.66	93.15	93.41
RS	90.14	90.45	90.61
ES	92.84	91.63	92.23
MC	93.08	92.54	92.81

由表 3 可知,BvSB 查询策略在本文数据集上取得了最好的结果,精确率达到 93.66%,召回率达到 93.15%,F1 值达到 93.41%。较随机采样策略在精确率、召回率和 F1 值上分别提高 3.52、2.7、2.8 个百分点。这是由于本文菊花数据集的类别中存在表型

性状相似的情况,例如第 1 类虎须和第 2 类天雨流星。BvSB 针对这些容易分成两类的样本可以取得更好的采样效果;相比于随机采样法,其它 3 种查询策略都取得了更好的效果。这是因为随机采样法从样本中随机抽取样本进行标注,而其它 3 种策略可以根据数据集特征选取更具备标注价值的样本进行标注,从而为模型提供贡献率更大的训练样本,取得更好的识别效果。综上,本文后续实验都将选用 BvSB 查询策略作为本文的最佳主动学习查询策略。

3.4.2 不同注意力模块对识别效果的影响

为验证不同注意力模块对本文模型识别效果的影响,将 CBAM^[21]、CAM (Channel attention module)^[27]、SAM (Spatial attention module)^[28]和 ECA (Efficient channel attetion module)^[29]4 种注意力模块分别插入主干网络 ResNet50 中,主动学习策略采用 BvSB 采样,实验结果如表 4 所示。

表 4 不同注意力模块的实验结果

Tab.4 Experiment results of different attention modules

	attention modules			%
注意力模块	精确率	召回率	F1 值	
CBAM	93.66	93.15	93.41	
CAM	91.55	90.87	91.21	
SAM	91.62	91.23	91.42	
ECA	93.15	92.96	93.05	

由表 4 可知,采用混合注意力模块 CBAM 在本文数据集上取得了最好的结果,精确率达到 93.66%,召回率达到 93.15%,F1 值达到 93.41%。较 CAM 在精确率、召回率和 F1 值上分别提高 2.11、2.8、2.2 个百分点;较 SAM 在精确率、召回率和 F1 值上分别提高 2.04、1.92、1.99 个百分点;较 ECA 在精确率、召回率和 F1 值上分别提高 0.51、0.19、0.36 个百分点。这是因为相比于其它 3 种注意力模块只考虑一种特征信息,CBAM 既考虑到了菊花图像空间特征,又兼顾其通道特征。所以本文采用 CBAM 作为主干网络 ResNet50 的注意力模块。

3.4.3 初始学习率参数对识别效果的影响

为验证不同初始学习率 η 对本文方法识别效果的影响,将初始学习率 η 分别设为 10^{-2} 、 10^{-3} 、 10^{-4} 和 10^{-5} ,4 种不同初始学习率 η 的损失值曲线如图 7 所示。

由图 7 可知,较低的初始学习率 η 会使模型收敛速度变得较为缓慢,而较高的初始学习率 η 虽然可以使得模型较快收敛,但最终收敛时的损失值无法达到最佳。当初始学习率 η 为 10^{-4} 时,模型不仅可以较快收敛,且收敛后的损失值最小。实验结果如表 5 所示。

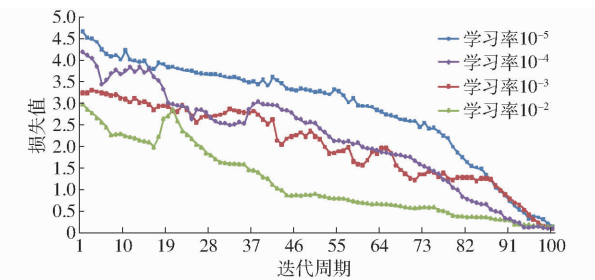


图 7 不同初始学习率的损失值曲线

Fig. 7 Loss graph of different initial learning rates

表 5 初始学习率对识别效果的影响

Tab. 5 Effect of initial learning rate on recognition			
%			
学习率 η	精确率	召回率	F1 值
10^{-2}	92.78	92.46	92.62
10^{-3}	93.21	92.85	93.03
10^{-4}	93.66	93.15	93.41
10^{-5}	92.87	93.09	92.98

由表 5 可知,当初始学习率 η 为 10^{-4} 时,本文方法的识别效果最佳,精确率达到 93.66%,召回率达到 93.15%,F1 值达到 93.41%。相比于初始学习率 η 为 10^{-2} 时,在精确率、召回率和 F1 值上效果分别提高 0.88、0.69、0.79 个百分点;相比于初始学习率 η 为 10^{-3} 时,在精确率、召回率和 F1 值上分别提高 0.45、0.3、0.38 个百分点;相比于初始学习率 η 为 10^{-5} 时,在精确率、召回率和 F1 值上分别提高 0.79、0.06、0.43 个百分点。综上,本文初始学习率 η 确定为 10^{-4} 。

3.4.4 优化器对识别效果的影响

为验证不同优化器对本文方法识别效果的影响,分别使用随机梯度下降 (Stochastic gradient descent, SGD)^[30]、SGDM^[31]、AdaGrad^[32] 和 Adam^[19] 4 种优化器来优化网络参数。实验结果如表 6 所示。

表 6 优化器对识别效果的影响

Tab. 6 Effect of optimizer on recognition			
%			
优化器	精确率	召回率	F1 值
Adam	93.66	93.15	93.41
SGD	93.54	93.02	93.28
SGDM	92.49	92.89	92.69
AdaGrad	92.97	92.47	92.72

由表 6 可知,Adam 优化器在精确率、召回率和 F1 值 3 项评价指标上的表现最佳,优于其它 3 种优化器,分别为 93.66%、93.15% 和 93.41%,且模型训练时间也是 4 种优化器中最短的。所以本文优化器选择 Adam。

3.4.5 消融实验

通过消融实验来验证注意力模块和主动学习策略对模型识别效果的影响。将本文使用的 ResNet50 模型作为基准网络,将加入注意力模块的 ResNet50 模型记为 ResNet50-CBAM, ResNet50-AL 为使用主动学习查询策略的基准网络,而本文方法 (ResNet50-CBAM-AL) 为既使用主动学习查询策略,又加入注意力模块 CBAM 的 ResNet50 模型。对于非主动学习方法 ResNet50, ResNet50-CBAM 使用随机采样法 (RS) 进行样本选择,采用主动学习方法的 ResNet50-AL 和 ResNet50-CBAM-AL 使用 BvSB 采样法进行样本选择。4 种方法损失值曲线如图 8 所示。

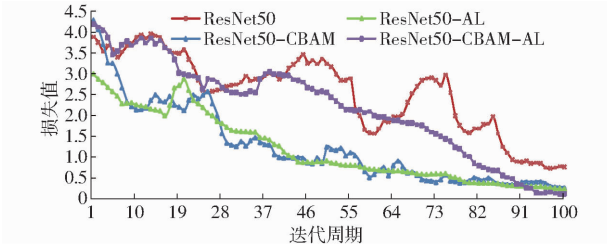


图 8 消融实验损失值曲线

Fig. 8 Loss graph of ablation study

由图 8 可知,直接使用 ResNet50 模型进行训练时损失值波动最大,且最后模型收敛时损失值最大;使用主动学习策略的模型相比于未使用主动学习策略的模型最终收敛时损失值更小,曲线更为平滑,这可能是因为主动学习策略选择更有标注价值的样本,相比于随机采样法更具目的性与高效性。本文方法最终收敛时损失值最小。4 种模型实验结果如表 7 所示。

表 7 消融实验结果

Tab. 7 Results of ablation experiment			
%			
模型	精确率	召回率	F1 值
本文方法	93.66	93.15	93.41
ResNet50	88.12	87.25	87.68
ResNet50-CBAM	90.14	90.45	90.29
ResNet50-AL	91.24	90.56	90.89

由表 7 可知,本文方法表现最佳。精确率达到 93.66%,召回率达到 93.15%,F1 值达到 93.41%。相比于 ResNet50 在精确率、召回率和 F1 值上分别提高 5.54、5.9、5.73 个百分点。相比于 ResNet50-CBAM,在精确率、召回率和 F1 值上分别提高 3.52、2.7、3.12 个百分点;相比于 ResNet50-AL,在精确率、召回率和 F1 值上分别提高 2.42、2.59、2.52 个百分点。说明本文方法是可行且有效的。

3.4.6 不同网络和本文方法对比

为验证本文方法的识别性能,将本文方法与

GoogleNet^[30]、SqueezeNet^[31] 和 VGG16^[32] 3 种深度卷积神经网络进行对比实验。其它网络模型使用随机采样法,实验结果如表 8 所示。

表 8 不同网络和本文方法实验结果

Tab. 8 Results of contrast of different networks and methods			%
模型	精确率	召回率	F1 值
本文方法	93. 66	93. 15	93. 41
GoogleNet	62. 47	62. 81	62. 64
SqueezeNet	82. 45	81. 89	82. 17
VGG16	88. 14	87. 66	87. 89

由表 8 可知,本文方法在精确率、召回率和 F1 值 3 项评价指标上都明显优于其它 3 种深度卷积神经网络。相比于 GoogleNet,本文方法在精确率、召回率和 F1 值上分别提高 31. 19、30. 34、30. 77 个百

分点;相比于 SqueezeNet,本文方法在精确率、召回率和 F1 值上分别提高 11. 21、11. 26、11. 24 个百分点;相比于 VGG16,本文方法在精确率、召回率和 F1 值上分别提高 5. 52、5. 49、5. 52 个百分点。说明本文方法与其它主流深度神经网络相比也具备较大优势。

3. 4. 7 模型特征可视化

目前,菊花表型智能识别过程缺乏足够的可解释性。通过梯度加权类激活映射(Gradient-weighted class activation mapping,Grad-CAM)^[33]绘制热力图与原图叠加,可较为直观展示网络是否学习到菊花图像的关键特征。网络经过多层卷积后得到特征层与预测值,再通过反向传播得到梯度信息并计算特征层中不同通道的权重,进行加权求和后经过激活函数输出得到 Grad-CAM。4 种模型的 Grad-CAM 图如图 9 所示。

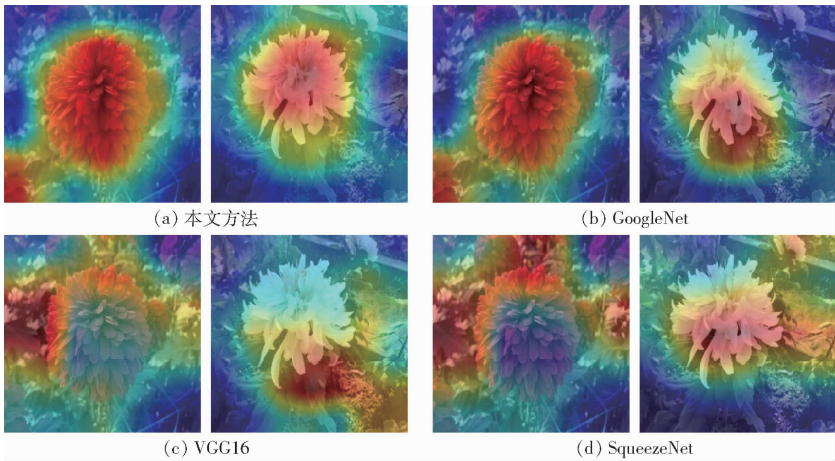


图 9 4 种模型的 Grad-CAM 图

Fig. 9 Grad-CAM diagram of four models

图 9 选取了最具代表性的 2 种菊花类型的 Grad-CAM 图,其可通过热力图的形式表现哪些区域的像素点对于模型输出具有更大的影响,颜色越亮,正向响应度越高。可见,本文提出的方法所生成的热力图更符合样本真实情况。原因在于:①本文方法使用主动学习策略选择更具标注价值的样本进行训练,能够更为准确地关注到不同种类菊花的表型特征。②本文方法融合了混合注意力模块 CBAM 以关注样本中的关键信息,更容易聚焦对样本有正影响的像素点激活区域。因此本文方法真实性更强,更好地解释了模型的识别过程。

4 结论

(1)提出一种基于深度主动学习与 CBAM 的细粒度菊花表型识别方法,结合 BvSB 主动学习策略与混合注意力机制模块 CBAM 实现菊花图像的智能识别。在本文菊花数据集上均优于其它深度学习模型,其中精确率、召回率以及 F1 值分别为 93. 66%、93. 15% 和 93. 41%。

(2)本文提出的方法可应用于标注成本高、标注难度较大的数据集,对细粒度花卉图像识别具有一定的参考性与示范性。

参 考 文 献

[1] ZHANG Q, LI J, WEN X, et al. Genome-wide identification and characterization analysis of RWP-RK family genes reveal their role in flowering time of *Chrysanthemum lavandulifolium*[J]. BMC Plant Biology, 2023, 23(1): 197-209.

[2] ZHANG Q, LI J, DENG C, et al. The mechanisms of optimal nitrogen conditions to accelerate flowering of *Chrysanthemum vestitum* under short day based on transcriptome analysis[J]. J. Plant Physiology, 2023, 285(1): 1-17.

[3] 朱丽,王新鹏,付海涛,等. 基于注意力机制的细粒度图像分类[J]. 吉林大学学报(理学版), 2023, 61(2): 371-376.

ZHU Li, WANG Xinpeng, FU Haitao, et al. Fine-grained image classification based on attention mechanism[J]. Journal of

- Jilin University (Science Edition), 2023, 61(2): 371–376. (in Chinese)
- [4] 袁培森,任守纲,翟肇裕,等. 基于半监督主动学习的菊花表型分类研究[J]. 农业机械学报, 2018, 49(9): 27–34.
YUAN Peisen, REN Shougang, ZHAI Zhaoyu, et al. Chrysanthemum phenotypic classification based on semi-supervised active learning[J]. Transactions of the Chinese Society for Agricultural Machinery, 2018, 49(9): 27–34. (in Chinese)
- [5] XIONG J, YU D, LIU S, et al. A review of plant phenotypic image recognition technology based on deep learning[J]. Electronics, 2021, 10(1): 1–19.
- [6] WANG J, TIAN Y, ZHANG R, et al. Multi-information model for large-flowered chrysanthemum cultivar recognition and classification[J]. Frontiers in Plant Science, 2022, 13(1): 1–13.
- [7] QI C, GAO J, PEARSON S, et al. Tea chrysanthemum detection under unstructured environments using the TC-YOLO model[J]. Expert Systems with Applications, 2022, 193(1): 1–13.
- [8] 徐光柱,朱泽群,尹思璐,等. 基于轻量级深层卷积神经网络的花卉图像分类系统[J]. 数据采集与处理, 2021, 36(4): 756–768.
XU Guangzhu, ZHU Zequn, YIN Silu, et al. Flower image classification system based on lightweight DCNN[J]. Journal of Data Acquisition and Processing, 2021, 36(4): 756–768. (in Chinese)
- [9] 齐振岭,刘羿璇,梁允泉,等. 基于 ConvNeXt-Tiny 的牡丹花卉分类识别研究[J]. 计算机时代, 2023, 2(1): 16–20.
QI Zhenling, LIU Yixuan, LIANG Yunquan, et al. Classification and recognition of peony flowers based on ConvNeXt-Tiny[J]. Computer Era, 2023, 2(1): 16–20. (in Chinese)
- [10] GUO N, GU K, QIAO J, et al. Improved deep CNNs based on nonlinear hybrid attention module for image classification[J]. Neural Networks, 2021, 140(1): 158–166.
- [11] 项和雨,邹斌,唐亮,等. 基于残差注意力网络模型的浮游植物识别[J]. 生态学报, 2021, 41(17): 6883–6892.
XIANG Heyu, ZOU Bin, TANG Liang, et al. Phytoplankton recognition based on residual attention network[J]. Acta Ecologica Sinica, 2021, 41(17): 6883–6892. (in Chinese)
- [12] 尚远航,余游江,吴刚. 基于混合注意力机制的植物病害识别[J]. 塔里木大学学报, 2021, 33(2): 94–103.
SHANG Yuanhang, YU Youjiang, WU Gang. Plant diseases recognition based on mixed attention mechanism[J]. Journal of Tarim University, 2021, 33(2): 94–103. (in Chinese)
- [13] 唐英鹏,黄圣君. 主动学习研究进展[J]. 中国基础科学, 2022, 24(3): 18–26, 34.
TANG Yingpeng, HUANG Shengjun. Recent progress on active learning[J]. China Basic Science, 2022, 24(3): 18–26, 34. (in Chinese)
- [14] 江显森. 图像分类中的深度主动学习研究综述[J]. 长江信息通信, 2021, 34(3): 161–163.
JIANG Xiansen. A survey of deep active learning in image classification[J]. Yangtze River Information and Communication, 2021, 34(3): 161–163. (in Chinese)
- [15] 赵小康,赵鑫,朱启兵,等. 一种基于无监督主动学习的苹果品质光谱无损检测模型构建方法[J]. 光谱学与光谱分析, 2022, 42(1): 282–291.
ZHAO Xiaokang, ZHAO Xin, ZHU Qibing, et al. A model construction method of spectral nondestructive detection for apple quality based on unsupervised active learning[J]. Spectroscopy and Spectral Analysis, 2022, 42(1): 282–291. (in Chinese)
- [16] YANG J, YANG Y, LI Y, et al. Image information contribution evaluation for plant diseases classification via inter-class similarity[J]. Sustainability, 2022, 14(17): 1–13.
- [17] WEI X, SONG Y, AODHA O M, et al. Fine-grained image analysis with deep learning: a survey[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022, 44(12): 8927–8948.
- [18] 任欢,王旭光. 注意力机制综述[J]. 计算机应用, 2021, 41(增刊1): 1–6.
REN Huan, WANG Xuguang. Review of attention mechanism[J]. Journal of Computer Applications, 2021, 41(Supp.1): 1–6. (in Chinese)
- [19] SHAN X, SHEN Y, CAI H, et al. Convolutional neural network optimization via channel reassessment attention module[J]. Digital Signal Processing, 2022, 123(1): 1–9.
- [20] ZHANG X, WANG Z. Spatial proximity feature selection with residual spatial-spectral attention network for hyperspectral image classification[J]. IEEE Access, 2023, 11(1): 23268–23281.
- [21] YANG H, GAO C, GE H, et al. Risk assessment of aviation DC series arc based on reconstructed CBAM-CNN[J]. Journal of Power Electronics, 2023, 23(5): 811–820.
- [22] SHARMA A K, NANDAL A, DHAKA A, et al. Enhanced watershed segmentation algorithm-based modified ResNet50 model for brain tumor detection[J]. BioMed Research International, 2022, 1(1): 1–14.
- [23] LUO Y, NIE Q, YANG D, et al. Robust optimal operation of active distribution network based on minimum confidence interval of distributed energy beta distribution[J]. Journal of Modern Power Systems and Clean Energy, 2021, 9(2): 423–430.
- [24] MACKEY C, TURNER B, BULLOCK M, et al. Margin sampling and survival outcomes in oral cavity and p16-positive oropharyngeal squamous cell carcinoma[J]. OTO Open, 2022, 6(3): 1–9.
- [25] OUYANG X, CHANG H, FENG B, et al. Application of an improved maximum entropy sampling method in hull form

optimization[J]. Ocean Engineering, 2023, 270(1): 1–21.

[26] WARD B. Resolving the raven paradox: simple random sampling, stratified random sampling, and inference to best explanation[J]. Philosophy of Science, 2022, 89(2): 360–377.

[27] 刘航,汪西莉. 自适应感受野机制遥感图像分割模型[J]. 中国图像图形学报, 2021, 26(2): 464–474.
LIU Hang, WANG Xili. Remote sensing image segmentation model based on an adaptive receptive field mechanism[J]. Journal of Image and Graphics, 2021, 26(2): 464–474. (in Chinese)

[28] 岳帆,谭勇,黄仕建,等. 基于全局金字塔和注意力机制的双流网络结构[J]. 信息技术与信息化, 2023, 1(3): 188–192.
YUE Fan, TAN Yong, HUANG Shijian, et al. A dual flow network architecture based on global pyramid and attention mechanism[J]. Information Technology and Informatization, 2023, 1(3): 188–192. (in Chinese)

[29] 方汀,刘艺超,唐哲,等. 基于高效通道注意力模块(ECA)和 YOLO v5 的图像检测方法研究[J]. 科学技术创新, 2023, 1(8): 88–91.
FANG Ting, LIU Yichao, TANG Zhe, et al. An image detection method based on ECA and YOLO v5[J]. Scientific and Technological Innovation, 2023, 1(8): 88–91. (in Chinese)

[30] FU Y, SONG J, XIE F, et al. Circular fruit and vegetable classification based on optimized GoogLeNet[J]. IEEE Access, 2021, 9(1): 113599–113611.

[31] BERNADO L S, DAMASEVICIUS R, LING S H, et al. Modified SqueezeNet architecture for parkinson’s disease detection based on keypress data[J]. Biomedicines, 2022, 10(11): 1–15.

[32] HUANG A, JIANG L, ZHANG J, et al. Attention – VGG16 – UNet: a novel deep learning approach for automatic segmentation of the median nerve in ultrasound images[J]. Quantitative Imaging in Medicine and Surgery, 2022, 12(6): 3138–3150.

[33] 朱炳宇,刘朕,张景祥. 融合 Grad – CAM 和卷积神经网络的 COVID – 19 检测算法[J]. 计算机科学与探索, 2022, 16(9): 2108–2120.
ZHU Bingyu, LIU Zhen, ZHANG Jingxiang. COVID – 19 detection algorithm combining Grad – CAM and convolutional neural network[J]. Computer Science and Exploration, 2022, 16(9): 2108–2120. (in Chinese)

~~~~~

(上接第 230 页)

[37] 热依拉·艾合买提,吾木提·艾山江,阿不都艾尼·阿不里,等. 基于机器学习的春小麦叶片水分含量高光谱估算[J]. 麦类作物学报, 2022,42(5):640–648.  
RAHILA Ahmat, UMUT Hasan, ABDUGHENI Abliz, et al. Hyperspectral estimation of spring wheat leaf water content based on machine learning[J]. Journal of Triticeae Crops, 2022,42(5):640–648. (in Chinese)

[38] AASEN H, HONKAVAARA E, LUCIEER A, et al. Quantitative remote sensing at ultra-high resolution with UAV spectroscopy: a review of sensor technology, measurement procedures, and data correction workflows[J]. Remote Sensing, 2018,10(7):1091.

[39] YIN C, LV X, ZHANG L, et al. Hyperspectral UAV images at different altitudes for monitoring the leaf nitrogen content in cotton crops[J]. Remote Sensing, 2022,14(11): 2576.

[40] LUO S, JIANG X, YANG K, et al. Multispectral remote sensing for accurate acquisition of rice phenotypes: impacts of radiometric calibration and unmanned aerial vehicle flying altitudes[J]. Frontiers in Plant Science, 2022, 13: 958106.

[41] CHUANMEI Z, JIANLI D, ZIPENG Z, et al. SPAD monitoring of saline vegetation based on Gaussian mixture model and UAV hyperspectral image feature classification[J]. Computers and Electronics in Agriculture, 2022, 200: 107236.

[42] HELEN S N, JOHN O, MBULISI S, et al. A comparative estimation of maize leaf water content using machine learning techniques and unmanned aerial vehicle (UAV)-based proximal and remotely sensed data[J]. Remote Sensing, 2021, 13(20):4091.

[43] CHLINGARYAN A, SUKKARIEH S, WHELAN B. Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: a review[J]. Computers and Electronics in Agriculture, 2018,151:61–69.

[44] SHI H, GUO J, AN J, et al. Estimation of chlorophyll content in soybean crop at different growth stages based on optimal spectral index[J]. Agronomy,2023,13(3): 663.