

基于改进注意力机制和多语义特征增强的 自然环境下枣品种识别方法

雷浩 苑迎春 许楠 何振学

(河北农业大学信息科学与技术学院, 保定 071001)

摘要: 针对目前自然环境下枣品种识别准确率较低的问题,提出了一种基于注意力机制和多语义特征增强的枣品种识别模型(ICBAM_MSFE_Res50)。该模型在 ResNet-50 基础上,引入改进注意力机制(Improved convolutional block attention module, ICBAM),ICBAM 采用一维卷积和多尺度空洞卷积对卷积块注意力模块(CBAM)进行改进,消除了特征图降维时的信息损失,降低了模型计算量和参数量,提高了模型对枣果区域细粒度特征的提取能力。同时,提出了多语义特征增强(Multi semantic feature enhancement, MSFE)模块,该模块通过枣果区域定位算法提取更多枣果局部显著特征,并采用显著性特征抑制算法迫使模型学习枣果次要特征,从而达到枣果多种语义特征学习。实验结果表明,在 20 类枣品种数据集上,本文模型准确率为 92.20%,与 ResNet-50 相比,提高 4.26 个百分点。对比 AlexNet、VGG-16、ResNet-18、InceptionV3 模型,准确率分别提高 15.84、9.22、6.86、3.55 个百分点。对比其他枣品种识别方法,本文方法在 20 种枣品种识别中表现最优,可为自然环境下枣品种识别研究提供参考。

关键词: 枣品种识别; 深度学习; 注意力机制; 多语义特征增强

中图分类号: TP391.41

文献标识码: A

文章编号: 1000-1298(2024)07-0270-10

OSID:



Jujube Variety Recognition Based on Improved Attention Mechanism and Multi-semantic Feature Enhancement

LEI Hao YUAN Yingchun XU Nan HE Zhenxue

(College of Information Science and Technology, Hebei Agricultural University, Baoding 071001, China)

Abstract: In response to the low accuracy of jujube variety recognition in current natural scenarios, a jujube variety recognition model was proposed based on attention mechanism and multi-semantic feature enhancement (ICBAM_MSFE_Res50). On the basis of ResNet-50, the attention mechanism ICBAM (improved convolutional block attention module) was introduced. ICBAM improved the convolutional block attention module (CBAM) by using one-dimensional convolution and multi-scale hole convolution, eliminating information loss during feature map dimensionality reduction, reducing the computational and parameter complexity of the model, and improving the model's ability to extract fine-grained features in jujube fruit regions. At the same time, a multi-semantic feature enhancement (MSFE) module was proposed, which extracted more local salient features of jujube fruit through jujube fruit region localization algorithm, and used saliency feature suppression algorithm to force the model to learn secondary features of jujube fruit, thereby achieving the learning of multiple semantic features of jujube fruit. The experimental results showed that the accuracy of the model on the dataset of 20 types of jujube varieties was 92.20%, which was 4.26 percentage points higher than that of ResNet-50. Compared with the AlexNet, VGG-16, ResNet-18, and InceptionV3 models, the accuracy was improved by 15.84, 9.22, 6.86, and 3.55 percentage points, respectively. Compared with other jujube variety recognition methods, this method still performed the best in the recognition of 20 types of jujube, which can provide reference for research on jujube variety recognition in natural scenarios.

Key words: jujube variety recognition; deep learning; attention mechanism; multi-semantic feature enhancement

收稿日期: 2023-11-03 修回日期: 2023-11-29

基金项目: 国家自然科学基金项目(62102130)和河北省自然科学基金项目(F2020204003)

作者简介: 雷浩(1999—),男,硕士生,主要从事深度学习与图像处理研究,E-mail: 2251081642@qq.com

通信作者: 苑迎春(1970—),女,教授,博士,主要从事智能信息处理与大数据技术研究,E-mail: nd_hd_ycy@163.com

0 引言

枣原产于中国,栽培历史悠久、种质资源丰富,具有重要的经济和生态价值^[1]。枣品种识别对枣果的种质资源研究、病害防治、自动采摘、分选销售等环节具有重要意义。因此,研究一种快速、准确的枣品种智能识别方法,可为枣产业高质量发展提供有力支撑,对智慧农业发展意义重大。

随着深度学习技术^[2]的不断发展和应用,卷积神经网络在图像识别领域逐渐占据主要地位。常见的卷积神经网络有 AlexNet^[3]、VGG^[4]、ResNet^[5]等。文献[6]以 3 种卷积神经网络为基学习器构建 Stacking 分类模型,在 5 个品类红枣上准确率达 92.38%。文献[7]在 VGGNet-16 网络基础上设计出一种双流网络结构,从形状特征和细粒度特征两个分支学习枣果特征,在 20 类枣上准确率可达 84.16%。文献[8]通过在 MobileNetV2 上进行改进,用于识别 8 种不同品种椰枣,准确率可达 99.0%。文献[9]提出一种基于 SqueezeNet 的双分支卷积神经网络用于识别干制红枣,两分支均采用 SqueezeNet 分别提取红枣特征,最后融合两分支提取到的特征进行分类。该方法在 4 种不同缺陷干制红枣上准确率高达 99.3%。文献[10]以枣果图像、枣叶片图像和叶片纹理图像作为模型输入,在 ResNet-18 网络的基础上构建三支网络结构,在自然环境下采集的 20 类枣上识别准确率为 91.68%。文献[11]提出一种缺陷红枣识别方法,通过在 AlexNet 上添加多尺度卷积模块,使网络充分学习红枣特征信息,该方法在 4 种红枣上平均识别率为 97.14%。

当前,越来越多的研究将注意力机制添加到卷积神经网络中。注意力机制^[12-14]是一种即插即用的模块,可以集成到任意网络的任意层中,能使模型更加注重挖掘图像显著特征,抑制无效信息,从而提升识别性能。文献[15]将注意力机制 SENet 引入 ResNet50 网络,构建用于识别田间自然环境下葡萄品种的分类模型。实验结果表明,注意力机制能有效降低背景干扰。文献[16]提出了一种融合通道和空间注意力的苹果品种鉴别模型,通过提取苹果表面的形状轮廓和颜色纹理特征,在 7 类苹果品种上准确率高达 96.78%。文献[17]提出一种苹果叶部病害分类模型,通过在 ResNet-18 网络中添加注意力机制 CBAM,强化网络对叶部病害区域的特征提取能力。文献[18]提出一种新的混合注意力模块,通过将 CBAM 中的通道注意力与空间注意力并行连接,有效提高了病虫害识别准确率。文

献[19]提出一种农作物病害识别模型,通过在 ResNet-18 中引入 Inception 模块和注意力机制 SENet,提高了特征的丰富度和模型对关键特征的提取能力。文献[20]在 ResNet-18 中引入 CBAM 注意力机制中的通道注意力模块,构建饲料原料种类识别模型,在测试集上准确率达 99.1%。

综上分析,卷积神经网络和注意力机制可为枣品种识别研究奠定技术基础。但自然场景下的枣品种识别存在枣果特征类间差异小、类内差异大、识别场景复杂等问题,仅采用以上技术方法难以有效识别枣品种图像,需要在此基础上对枣品种识别任务设计合适的分类模型。因此,为解决上述问题,本文提出一种基于注意力机制和多语义特征增强的枣品种识别方法。

本文以 20 类枣品种为对象,在 ResNet-50 基础上,构建基于改进注意力机制(ICBAM)和多语义特征增强(MSFE)的枣品种识别模型 ICBAM_MSFE_Res50。首先,针对 CBAM 存在的特征信息丢失、破坏通道特征与注意力权重间的直接对应关系问题,在通道注意力中引入一维卷积代替多层感知机结构,以保留特征通道与通道权重间的一一对应关系和相邻特征通道间的信息交互,并降低模型的计算量和参数量。然后,针对 CBAM 存在的卷积核特征提取尺度单一问题,在空间注意力中引入多尺度空洞卷积通路,使模型能充分提取全局空间上下文信息和局部细微特征,提高模型对枣果区域细粒度特征的提取能力。最后,针对枣品种图像的细粒度特征多存在于局部细节区域的特点,提出多语义特征增强模块,使用枣果区域定位算法来放大枣果局部细节区域的显著性特征,并使用显著性特征抑制算法迫使模型学习枣果次要特征,从而达到枣果多种语义特征学习。

1 实验数据

本文枣图像采集于河北省沧县国家枣树良种繁育基地,采集设备为小米智能手机和尼康 D7500 单反数码相机,分辨率分别为 2 250 像素 × 4 000 像素和 2 784 像素 × 1 856 像素。拍摄环境涵盖晴天、阴天、雨天等自然场景。拍摄时,相机镜头对准枣果,使其居于图像中间区域。每幅图像上仅包含单个完整枣果或多个分离枣果且都包含一定的复杂背景,例如枣果叶片、树枝、树干等。选取茶壶枣、朝阳圆枣等 20 类枣品种构建数据集,部分枣品种图像如图 1 所示。

共采集 20 类枣品种图像 2 159 幅,按照比例 8:2 划分成训练集和测试集,训练集图像 1 736 幅,

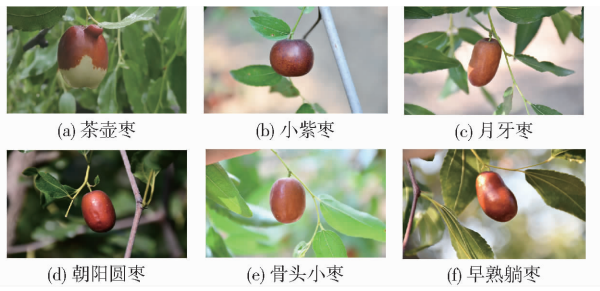


图1 部分枣品种图像

Fig. 1 Images of some jujube varieties

测试集图像 423 幅。对训练集中的图像进行数据增强,包括调整图像色度、对比度、随机翻转图像等 9 种增强方式,增强后的枣品种数据集分布如表 1 所示。在输入到模型之前,图像首先通过等比缩放将图像最小边缩放至 256 像素,保证图像不发生变形扭曲。再通过中心裁剪调整图像尺寸为 224 像素 \times 224 像素,最后进行归一化处理。

2 枣品种识别模型

本文构建的 ICBAM_MSFE_Res50 如图 2 所示。模型在 ResNet-50 的基础上引入改进注意力机制 ICBAM 和多语义特征增强模块进行改进。其中,改进注意力机制 ICBAM 添加到 ResNet-50 的卷积层后。多语义特征增强模块则由枣果区域定位和显著性特征抑制两部分组成。模型在训练过程中,主要

表 1 枣品种数据集分布
Tab. 1 Distribution of jujube variety dataset

枣品种	图像采集数量	训练集数量	测试集数量	训练集增强后数量
茶壶枣	124	100	24	1 000
朝阳圆枣	115	92	23	920
大荔龙枣	110	88	22	880
大木枣	93	75	18	750
短里长红	87	70	17	700
官滩枣	164	132	32	1 320
骨头小枣	105	84	21	840
葫芦长红	76	61	15	610
晋枣	101	81	20	810
辣椒枣	73	59	14	590
磨盘枣	59	48	11	480
山东龙枣	146	117	29	1 170
酥圆铃	103	83	20	830
襄汾圆枣	106	85	21	850
小枣	174	140	34	1 400
小紫枣	102	82	20	820
宣城圆枣	94	76	18	760
圆铃小枣	77	62	15	620
月牙枣	146	117	29	1 170
早熟躺枣	104	84	20	840
总计	2 159	1 736	423	17 360

分为 3 阶段:①原图首先经过 ResNet-50 提取全局特征,输出特征图送入多语义特征增强模块中进行

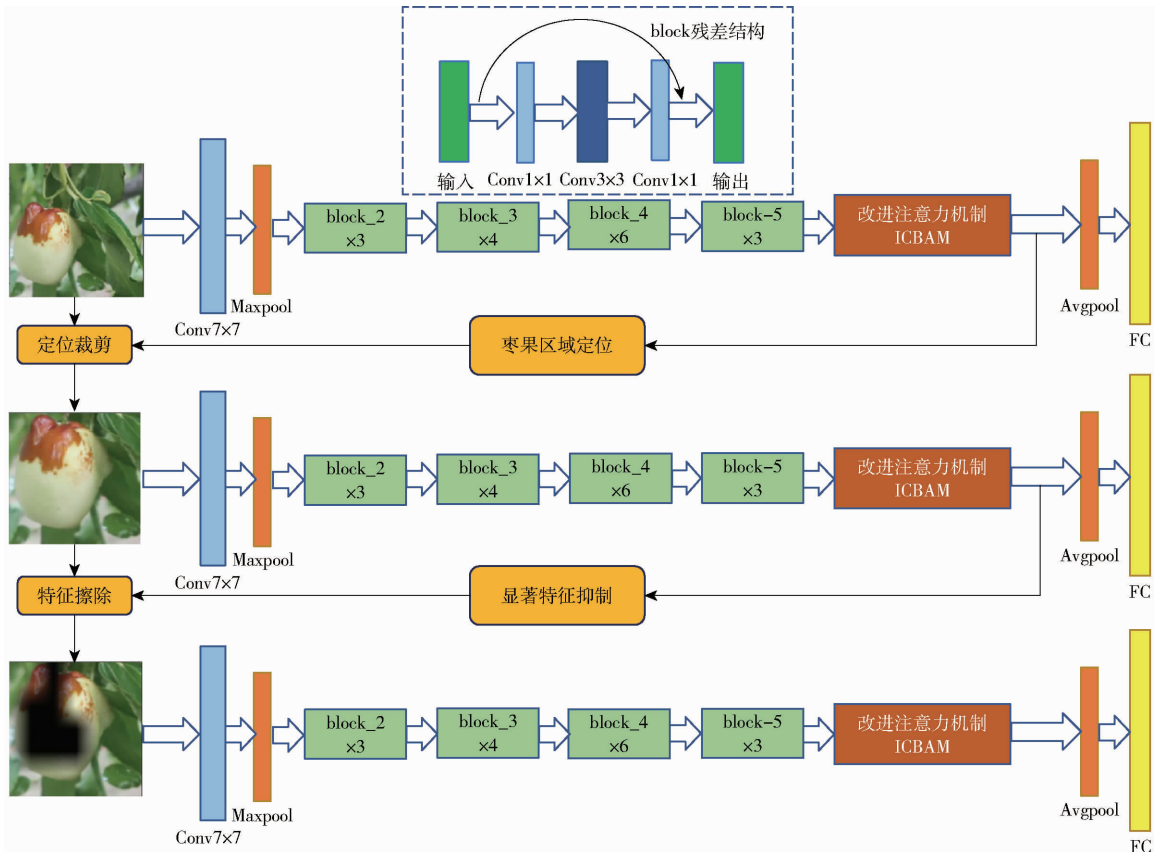


图2 ICBAM_MSFE_Res50 模型结构

Fig. 2 ICBAM_MSFE_Res50 model structure

枣果区域定位,并在原图上进行裁剪和上采样后得到枣果裁剪图。②将得到的枣果裁剪图送入 ResNet-50 中学习枣果细微特征,同时将输出特征图送入多语义特征增强模块中进行显著性特征抑制。③将枣果显著特征抑制图再送到 ResNet-50 中充分学习其他次要特征。图中 3 个分支 ResNet-50 为同一个网络,共享全部参数。网络总损失是 3 个分支损失之和,从而共同优化模型。

2.1 残差网络

随着卷积神经网络层数的不断加深,网络难以训练,容易引发梯度爆炸和梯度消失等问题,从而导致网络性能退化。为解决这一问题,HE 等^[5]在 2016 年提出了 ResNet 网络,该网络通过堆叠残差结构加深网络深度的同时,依旧可以取得良好的性能。残差结构如图 3 所示。

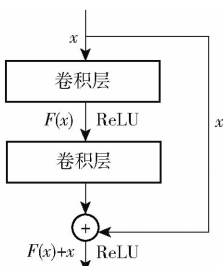


图 3 残差结构

Fig. 3 Residual structure

假设 x 是残差结构输入, $H(x)$ 是输出。网络并不是直接学习输入到输出的映射函数 $H(x)$, 而是学习输出与输入之间的差异, 即残差映射: $F(x) = H(x) - x$ 。通过恒等映射通路, 将浅层信息与当前残差块提取到的信息相结合, 不仅增强了模型特征表征能力, 而且有效避免网络出现梯度消失、难以收敛等问题。本文选择 ResNet-50 作为模型的基础网络, 并对其进行改进。

2.2 改进注意力机制 ICBAM

在自然场景下采集的枣品种图像中, 存在叶片、树枝等噪声信息, 会对模型的识别产生负面影响。注意力模块 CBAM^[13] 的提出可以达到抑制复杂背景干扰和提取关键像素区域的目的。CBAM 由通道注意力和空间注意力组成, 通道注意力主要筛选特征图通道维度上的重要特征, 能够抑制复杂背景等干扰信息。但是它采用由两层全连接层构成的多层感知机去学习特征通道维度的信息, 并利用缩放系数对特征通道进行先降维再升维的方式减少模型参数, 这破坏了特征通道与通道权重间的直接对应关系, 部分特征信息也在降维过程中受到损失。同时, 使用全连接层不可避免地会给模型带来大量参数。空间注意力模块重点关注特征图中对分类任务起关

键作用的像素区域, 其采用的卷积核大小决定了其特征提取性能, 大尺度卷积核有更大的感受野, 可更好地聚合空间上下文信息, 小尺度卷积核则可有效提取局部细微特征, 因此, 该模块仅采用一种尺度的卷积核获取空间位置信息的方法使得特征提取不全面。

通过分析 CBAM 在通道和空间注意力模块中的不足, 提出了改进注意力机制 ICBAM, 其在通道注意力上使用一维卷积代替多层感知机结构, 在空间注意力上采用多尺度空洞卷积通路。具体改进描述如下:

(1) 通道注意力改进

ICBAM 的通道注意力模块如图 4 所示。其具体流程如下: 给定输入特征图 $F \in \mathbf{R}^{C \times H \times W}$, 其中 C 表示特征图通道数, $H \times W$ 表示特征图尺寸(长 \times 宽), 首先分别在空间维度上进行全局平均池化和全局最大池化。全局平均池化突出图像全局特征信息, 最大池化突出图像显著特征信息。将两种池化的结果在通道维度上相加融合得到 F_{add} , 以获得更加全面的特征信息。然后, 使用卷积核长度为 k 的一维卷积来映射 F_{add} 中 k 个相邻通道间的特征信息。最后通过 Sigmoid 激活函数将值归一化到 $0 \sim 1$ 区间, 生成通道注意力权重 M_C 。其公式为

$$M_C(F) = \sigma(f_{1D}^k(\text{Avgpool}(F) + \text{Maxpool}(F))) \quad (1)$$

式中, f_{1D}^k 表示卷积核大小为 k 的一维卷积操作, 可根据输入特征图的通道数自适应确定一维卷积核 k 的大小, 公式为

$$k = \left\lfloor \frac{\text{lb}C}{2} + \frac{1}{2} \right\rfloor_{\text{odd}} \quad (2)$$

式中, $\left\lfloor \frac{\text{lb}C}{2} + \frac{1}{2} \right\rfloor_{\text{odd}}$ 表示向上取与 $\frac{\text{lb}C}{2} + \frac{1}{2}$ 最接近的奇数。

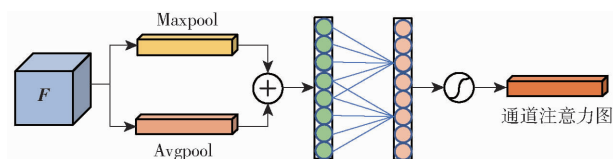


图 4 ICBAM 的通道注意力模块

Fig. 4 Channel attention module for ICBAM

(2) 空间注意力改进

ICBAM 的空间注意力模块如图 5 所示。其具体流程如下: 给定输入特征图 $F \in \mathbf{R}^{C \times H \times W}$, 分别在通道维度上进行全局平均池化和全局最大池化, 并将得到的两个特征图在通道维度上进行拼接融合得到 F_{concat} 。接着将 F_{concat} 输入到多尺度卷积通路分别提取不同感受野的特征信息, 该结构包括一个 1×1

的卷积层和3个 3×3 的空洞卷积层(扩张因子分别为1、2、3)。然后将多尺度卷积通路的输出再进行拼接融合,并使用 1×1 卷积层进行通道降维。最后通过Sigmoid激活函数将值归一化到0~1区间,生成空间注意力权重 M_s 。其过程公式为

$$\mathbf{A}(\cdot) = [\text{Conv}_{(1,1)}(\cdot), \text{Conv}_{(3,1)}(\cdot), \text{Conv}_{(3,2)}(\cdot), \text{Conv}_{(3,3)}(\cdot)] \quad (3)$$

$$\mathbf{M}_s(\mathbf{F}) = \sigma(f^{1 \times 1}(\text{concat}(\mathbf{A}[\text{Avgpool}(\mathbf{F}); \text{Maxpool}(\mathbf{F})]))) \quad (4)$$

式中 \mathbf{A} ——多尺度卷积通路

$\text{Conv}_{(i,j)}$ ——采用卷积核原始大小为 i 、扩张因子为 j 的空洞卷积核进行卷积操作

$f^{1 \times 1}$ ——卷积核大小为 1×1 的卷积操作

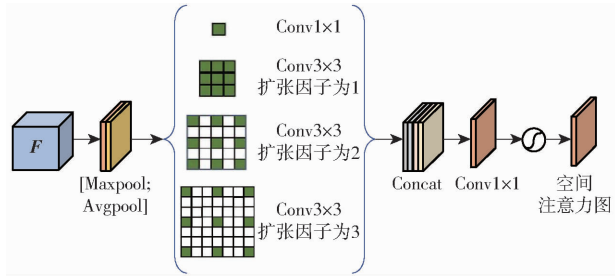


图5 ICBAM的空间注意力模块

Fig. 5 Spatial attention module for ICBAM

2.3 多语义特征增强模块 MSFE

多语义特征增强模块包含枣果区域定位和枣果显著性特征抑制两部分,通过枣果区域定位可使模型更加关注枣果局部显著性特征,在此基础上再进行显著性特征抑制,使模型也充分学习次要特征,以此达到多种语义特征的学习。

(1) 枣果区域定位

卷积神经网络在提取图像特征时,同一通道中不同位置学习到图像不同区域的特征,同一像素位置在不同通道的特征响应也各有不同。在通道维度上聚合特征图后,某些位置的特征激活响应越高,则该特征的重要程度就越大,而且其属于枣果对象的可能性就越大。因此,本研究通过聚合枣品种图像的深度卷积特征,用于定位图像中的枣果区域,具体过程如下:假设 \mathbf{F} 是ResNet-50最后一层卷积层经过改进注意力机制ICBAM后的输出, \mathbf{F}_i 表示特征图 \mathbf{F} 的第 i 个特征通道。在通道维度上对 \mathbf{F} 的对应位置求和得到激活响应图 \mathbf{A} ,其公式表示为

$$\mathbf{A} = \sum_{i=0}^{c-1} \mathbf{F}_i \quad (5)$$

激活图 \mathbf{A} 中每个位置的响应值代表了其特征

重要程度,可有效反映目标对象的空间分布。根据激活图 \mathbf{A} 计算阈值 a ,用于判断 \mathbf{A} 中的位置 (i,j) 是否属于目标对象的一部分,阈值 a 的计算公式为

$$a = \frac{\sum_{i=0}^{H-1} \sum_{j=0}^{W-1} \mathbf{A}(i,j)}{HW} \quad (6)$$

将响应值小于阈值 a 的区域作为无用背景区域忽略掉,大于阈值的响应区域作为枣果对象区域保留下来。其过程为

$$M_{(i,j)} = \begin{cases} 1 & (\mathbf{A}(i,j) > a) \\ 0 & (\mathbf{A}(i,j) \leq a) \end{cases} \quad (7)$$

目标对象通常位于掩码 $M_{(i,j)}$ 的最大连通分量中,因此选取 $M_{(i,j)}$ 中最大连通分量的最小外接矩形作为目标对象的定位区域,最后将该区域在原图上进行裁剪并放大到原图尺寸,获得枣果目标定位图像。枣果定位图像不仅放大了原始图像中的枣果对象,排除掉了大量背景噪声,使网络更加注重对枣果本身的学习,而且针对枣品种图像中的判别特征常存在于局部细微区域的问题,实现了枣果判别部位和细微特征的放大,更有利于网络提取特征。

(2) 显著性特征抑制

对于目标图像,现有的部分细粒度图像分类方法只学习到图像中最显著的特征区域,忽略了同样具有重要作用的次要特征区域。该类方法忽略了特征区域之间的联系,达不到特征的充分学习。因此,为增强模型提取到的语义信息,通过对输入图像中的显著性特征区域进行抑制,迫使模型学习次要区域的特征,使图像特征学习更加全面。

具体来说,目标定位图像经过特征提取层和注意力模块后得到特征图 $\mathbf{F}_{local} \in \mathbf{R}^{C \times H \times W}$,同样采用式(5)获取到 \mathbf{F}_{local} 对应的激活响应图 \mathbf{A}_{local} ,其中激活值较大的区域表示其属于显著性特征区域,相反则为次要特征区域。为使模型重点关注到除高响应特征区域外的其他特征,根据式(6)设置阈值 a_{local} 得到显著特征抑制掩码 M_{mask} ,其公式为

$$M_{mask} = \begin{cases} 1 & (\mathbf{A}_{local}(i,j) < a_{local}) \\ 0 & (\mathbf{A}_{local}(i,j) \geq a_{local}) \end{cases} \quad (8)$$

将 M_{mask} 上采样到原图尺寸后,再与枣果定位图像相乘,得到显著性特征区域被抑制的新图像。将新图像重新输入到模型中进行训练,促使模型学习其他区域的细粒度特征。

3 实验结果与分析

3.1 实验设置

本文实验环境参数如表2所示。

表2 实验环境参数

Tab.2 Experimental environment parameters

实验环境	参数
操作系统	Windows 10
处理器	Inter(R) Core(TM) i7-8700K@ 3.70GHz
运行内存	16GB
显卡	NVIDIA GeForce GTX 1080 Ti
编程语言	Python 3.6
深度学习框架	PyTorch
批处理大小	32
迭代次数	50
优化器	SGD
初始学习率	0.001
初始化参数	ImageNet 预训练参数

3.2 与其他卷积神经网络对比

为验证本文模型在枣品种识别研究上的有效性,选取在图像识别领域具有代表性的卷积神经网络进行对比,其中包括 AlexNet、VGG-16、ResNet-18、ResNet-50、InceptionV3^[21]。在训练和测试阶段,保证所有模型的各项实验参数设置相同。不同模型在测试集上的准确率、参数量和内存占用量对比如表3所示。在各类枣品种上准确率对比如表4所示。

表3 各类神经网络性能对比

Tab.3 Comparison of various neural networks

CNN 类型	准确率/%	参数量	内存占用量/MB
AlexNet	76.36	5.7090×10^7	217.0
VGG-16	82.98	1.3430×10^8	512.0
ResNet-18	85.34	1.1187×10^7	42.7
ResNet-50	87.94	2.3550×10^7	90.1
InceptionV3	88.65	2.1827×10^7	83.5
ICBAM_MSFE_Res50	92.20	2.3550×10^7	90.1

表4 不同神经网络在各类枣品种上的准确率

Tab.4 Accuracy of different neural networks on various jujube varieties

枣品种	AlexNet	VGG-16	ResNet-18	ResNet-50	InceptionV3	ICBAM_MSFE_Res50	%
茶壶枣	95.83	91.67	91.67	91.67	91.93	100	
朝阳圆枣	47.83	69.57	73.91	82.61	77.99	86.96	
大荔龙枣	68.18	81.82	81.82	90.91	76.70	95.45	
大木枣	83.33	83.33	77.78	100	94.44	100	
短里长红	52.94	70.59	58.82	76.47	68.20	76.47	
官滩枣	90.62	93.75	93.75	93.75	94.73	100	
骨头小枣	61.90	47.62	80.95	71.43	67.71	85.71	
葫芦长红	66.67	46.67	66.67	80.00	82.29	73.33	
晋枣	50.00	85.00	85.00	80.00	73.59	85.00	
辣椒枣	71.43	92.86	85.71	85.71	87.95	92.86	
磨盘枣	100	100	100	100	99.43	100	
山东龙枣	93.10	96.55	100	100	99.68	100	
酥圆铃	55.00	80.00	80.00	75.00	81.25	90.00	
襄汾圆枣	90.48	100	85.71	85.71	96.88	95.24	
小枣	85.29	94.12	91.18	88.24	94.94	91.18	
小紫枣	80.00	70.00	95.00	95.00	94.38	95.00	
宣城圆枣	94.44	100	100	100	92.19	94.44	
圆铃小枣	73.33	93.33	86.67	93.33	91.25	93.33	
月牙枣	86.21	75.86	86.21	89.66	88.69	96.55	
早熟躺枣	60.00	75.00	70.00	75.00	80.94	80.00	

从表3可知,除本文提出的模型 ICBMA_MSFE_Res50 外,其他对比网络识别准确率均低于 90%,本文模型在测试集上准确率最高,可达到 92.20%。其中,AlexNet 识别准确率最低,仅为 76.36%,同时模型参数量和内存占用量较大,ICBAM_MSFE_Res50 比其准确率提升 15.84 个百分点。VGG-16 的模型参数量和内存占用量均为最大,ICBAM_MSFE_Res50 比其准确率提升 9.22 个百分点。与改进前的 ResNet-50 相比,在其基础上添加改进的注意力机制和多语义特征增强模块后,识别准确率

提升 4.26 个百分点,而模型参数量和内存占用量几乎不变。ICBAM_MSFE_Res50 比 InceptionV3,准确率提升 3.55 个百分点,虽然在参数量和内存占用量上略有增加,但模型分类效果显著提升,达到了模型识别性能和复杂度的有效平衡。分析表明本文模型综合性能优于其他对比网络,更适用于枣品种识别。

由表4可知,ICBAM_MSFE_Res50 相比于其他对比网络,在其中 15 类枣品种上的识别准确率取得最优。其中 8 类枣品种识别准确率超过 95.00%,有 5 类枣品种准确率达到 100%,分别为茶壶枣、大

木枣、官滩枣、磨盘枣和山东龙枣。其他5类枣准确率虽然不是最优,但也相对较高,有3类枣准确率超过90%。不同模型在磨盘枣和山东龙枣上的识别准确率都相对较高,原因是这两类枣的显著性特征较为明显,模型容易识别。20类枣中仅有2类枣品种准确率低于80%,分别为短里长红和葫芦长红,这两类枣在其他对比网络上的识别准确率也相对较低,原因是这两类枣果特征较不明显,模型识别效果差,但在短里长红上,本文模型仍取得最优准确率。综合整体表现,ICBAM_MSFE_

Res50能较好地完成枣品种分类任务。

3.3 消融实验

为验证ICBAM_MSFE_Res50中各个改进点的有效性,设置7个实验组在测试集上进行对比实验,分别为:ResNet-50(实验组A);在ResNet-50基础上分别添加注意力机制SENet、ECANet和CBAM(实验组B、C、D);在ResNet50上分别添加改进注意力机制ICBAM和多语义特征增强模块(实验组E和F);以及本文提出的ICBAM_MSFE_Res50(实验组G)。实验对比如表5所示。

表5 消融实验

Tab.5 Ablation experiment

实验组	模型	准确率/%	参数量	内存占用量/MB
A	ResNet-50	87.94	2.355×10^7	90.1
B	ResNet-50 + SENet	88.42	2.407×10^7	92.1
C	ResNet-50 + ECANet	88.89	2.355×10^7	90.1
D	ResNet-50 + CBAM	89.60	2.460×10^7	94.1
E	ResNet-50 + 改进注意力机制 ICBAM	90.07	2.355×10^7	90.1
F	ResNet-50 + 多语义特征增强模块	91.02	2.355×10^7	90.1
G	ICBAM_MSFE_Res50	92.20	2.355×10^7	90.1

对比实验组A、B、C、D、E可知,在ResNet-50的基础上添加注意力机制后,通过提升枣果特征权重,减少复杂背景干扰,可有效提升枣品种识别准确率。本文提出的ICBAM相比于其他3种注意力机制,准确率最高,可达90.07%,相比于ResNet-50提升2.13个百分点。相比于改进前的CBAM,通过在通道和空间注意力模块中分别引入高效一维卷积和多尺度卷积通路后,准确率提升0.47个百分点。而ICBAM在模型参数量和内存占用量相比于CBAM大大减小,与ResNet-50相当。

实验组F在ResNet-50基础上添加了多语义特征增强模块,通过枣果目标定位和显著性特征抑制使模型特征学习更加全面,识别准确率可达91.02%,相比于ResNet-50提升3.08个百分点,相比于A~E实验组,提升效果最为明显。同时这也说明,充分学习枣果全局特征和局部细微特征可有效区分各类枣品种,对识别准确率的提升贡献最大。而且,多尺度增强模块并不会额外增加模型的参数量和权重。

由上分析,分别添加改进注意力机制ICBAM和多语义特征增强模块均能有效提升网络的准确率。ICBAM_MSFE_Res50在ResNet-50基础上,融合改进注意力机制ICBAM和多语义特征增强模块,相比于实验组E、F准确率进一步提升至92.20%,相比于ResNet-50提升4.26个百分点,提升效果显著。模型参数量和内存占用量相比于ResNet-50基本

不变。表明本文模型在枣品种识别上的优越性。

3.4 混淆矩阵

混淆矩阵可直观表明模型在各个品种上的分类效果,是评判模型识别性能的重要指标之一。将训练好的ResNet-50和ICBAM_MSFE_Res50在测试集上生成混淆矩阵,如图6所示。

混淆矩阵中,每个枣品种对角线上的值越大,则表示识别效果越好。分析两个混淆矩阵可知,在ICBAM_MSFE_Res50中,共有12类枣品种的识别正确个数有所提升。其中,骨头小枣和酥圆铃的提升最大,识别准确率分别提升14.28、15.00个百分点。在酥圆铃识别中,通过模型对枣果全局和局部细微特征的充分学习,消除了对茶壶枣、大荔龙枣和圆铃小枣的错判。在茶壶枣、官滩枣、襄汾圆枣和月牙枣上识别正确个数均有增加,其中茶壶枣和官滩枣达到全部识别正确。在葫芦长红上识别效果略有不足,因为该枣与辣椒枣的形状较为相似,不易区分,导致模型错判为辣椒枣的个数较多。另外,模型识别错误的原因还可能是受复杂背景和拍摄光线的影响较大,部分枣果因存在缺陷也会导致识别错误。综合分析,ICBAM_MSFE_Res50在枣品种识别上有较好的分类性能。

根据上述混淆矩阵计算出两种模型对应各个枣品种精确率、召回率和F1值,计算结果如表6所示。

本文提出的ICBAM_MSFE_Res50在测试集上

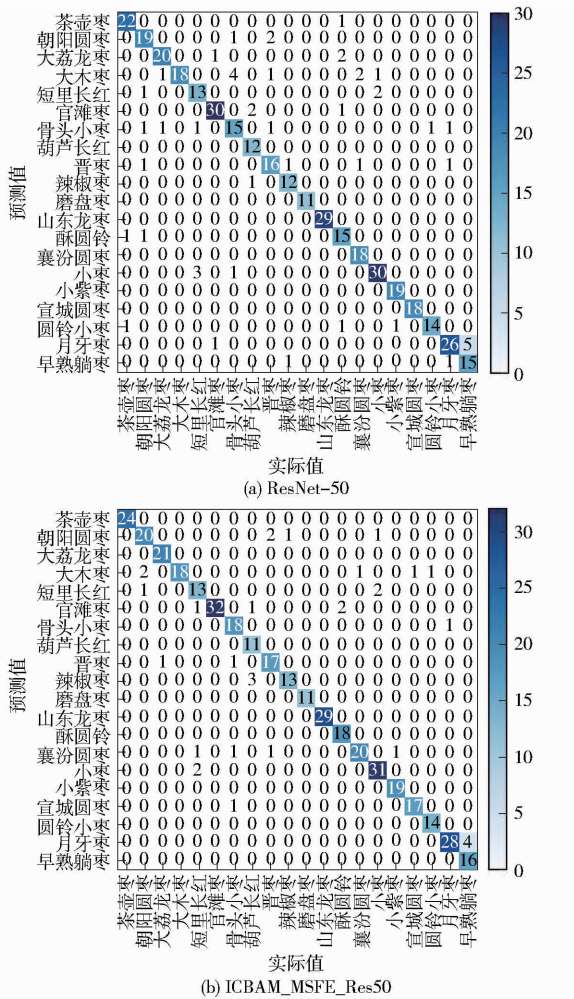


图 6 ResNet-50 和 ICBAM_MSFE_Res50 的混淆矩阵

Fig.6 ResNet-50 and ICBAM_MSFE_Res50 confusion matrix

平均精准率为 92.82%、召回率为 91.58%、F1 值为 91.86%；改进前 ResNet-50 在测试集上平均精确率为 88.81%、召回率为 87.72%、F1 值为 87.92%。在上述 3 种指标中,精确率越高表示模型越不会将其他品种的枣类错误地识别为特定品种,即模型能有效区分相似品种。召回率越高越表示模型能有效区分同一品种中的不同特征的枣。F1 值则综合反映模型解决枣果类间差异小、类内差异大的效果。本文提出的 ICBAM_MSFE_Res50 相比于 ResNet-50 均有一定提升,表明本文提出的模型具有较好的识别效果,能有效解决枣品种识别类间差异小、类内差异大的问题。

3.5 其他枣品种识别方法对比

表 7 给出了不同枣品种分类方法的识别结果。文献[22-23]中采用传统机器学习方法和高光谱成像技术构建识别模型。对于识别数量较少的枣品种,可取得较好的识别效果。但传统机器学习方法特征设计、选取过程复杂,高光谱图像采集成本高,难以推广到识别更多枣品种数量上。文献[6-7, 10, 24]采用深度学习方法,在品种数量较大时,依

表 6 评价指标对比

Tab.6 Comparison of evaluation indexes %

枣品种	ResNet-50			ICBAM_MSFE_Res50		
	精确率	召回率	F1 值	精确率	召回率	F1 值
茶壶枣	95.65	91.67	93.62	100	100	100
朝阳圆枣	82.61	82.61	82.61	83.33	86.96	85.11
大荔龙枣	86.96	90.91	88.89	100	95.45	97.67
大木枣	66.67	100	80.00	78.26	100	87.80
短里长红	81.25	76.47	78.79	81.25	76.47	78.79
官滩枣	90.91	93.75	92.31	88.89	100	94.12
骨头小枣	71.43	71.43	71.43	94.74	85.71	90.00
葫芦长红	100	80.00	88.89	100	73.33	84.61
晋枣	80.00	80.00	80.00	89.47	85.00	87.18
辣椒枣	92.31	85.71	88.89	81.25	92.86	86.67
磨盘枣	100	100	100	100	100	100
山东龙枣	100	100	100	100	100	100
酥圆铃	88.24	75.00	81.08	100	90.00	94.74
襄汾圆枣	100	85.71	92.31	83.33	95.24	88.89
小枣	88.24	88.24	88.24	93.94	91.18	92.54
小紫枣	100	95.00	97.44	100	95.00	97.44
宣城圆枣	100	100	100	94.44	94.44	94.44
圆铃小枣	82.35	93.33	87.50	100	93.33	96.55
月牙枣	81.25	89.66	85.25	87.50	96.55	91.80
早熟躺枣	88.24	75.00	81.08	100	80.00	88.89
平均值	88.81	87.72	87.92	92.82	91.58	91.86

表 7 各类枣品种分类方法识别结果

Tab.7 Comparison of identification methods for various jujube varieties

分类方法	品种数	准确率/
	量/种	%
基于近红外高光谱成像技术的干制红枣识别 ^[22]	4	100
GA-SVM ^[23]	4	96.00
基于 Stacking 模型的红枣品种分类 ^[6]	5	92.38
基于度量学习的枣品种识别方法 ^[7]	20	84.16
融合枣果与叶片信息的枣品种识别方法 ^[10]	20	91.68
基于 MobileNetV3 的枣品种细粒度分类多任务模型 ^[24]	20	91.79
ICBAM_MSFE_Res50	20	92.20

旧可以取得较高准确率。但文献[10]需要融合枣的多器官特征信息,这对数据集的采集和预处理要求较高。文献[24]手动裁剪出了枣图像中的枣果,弱化了自然环境下复杂背景影响识别效果这一问题。本文针对自然场景下的枣品种识别问题,采用注意力机制和多语义特征增强模块,充分增强枣果特征表示,减弱复杂背景干扰。在与其他方法同为识别 20 类枣品种上,取得最高分类准确率,表明本文模型更适用于枣品种识别。

3.6 模型可视化

为更好地理解模型学习过程,随机选取 4 类枣

品种图像,在原图基础上,可视化枣果定位图、枣果裁剪图和显著性特征抑制图。其中枣果定位图为在原图的基础上定位到枣果区域。枣果裁剪图是将定位到的枣果区域进行裁剪并放大到指定尺寸。显著性特征抑制图则是在枣果裁剪图的基础上对显著特征区域进行擦除。4类枣品种图像可视化结果如图7所示。

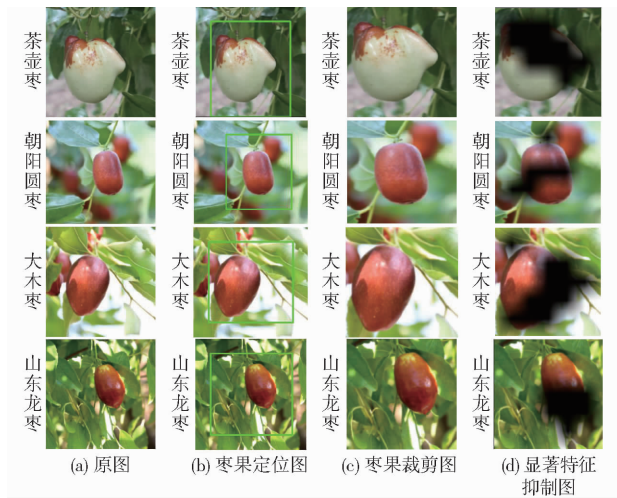


图7 4类枣品种图像可视化

Fig.7 Image visualization of four types of jujube varieties

由图7可知,在这4类枣品种中,通过枣果区域定位方法都能较好地定位到枣果区域,排除了一定的复杂背景。将定位到的枣果进行裁剪并上采样到指定图像输入尺寸,从而放大枣果更多局部细节特征信息。在茶壶枣中,显著性特征为枣果上半部类似“茶壶嘴”的凸起部分;在朝阳圆枣中,枣果表面较为粗糙;在大木枣和山东龙枣中,枣果底部为锥形。综合分析图7c、7d可知,模型能在枣果裁剪图的基础上学习到这些显著性特征,并在显著性特征抑制图中对这些特征进行擦除,迫使模型学习枣果其他位置的特征信息,使特征学习更加全面。

3.7 公共数据集结果分析

为验证本文方法的适用性,将本文方法应用到细粒度图像识别任务中常见的3类公共数据集上,这3类数据集分别为CUB-200-2011、Stanford Cars和FGVC-Aircraft。这些数据集同样具有类别数多、图像特征类间差异大、类间差异小、易受复杂背景干扰、图像识别难度大等特点。这3类数据集分布如表8所示。本文方法在这3类公共数据集上的分类结果如表9所示。

表8 3类公共数据集分布

Tab.8 Distribution of three types of public datasets

公共数据集	识别对象	类别数	训练集数量	测试集数量
CUB-200-2011	鸟类	200	5 994	5 794
Stanford Cars	汽车	196	8 144	8 041
FGVC-Aircraft	飞机	100	6 667	3 333

表9 3类公共数据集分类结果

Tab.9 Classification results for three types of public datasets

模型	准确率			%
	CUB-200-2011	Stanford Cars	FGVC-Aircraft	
B-CNN ^[25]	84.1	91.3	84.1	
ResNet-50	84.5	92.2	90.3	
RA-CNN ^[26]	85.4	92.5	88.4	
MA-CNN ^[27]	86.5	92.8	89.9	
NTS-Net ^[28]	87.5	93.9	91.4	
MC-Loss ^[29]	87.3	93.7	92.6	
DCL ^[30]	87.8	94.5	93.0	
CIN ^[31]	88.1	94.5	92.8	
ICBAM_MSFE_	88.5	95.3	94.0	
Res50				

从表9可以看出,与其它经典细粒度图像分类方法相比,本文方法在3类数据集上均取得最高准确率。在CUB-200-2011数据集上获得88.5%的准确率,在Stanford Cars上获得95.3%的准确率,在FGVC-Aircraft上获得94.0%的准确率,充分表明本文提出模型在不同数据集上的优越性。

4 结论

(1)针对自然场景下枣品种识别存在的问题,本文提出一种融合改进注意力机制ICBAM和多语义特征增强的枣品种识别模型。ICBAM在CBAM的基础上利用一维卷积和多卷积特征提取通路进行改进,增强了模型对枣果细粒度特征的提取能力,同时ICBAM参数量相比于CBAM有所减少。多语义特征增强模块的引入,使模型对枣果局部细微特征等多语义信息得到充分学习,模型相比改进前,识别性能显著提升。

(2)为验证本文模型的有效性,在20类枣品种数据集上,ICBAM_MSFE_Res50准确率可达92.20%,识别性能优于主流的卷积神经网络。同时相比于其他枣品种分类方法,本文模型在识别20类枣品种上具有最高的识别准确率。

参考文献

- [1] 刘孟军,王玖瑞. 新中国果树科学研究70年——枣[J]. 果树学报,2019,36(10):1369-1381.
LIU Mengjun, WANG Jiurui. Fruit scientific research in New China in the past 70 years; Chinese jujube[J]. Journal of Fruit Science, 2019, 36(10): 1369-1381. (in Chinese)

- [2] LECUN Y, BENGIO Y, HINTON G. Deep learning[J]. *Nature*, 2015, 521(7553): 436 - 444.
- [3] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks[J]. *Commun. ACM*, 2017, 60(6): 84 - 90.
- [4] SIMONYAN K, ZISSERMA A. Very deep convolutional networks for large-scale image recognition[J]. *arXiv preprint arXiv: 1409.1556*, 2014.
- [5] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]// *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016: 770 - 778.
- [6] 余游江, 喻彩丽, 尚远航, 等. 基于 Stacking 模型的红枣品种分类识别[J]. *北方园艺*, 2022(8): 139 - 148.
YU Youjiang, YU Caili, SHANG Yuanhang, et al. Classification and recognition of jujube varieties based on Stacking model fusion[J]. *Northern Horticulture*, 2022(8): 139 - 148. (in Chinese)
- [7] MENG X, YUAN Y C, TENG G F, et al. Deep learning for fine-grained classification of jujube fruit in the natural environment[J]. *Journal of Food Measurement and Characterization*, 2021, 15(5): 4150 - 4165.
- [8] ALBARRAK K, GULZAR Y, HAMID Y, et al. A deep learning-based model for date fruit classification[J]. *Sustainability*, 2022, 14(10): 6339.
- [9] GENG L, XU W, ZHANG F, et al. Dried jujube classification based on a double branch deep fusion convolution neural network[J]. *Food Science and Technology Research*, 2018, 24(6): 1007 - 1015.
- [10] 孟惜, 苑迎春, 刘博, 等. 融合枣果与叶片信息的枣品种识别方法研究[J]. *河北农业大学学报*, 2022, 45(2): 107 - 113.
MENG Xi, YUAN Yingchun, LIU Bo, et al. Classification of jujube by integrating fruit with leaf information[J]. *Journal of Hebei Agricultural University*, 2022, 45(2): 107 - 113. (in Chinese)
- [11] 方双, 赵凤霞, 楚松峰, 等. 基于多尺度卷积神经网络的缺陷红枣检测方法[J]. *食品与机械*, 2021, 37(2): 158 - 163, 168.
FANG Shuang, ZHAO Fengxia, CHU Songfeng, et al. Defective jujube detection technology based on multi-scale convolutional neural network[J]. *Food & Machinery*, 2021, 37(2): 158 - 163, 168. (in Chinese)
- [12] 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.
- [13] WOO S, PARK J, LEE J Y, et al. CBAM: convolutional block attention module[C]// *Proceedings of the European Conference on Computer Vision (ECCV)*, 2018: 3 - 19.
- [14] WANG Q, WU B, ZHU P, et al. ECA - Net: efficient channel attention for deep convolutional neural networks[C]// *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020: 11534 - 11542.
- [15] 苏宝峰, 沈磊, 陈山, 等. 基于注意力机制的葡萄品种多特征分类方法[J]. *农业机械学报*, 2021, 52(11): 226 - 233, 252.
SU Baofeng, SHEN Lei, CHEN Shan, et al. Multi-features identification of grape cultivars based on attention mechanism[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2021, 52(11): 226 - 233, 252. (in Chinese)
- [16] 耿磊, 黄亚龙, 郭永敏. 基于融合注意力机制的苹果品种分类方法[J]. *农业机械学报*, 2022, 53(6): 304 - 310, 369.
GENG Lei, HUANG Yalong, GUO Yongmin. Apple variety classification method based on fusion attention mechanism[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2022, 53(6): 304 - 310, 369. (in Chinese)
- [17] 姜红花, 杨祥海, 丁睿柔, 等. 基于改进 ResNet18 的苹果叶部病害多分类算法研究[J]. *农业机械学报*, 2023, 54(4): 295 - 303.
JIANG Honghua, YANG Xianghai, DING Ruirou, et al. Identification of apple leaf diseases based on improved ResNet18[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2023, 54(4): 295 - 303. (in Chinese)
- [18] 王美华, 吴振鑫, 周祖光. 基于注意力改进 CBAM 的农作物病虫害细粒度识别研究[J]. *农业机械学报*, 2021, 52(4): 239 - 247.
WANG Meihua, WU Zhenxin, ZHOU Zuguang. Fine-grained identification research of crop pests and diseases based on improved CBAM via attention[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2021, 52(4): 239 - 247. (in Chinese)
- [19] 黄林生, 罗耀武, 杨小冬, 等. 基于注意力机制和多尺度残差网络的农作物病害识别[J]. *农业机械学报*, 2021, 52(10): 264 - 271.
HUANG Linsheng, LUO Yaowu, YANG Xiaodong, et al. Crop disease recognition based on attention mechanism and multi-scale residual network[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2021, 52(10): 264 - 271. (in Chinese)
- [20] 牛智有, 于重洋, 吴志陶, 等. 基于改进 ResNet18 模型的饲料原料种类识别方法[J]. *农业机械学报*, 2023, 54(2): 378 - 385, 402.
NIU Zhiyou, YU Chongyang, WU Zhitao, et al. Identification of feed raw material type based on improved ResNet18 model[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2023, 54(2): 378 - 385, 402. (in Chinese)
- [21] SZEGEDY C, VANHOUCKE V, IOFFE S, et al. Rethinking the inception architecture for computer vision[C]// *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016: 2818 - 2826.

53(8):213–219,239.

YANG Shuqin, WANG Pengfei, WANG Shuai, et al. Detection of wheat lodging in UAV remote sensing image based on Multi-head Self-attention DeepLab v3 + [J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(8):213–219,239. (in Chinese)

- [26] 常晗,郭树欣,张海洋,等. 基于改进 CBAM-DeepLab V3+ 的苹果种植面积提取[J]. 农业机械学报,2023,54(增刊2):206–213.
CHANG Han, GUO Shuxin, ZHANG Haiyang, et al. Apple planting area extraction based on improved DeepLab V3 + [J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(Supp. 2):206–213. (in Chinese)
- [27] WOO S, PARK J, LEE J, et al. CBAM: convolutional block attention module [C] // European Conference on Computer Vision, Munich, Germany: Springer, 2018:3–19.
- [28] HERTEM V T, TELLO S A, VIAZZI S, et al. Implementation of an automatic 3D vision monitor for dairy cow locomotion in a commercial farm[J]. Biosystems Engineering, 2018, 173: 166–175.
- [29] SCHLAGETER-TELLO A, HERTEM V T, BOKKERS E A, et al. Performance of human observers and an automatic 3-dimensional computer-vision-based locomotion scoring method to detect lameness and hoof lesions in dairy cows[J]. Journal of Dairy Science, 2018, 101(7): 6322–6335.
- [30] 李前,初梦苑,康熙,等. 基于计算机视觉的奶牛跛行识别技术研究进展[J]. 农业工程学报,2022,38(15):159–169.
LI Qian, CHU Mengyuan, KANG Xi, et al. Research progress of lameness recognition technology in dairy cows using computer vision[J]. Transactions of the CASE, 2022, 38(15):159–169. (in Chinese)
- [31] HARRER S. Attention is not all you need; the complicated case of ethically using large language models in healthcare and medicine[J]. eBioMedicine, 2023, 90:1045122.
- [32] JI Y, ZHOU Z, LIU H, et al. DNABERT: pre-trained bidirectional encoder representations from transformers model for DNA-language in genome [J]. Bioinformatics, 2021, 37(15):2112–2120.
- [33] 姚冲,李前,刘刚,等. 基于 NAS-Res 的局部遮挡荷斯坦奶牛个体识别[J]. 农业机械学报,2023,54(增刊1):252–259.
YAO Chong, LI Qian, LIU Gang, et al. Individual identification of partially occluded Holstein cows based on NAS-Res[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(Supp. 1):252–259. (in Chinese)
- [34] WU D, WANG Y, HAN M, et al. Using a CNN-LSTM for basic behaviors detection of a single dairy cow in a complex environment[J]. Computers and Electronics in Agriculture, 2021, 182: 106016.

(上接第 279 页)

- [22] 樊阳阳,裴正军,陈俭,等. 基于近红外高光谱成像技术的干制红枣品种鉴别[J]. 光谱学与光谱分析, 2017, 37(3):836–840.
FAN Yangyang, QIU Zhengjun, CHEN Jian, et al. Identification of varieties of dried red jujubes with near-infrared hyperspectral imaging[J]. Spectroscopy and Spectral Analysis, 2017, 37(3): 836–840. (in Chinese)
- [23] 苏军,饶元,张敬尧,等. 基于 GA 优化 SVM 的干制红枣品种分类方法[J]. 洛阳理工学院学报(自然科学版), 2018, 28(4):65–69,93.
SU Jun, RAO Yuan, ZHANG Jingyao, et al. Study on classification method of jujube varieties based on GA optimized SVM [J]. Journal of Luoyang Institute of Science and Technology (Natural Science Edition), 2018, 28(4): 65–69,93. (in Chinese)
- [24] ZHANG R C, YUAN Y C, MENG X, et al. A multitask model based on MobileNetV3 for fine-grained classification of jujube varieties[J]. Journal of Food Measurement and Characterization, 2023, 17(5):4305–4317.
- [25] LIN T Y, ROYCHOWDHURY A, MAJI S. Bilinear CNNs for fine-grained visual recognition[J]. arXiv Preprint arXiv: 1504.07889, 2015.
- [26] FU J, ZHENG H, MEI T. Look closer to see better: recurrent attention convolutional neural network for fine-grained image recognition [C] // 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu: IEEE, 2017: 4476–4484.
- [27] ZHENG H, FU J, MEI T, et al. Learning multi-attention convolutional neural network for fine-grained image recognition [C] // 2017 IEEE International Conference on Computer Vision (ICCV). Venice: IEEE, 2017: 5219–5227.
- [28] YANG Z, LUO T, WANG D, et al. Learning to navigate for fine-grained classification [J]. arXiv Preprint arXiv: 1809.00287, 2018.
- [29] CHANG D, DING Y, XIE J, et al. The Devil is in the channels: mutual-channel loss for fine-grained image classification [J]. IEEE Transactions on Image Processing, 2020, 29: 4683–4695.
- [30] CHEN Y, BAI Y, ZHANG W, et al. Destruction and construction learning for fine-grained image recognition [C] // 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Long Beach: IEEE, 2019: 5152–5161.
- [31] GAO Y, HAN X, WANG X, et al. Channel interaction networks for fine-grained image categorization [C] // Proceedings of the AAAI Conference on Artificial Intelligence, 2020: 10818–10825.