

## 基于改进 ResNet18 的干香菇等级识别<sup>\*</sup>

王莉<sup>1</sup> 董鹏豪<sup>1</sup> 王瞧<sup>1,2</sup> 牛群峰<sup>1</sup>

(1. 河南工业大学电气工程学院 郑州 450000; 2. 河南工业大学粮食信息处理与控制重点实验室 郑州 455000)

**摘要:**为解决干香菇等级识别技术复杂及识别精度不高的问题,提出了一种基于残差神经网络 ResNet18 的干香菇等级识别方法。首先将传统的 ResNet18 中 Stem 的  $7 \times 7$  卷积层替换为 3 个  $3 \times 3$  卷积层串联,保证在感受野保持不变的情况下进一步减小计算量;其次针对残差块中线性变换和非线性变换不足的问题,引入融合非对称卷积和 h-swish 激活函数,增加了模型的复杂性,使其能够进行更深层次的特征学习;最后在 ResNet18 骨干网络中引入高效通道注意力机制,加强模型提取特征的能力;实验结果表明,改进后的 ResNet18 网络模型准确度达 97.04%,相比 ResNet18 网络模型方法提升了 4.81%,且性能优于 VGG16、MobileNetV2、DenseNet121、ResNet34 等网络模型方法,可提高干香菇等级的识别精度,单幅图像的检测时间为 5.91 ms,对于香菇智能分拣过程中的等级识别具有借鉴意义。

**关键词:**干香菇分级;机器视觉;ResNet18;高效通道注意力机制

**中图分类号:** TP391 **文献标识码:** A **国家标准学科分类代码:** 520.2

## Dried shiitake mushroom grade recognition based on improved ResNet18

Wang Li<sup>1</sup> Dong Penghao<sup>1</sup> Wang Qiao<sup>1,2</sup> Niu Qunfeng<sup>1</sup>

(1. College of Electrical Engineering, Henan University of Technology, Zhengzhou 450000, China; 2. Key Laboratory of Grain Information Processing and Control, Henan University of Technology, Zhengzhou 450000, China)

**Abstract:** To solve the problems of complexity and low recognition accuracy of dried shiitake mushroom grade recognition technology, a method of dried shiitake mushroom grade recognition based on residual neural network ResNet18 is proposed. Firstly, the  $7 \times 7$  convolutional layer of Stem in the traditional ResNet18 is replaced by three  $3 \times 3$  convolutional layers in series, which ensures that the computational amount is further reduced while the sensory field remains unchanged. Secondly, to address the problem of insufficient linear and nonlinear transformations in the residual block, fused asymmetric convolution and h-swish activation function are introduced, which increases the complexity of the model and enables it to carry out a deeper level of feature learning. Finally, an efficient channel attention mechanism is introduced into the ResNet18 backbone network to strengthen the ability of the model to extract features. The experimental results show that the improved ResNet18 network model has an accuracy of 97.04%, which is 4.81% higher compared to the ResNet18 network modeling method, and outperforms VGG16, MobileNetV2, DenseNet121, ResNet34 and other network model methods, which can improve the recognition accuracy of dried shiitake mushroom grades, and the detection time of a single image is 5.91 ms, which is useful for grade recognition in the intelligent sorting process of dried shiitake mushrooms.

**Keywords:** dried shiitake mushroom grading; machine vision; ResNet18; efficient channel attention mechanism

### 0 引言

香菇是担子菌门、蘑菇纲、蘑菇目、口蘑科和香菇属的

一个属,是一种食用历史悠久且首次驯化栽培的优良食用菌<sup>[1]</sup>。其肉质肥厚细腻,味道鲜美,香气独特,且药食两用,含有丰富的营养物质以及活性物质<sup>[2-3]</sup>,近年来国内外

收稿日期:2023-08-02

<sup>\*</sup> 基金项目:河南工业大学创新基金(2022ZKCJ03)、河南省科技研究计划(2013000210100)项目资助

对香菇需求呈明显上升趋势。常见干香菇主要分为干花菇、干厚菇、干薄菇3种,不同种类的干香菇可根据外观特征分为特级、一级、二级3种等级,等级越高代表着干香菇品质越好。目前干香菇等级的识别主要依赖于人工分选法、称重和基于筛孔的机械分级方法。人工分选法是通过有经验的工人进行人工识别,该方法效率低,准确性受人工经验影响波动较大。称重法是根据香菇的重量进行分选,效率上虽然有所提升,但无法利用干香菇顶盖的纹理特征来区分干花菇。基于筛孔的机械分级方法是通过筛孔尺寸的不同来实现对干香菇的分级,该方法处理快速但同样无法利用干香菇顶盖的纹理特征,另外,部分异形的干香菇可能无法适应固定尺寸的筛孔,导致分级不准确。

近年来,深度学习<sup>[4]</sup>在各个领域内掀起了研究热潮。深度学习结合机器视觉已经广泛的应用到各个场景之中,其优秀的特征提取能力极大地减少了图像处理任务的工作量<sup>[5]</sup>。汪威等<sup>[6]</sup>提出了一种基于机器视觉和图像处理的去柄鲜香菇智能分级方法,对香菇各区域的圆度、直径、均匀性等几何特征进行量化,设置不同的分级标准,最终实现香菇的分级,但主要是针对鲜香菇分级。陈红等<sup>[7]</sup>从蘑菇帽中截取合适的纹理区域,利用灰度直方图统计、灰度共生矩阵、高斯马尔可夫随机场模型和分形维数模型进行特征提取,最后构造了K最近邻分类器对图像进行特征分类,但主要是对不同种类的干香菇进行分类。时宇等<sup>[8]</sup>设计了一种基于菌盖开裂纹理特征的干香菇分类方法,该方法通过对干香菇的纹理特征进行分析,建立了一种品质因数计算公式,从而对干香菇菌盖开裂纹理的进行量化计算得到分类指标。但其数据集原始样本过少,对宏观上差异较小的干香菇难以区分,泛化能力差。

本文以干香菇为研究对象,采集干香菇图像并制作数据集,对ResNet18模型进行网络改进,并与其他传统神经网络模型(VGG16<sup>[9]</sup>、MobileNetV2<sup>[10]</sup>、DenseNet121<sup>[11]</sup>、ResNet34<sup>[12]</sup>、ResNet18<sup>[13]</sup>)进行对比分析,验证本文模型对于干香菇识别的性能提升。通过对网络改进,实现快速准确地识别不同等级的干香菇,为改进人工分拣方式提供解决方案,提升干香菇品质保证。

## 1 材料与方法

### 1.1 图像采集系统

图像采集系统如图1所示,主要包括计算机、载物平台、工业相机、镜头、环形光源、白色平衡卡等。相机选用的型号是海康MV-CE100-30GC 1000万像素工业相机,配备MVL-HF1224M-10MP 12 mm焦距的海康光学工业镜头。为了突出干香菇的颜色及纹理细节特征,拍摄背景选用白色平衡卡,根据实际图像拍摄效果,拍摄物距定为45 cm。

以河南省西峡县地区的干香菇为研究对象,按照香菇



图1 图像采集平台

国家农业行业标准进行分级,采集干花菇特级、干花菇一级、干花菇二级、干厚菇特级、干厚菇一级、干薄菇特级共6个等级1500个干香菇。采集的6种不同等级干香菇如图2所示。

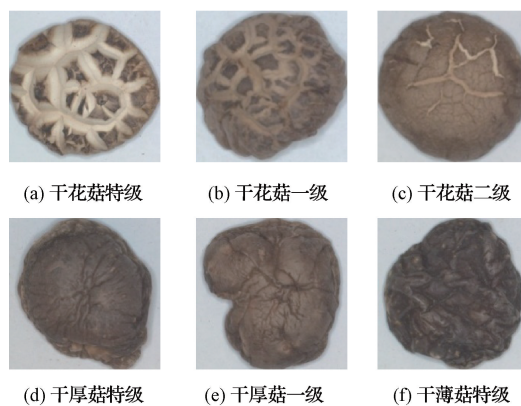


图2 不同等级干香菇类型

获取到所有干香菇图像后,人工剔除模糊、拍摄背景不完整等不合格的干香菇图像,筛选后得到1355张干香菇图像,其中干花菇特级184张、干花菇一级223张、干花菇二级212张、干厚菇特级249张、干厚菇一级222张、干薄菇特级265张。

### 1.2 图像预处理

干香菇原始采集图像尺寸为 $3\,840 \times 2\,748$  pixels,整体图像尺寸较大,而干香菇目标所占区域较小,背景中包含了较多无关信息,为了加快模型的训练以及收敛速度,需要对原始采集图像做进一步预处理工作,精准定位干香菇所占区域。预处理流程如图3所示。

首先对采集图像进行灰度化,然后通过最大类间方差法<sup>[14]</sup>分析得到区分干香菇所在前景区域与无关背景的阈值,利用此阈值将灰度图像转化为二值图像;接着使用OpenCV中的轮廓提取函数对干香菇主体轮廓进行提取,并进一步裁剪出包围干香菇最大轮廓的最小外接矩形。为了避免干香菇图像比例失调,根据最小外接矩形的长宽情况,将裁剪得到的干香菇图像均匀地扩展为最小外接正方形。这样做可以保证裁剪后的图像具有一致的长宽尺

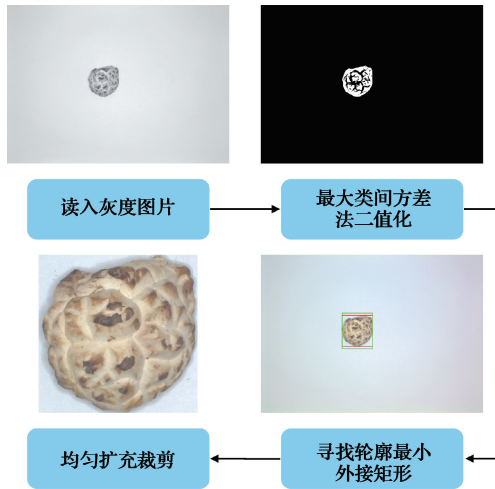


图3 图像预处理流程

寸,符合模型学习的要求。

### 1.3 数据增强

鉴于训练模型时图像数量不足可能导致过拟合现象,为了提升训练模型的泛化能力,采取数据增强的方法对干香菇图像数据量进行扩充。考虑到不同等级的干香菇在颜色上的差异较小,为避免影响分级精度,故使用第三方Python库imgaug对现有数据集进行水平翻转、垂直翻转、等比例缩放、添加高斯噪声增强操作,每种方法处理后图像数据增加1倍,增强后数据集图像共6 775幅。部分干香菇图像增强方法处理效果如图4所示。将数据集以7:3的比例随机划分训练集和测试集,表1为增强后干香菇数据集。

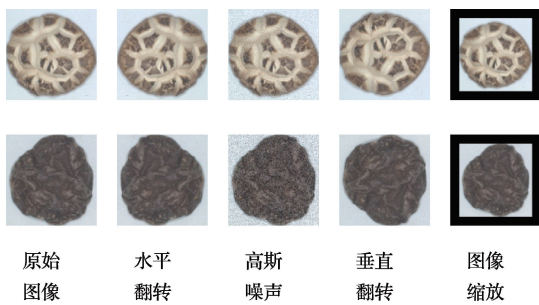


图4 部分干香菇样本数据增强效果

表1 增强后干香菇数据集

干香菇类别	训练集/张	测试集/张
干花菇特级	644	276
干花菇一级	781	334
干花菇二级	742	318
干厚菇特级	872	373
干厚菇一级	777	333
干薄菇特级	928	397
合计	4 744	2 031

## 2 干香菇等级识别模型

本文所提出的改进型ResNet18,主要包含4个部分:1)ResNet18作为骨干网络模块,负责特征提取和信息传递;2)Stem模块使用3个 $3\times 3$ 卷积层串联替换 $7\times 7$ 卷积层,增加网络深度的同时减少计算成本;3)改进残差模块,使用融合非对称卷积层替换 $3\times 3$ 卷积层,有助于在特征表示中获得更丰富的信息;4)在模型中融入高效通道注意力机制,增强网络的特征表达能力,使得模型能够更好地学习图像中的复杂特征。

### 2.1 ResNet18 骨干网络

为保证干香菇分级任务的高准确率和实时性,选取网络层数适中,收敛速度快,能够平衡训练速度和网络深度的ResNet18作为骨干网络,该网络是一种卷积神经网络(convolutional neural network,CNN)结构,于2015年作为残差网络(residual network,ResNet)系列的一部分被提出,旨在通过使用跳跃连接或快捷连接来解决深度神经网络中梯度消失的问题。

ResNet18网络结构如图5所示,该网络包含18个层,其中包括16个卷积层和2个全连接层。卷积层被组织成多个残差块,每个块包含多个卷积层、批归一化和ReLU激活函数。ResNet18骨干网络中,输入网络为一幅 $224\times 224$ 的三通道RGB图像,经过特征提取和降采样后得到一个 $N\times 6\times 7\times 7$ ( $N$ 为一个mini-batch尺寸)大小的特征图。随后通过自适应平均池化层,将特征图继续降采样为 $N\times 6\times 1\times 1$ ,最后将该提取到的干香菇细节特征输入Softmax分类器计算不同类型干香菇等级预测结果。

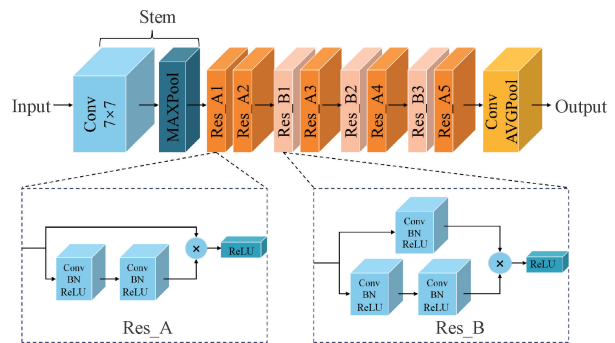


图5 ResNet18 骨干网络

### 2.2 Stem 卷积核替换

ResNet18结构中,Stem模块是网络的初始部分,负责将输入图像进行一系列卷积和池化操作,从而将图像转换为一组特征图。Stem阶段在整个网络中起着提取初级特征的作用,这些特征会随后在后续的层中不断进行进一步的特征提取和组合,同时降低输入图像的分辨率,减少计算成本。Stem模块由步幅为2的 $7\times 7$ 卷积层和一个步幅为2的 $3\times 3$ 最大池化层组成。模型输入是 $224\times 224$



的图像,感受野使用较大的  $7 \times 7$  卷积核,尽可能保留特征信息。然后将  $7 \times 7$  卷积层和  $3 \times 3$  最大池化层的步幅设为 2,进行降采样,减少计算量。通过上述操作,图像输入分辨率降低到  $56 \times 56$ 。

感受野的计算公式如下:

$$F_i = (F_{i-1} - 1) \times \text{Stride} + K_{\text{size}} \quad (1)$$

式中:  $F_i$  为第  $i$  层感受野;  $\text{Stride}$  为第  $i$  层步距;  $K_{\text{size}}$  为卷积核或池化核尺寸。

假设网络输入的感受野大小为 1,经过一层步幅为 2 的  $7 \times 7$  卷积层,感受野大小为:

$$F_1 = (1 - 1) \times 2 + 7 = 7 \quad (2)$$

保持网络输入的感受野大小为 1,网络第 1 层经过一层步幅为 2 的  $3 \times 3$  卷积层,后面串联两个步幅为 1 的  $3 \times 3$  卷积层,感受野计算过程如下:

$$F_1 = (1 - 1) \times 2 + 3 = 3 \quad (3)$$

$$F_2 = (3 - 1) \times 1 + 3 = 5 \quad (4)$$

$$F_3 = (5 - 1) \times 1 + 3 = 7 \quad (5)$$

由式(2)~(5)可知,用 3 个  $3 \times 3$  卷积层串联替换初始的  $7 \times 7$  卷积层,感受野保持不变。此外,假设输入输出的通道数都为  $C$ ,将卷积核尺寸乘以输入和输出的通道数即可计算出该层卷积层的参数量,那么使用  $7 \times 7$  卷积层所需参数为:

$$P_1 = 7 \times 7 \times C \times C = 49C^2 \quad (6)$$

堆叠 3 个  $3 \times 3$  卷积层所需参数为:

$$P_2 = 3 \times 3 \times C \times C + 3 \times 3 \times C \times C + 3 \times 3 \times C \times C = 27C^2 \quad (7)$$

由式(6)和(7)可以看出,用 3 个  $3 \times 3$  卷积层串联替换初始的  $7 \times 7$  卷积层可大幅降低参数量。同时,深层网络通常具有更强的特征提取能力,可以更好地捕捉干香菇图像中的细节和高级特征。因此,本文将 ResNet18 网络 Stem 中的一个  $7 \times 7$  卷积层替换为 3 个  $3 \times 3$  卷积层串联,如图 6 所示。

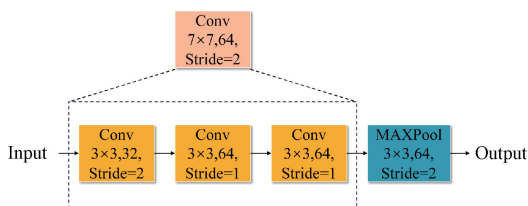


图 6 Stem 模块改进

### 2.3 改进残差模块

ResNet 网络模型的残差块采用  $3 \times 3$  卷积操作,用于提取输入特征的高级表示。这些特征经过 ReLU 激活函数进行非线性变换,并随后进行批量归一化(batch normalization, BN)操作,以增强模型对不同尺度的特征的学习能力。尽管 ResNet 模型在特征提取方面取得了显著的进展,但 ResNet 的残差块由线性变换和非线性变换相结合的这种构成方式可能导致特征向量的表示能力受限。

干香菇的不同等级主要受颜色和纹理轮廓细节的影响,这两个因素对于香菇等级的判定至关重要。不同等级的干香菇存在相似之处,如干厚菇二级和干薄菇特级都存在颜色较深、轮廓表面隆起的小细节纹理,这导致识别问题加大难度。若采用原有的网络结构去提取干香菇细节特征容易造成特征信息丢失,影响识别精度。针对以上问题对残差块 Res\_A 和 Res\_B 做出以下改进。

首先,针对  $3 \times 3$  卷积核特征提取能力不强的问题,引入融合非对称卷积<sup>[15]</sup>。该卷积操作是利用  $3 \times 3$ 、 $1 \times 3$  和  $3 \times 1$  并行的卷积核对图像进行平行特征提取,并做 BN 处理。如尺寸为  $U \times V$ 、通道数为  $C$  的特征图输入到尺寸为  $H \times W$ 、通道数为  $D$  的卷积核。其中,  $\mathbf{M} \in \mathbf{R}^{U \times V \times C}$  表示输入,  $\mathbf{F} \in \mathbf{R}^{H \times W \times C}$  表示卷积核,  $\mathbf{O} \in \mathbf{R}^{R \times T \times D}$  代表输出特征图。对于这个层的第  $j$  个卷积核,相应的输出特征映射通道计算方式为:

$$O_{:,j} = \sum_{k=1}^C M_{:,k} * F_{:,k}^{(j)} \quad (8)$$

式中:  $*$  表示二维卷积算子;  $M_{:,k}$  表示第  $k$  个通道尺寸是  $U \times V$  的特征图;  $F_{:,k}^{(j)}$  表示第  $k$  个通道尺寸是  $H \times W$  的特征图。输出特征经过 BN 处理,输出特征映射通道计算方式为:

$$O_{:,j} = \left( \sum_{k=1}^C M_{:,k} * F_{:,k}^{(j)} - \mu_j \right) \frac{\gamma_j}{\sigma_j} + \beta_j \quad (9)$$

式中:  $\mu_j$  和  $\sigma_j$  表示通道平均值和标准差;  $\beta_j$  和  $\gamma_j$  表示缩放系数和偏移量。卷积具有可加性,经过融合非对称卷积和 BN 处理后,输出特征映射通道计算方式为:

$$O_{:,j} + \tilde{O}_{:,j} + \hat{O}_{:,j} = \sum_{k=1}^C M_{:,k} * F_{:,k}^{(j)} + b_j \quad (10)$$

式中:  $O_{:,j}$ 、 $\tilde{O}_{:,j}$  和  $\hat{O}_{:,j}$  分别表示  $3 \times 3$ 、 $1 \times 3$  和  $3 \times 1$  卷积核的输出;  $b_j$  表示偏置。

加入融合非对称卷积和 BN 操作后,可以提高模型线性变换和特征细节提取的能力。

其次对激活函数进行修改。激活函数给网络模型带来了非线性因素,增强神经网络的表达能力,从而提高最终的分级性能<sup>[16]</sup>。ReLU 激活函数属于分段函数,其函数表达式为:

$$\text{ReLU}(x) = \max(0, x) \begin{cases} x, & x \geq 0 \\ 0, & x < 0 \end{cases} \quad (11)$$

式中:  $x$  表示输入特征向量。

由式(11)可以看出,输入值  $x < 0$  时,函数的值恒为 0,此时函数失去了激活作用,导致相应的参数不会被更新,干香菇图像部分信息就会丢失,影响识别精度。为了解决这一问题,本文使用 h-swish 激活函数代替残差块中的 ReLU 激活函数, h-swish 激活函数表达式为:

$$h\text{-swish}(x) = x \cdot \frac{\text{ReLU6}(x + 3)}{6} \quad (12)$$

ReLU6 激活函数的数学表达式为:

$$\text{ReLU6}(x) =$$



$$\min(6, \text{ReLU}(x)) \begin{cases} 0, & x \leq 0 \\ x, & 0 < x < 6 \\ 6, & x > 6 \end{cases} \quad (13)$$

h-swish 激活函数引入 ReLU6 激活函数,既保留了 ReLU 激活函数在正半轴良好的稀疏表达能力,又增强了负半轴部分数值,使模型具有更好的非线性特性、灵活的参数控制能力和更稳定的梯度传播。改进后的残差网络结构如图 7 所示。

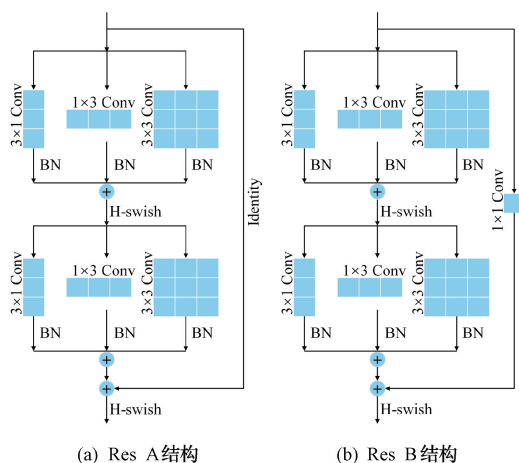


图 7 改进残差网络结构

## 2.4 高效通道注意力机制模块

近年来,注意力机制在基于深度学习的卷积神经网络分类任务中得到了广泛的应用。通道注意力网络(squeeze-and-excitation networks, SENet)是由 Hu 等<sup>[17]</sup>在 2017 年提出的一种深度学习网络结构,它引入了注意力机制来增强模型对重要特征的关注。SENet 网络结构如图 8 所示。

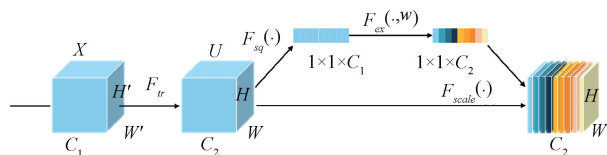


图 8 SENet 结构

然而,SENet 中的降维虽然可以降低模型的复杂度,但对获取通道之间的关联效率低,会降低通道注意力预测的效果。针对以上问题,Wang 等<sup>[18]</sup>提出了一种不降维的局部跨通道信息交互策略(efficient channel attentional network, ECANet)。ECANet 模块在进行全局平均池化操作后,并不直接将特征图降维为一个数值。相反,它引入了一个可学习的缩放因子(scale factor),这个缩放因子是一个通道维度的向量,它的长度与通道数相等。通过学习得到通道维度的缩放因子,来实现对通道的注意力加权,从而增强重要特征的表达。这样的设计使得 ECANet 在计算效率上更加高效,同时能够获得类似 SENet 的注意力增强效果。ECANet 结构如图 9 所示。

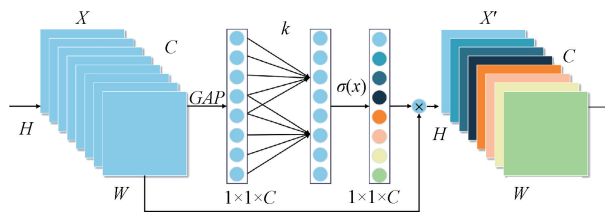


图 9 ECA 结构

ECA 结构 ECANet 的实现可以分为两个步骤,第 1 步对输入特征图进行全局平均池化,将每个通道的特征图降维至一个数值,得到  $1 \times 1 \times C$  的全局描述特征,表示每个通道的全局重要性。在第 1 步得到的全局描述特征基础上,通过卷积核尺寸为  $k$  的  $1 \times 1$  卷积来进行部分通道特征学习,再通过 sigmoid 激活函数作用计算所有通道的缩放因子  $\omega$ 。它的长度与通道数  $C$  相等。然后,将每个通道的特征图与对应的缩放因子相乘,以增强重要特征的表达,抑制不重要的特征。这种注意力机制在计算效率上较高,同时能够提高模型的性能和泛化能力。其中每个通道的权重  $\omega$  如下:

$$\omega = \sigma(C_1 D_k(y)) \quad (14)$$

式中:  $\sigma$  为 sigmoid 激活函数;  $C_1 D$  代表一维卷积;  $y$  表示通道;  $k$  为卷积核大小,  $k$  的大小和通道自适应函数有关,如式(15)所示。

$$C = 2^{(r \times k - b)} \quad (15)$$

给定通道维度  $C$ , 卷积核  $k$  的大小可以根据式(16)来确定。

$$k = \left\lceil \frac{\log_2 C}{\gamma} + \frac{b}{\gamma} \right\rceil_{\text{odd}} \quad (16)$$

式中:  $\lceil X \rceil_{\text{odd}}$  表示  $k$  的值只能为奇数,为了减少训练过程的时间和计算成本,设置超参数  $\gamma=2, b=1$ 。

激活函数使用整流线性单元(ReLU6)控制模型的复杂性,有效地抑制梯度爆炸问题,同时减少过拟合的风险。ReLU6 的函数解析式为:

$$\text{ReLU6}(x) = \min(\max(0, x), 6) \quad (17)$$

将 ECANet 模型嵌入到 ResNet18 骨干网络中,每经过一个 Res\_A 或 Res\_B 残差模块之后就通过一个 ECANet 网络进行通道学习,干香菇等级识别最终网络结构如图 10 所示,命名为 ECA-AC-ResNet18。

## 3 实验与结果分析

### 3.1 实验平台

实验环境为 Windows11 操作系统, GPU 型号为 GeForce RTX 2070, 显存 8 G, 处理器使用的是 Inter(R) Core(TM) i5-12400CPU@4.4 GHz, 运行内存 16 G。模型开发环境使用 Pycharm, Python 版本为 3.9, 通过 Pytorch 深度学习框架进行模型搭建, 算法使用 CUDA 和 CUDnn 库进行加速运算。

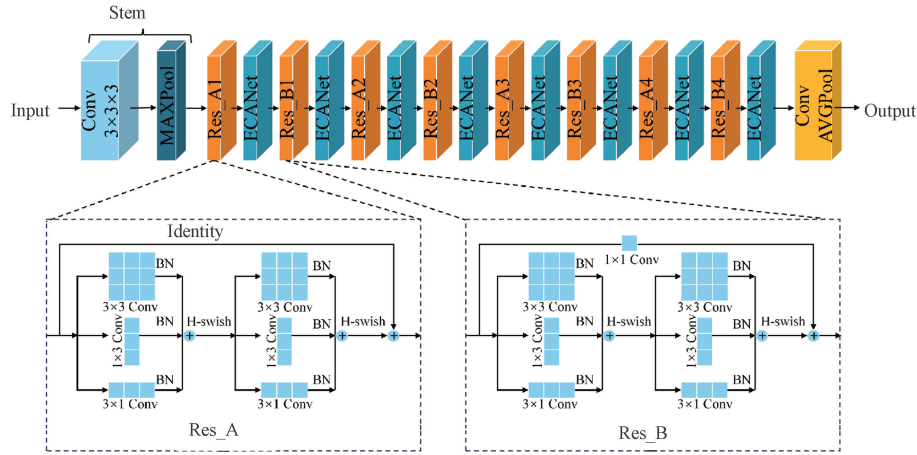


图 10 ECA-AC-ResNet18 网络结构

### 3.2 参数设置

干香菇图像在送入模型之前像素尺寸不一,因此使用双线性插值<sup>[19]</sup>来尽可能保证图像缩放后的质量,干香菇数据集中的每个图像都被缩放到  $224 \times 224$  pixels 大小。对每个通道的数据进行标准化处理,使用均值 0.5 和标准差 0.5 进行标准化。这样做可以使不同特征具有相同的尺度,从而提高模型的学习效率。此外,对输入的训练集图像进行随机打乱,以减少图像顺序对模型的影响。通过多次迭代反馈的梯度下降,使得模型能够跳出局部最优,达到全局最优的分级模型。

模型优化器选择使用 Adam 优化器。通过自适应学习策略 Adagrad 结合动量下降算法,能够适应稀疏梯度且缓解梯度震荡问题,使模型性能更加稳定达到收敛。设置初始学习率为 0.000 1, batchsize 设置为 32, 迭代次数为 50 次。每迭代一轮 epoch 后,通过验证集来测试模型的准确性,并保存每轮 epoch 生成的模型和验证结果。损失函数使用交叉熵损失函数<sup>[20]</sup>:

$$C_{loss} = -\frac{1}{m} \sum_{i=1}^m \sum_{j=1}^4 y_{ji} \log(\hat{y}_{ji}) \quad (18)$$

式中:  $m$  为当前一批输入网络的图像数量;  $y_{ji}$  为真实的标签;  $\hat{y}_{ji}$  为预测的标签;  $C_{loss}$  为损失值。一般训练过程中交叉熵的损失值越小,表示模型的预测结果与目标值之间的差距越小,即模型的学习效果越好。

### 3.3 模型评价指标

本文使用模型在干香菇测试集上的准确率(accuracy, ACC)、精确率(precision, P)、召回率(recall, R)、 $F_1$  分数作为评价指标,用来评估模型的性能。其中,  $F_1$  是融合了查准率和查全率的综合指标,  $F_1$  值分数越高,表示模型整体越好<sup>[21]</sup>。

ACC 表示预测正确的样本占总样本的比例,公式为:

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \quad (19)$$

P 表示预测为正的样本实际为正样本的比例,公

式为:

$$P = \frac{TP}{TP + FP} \quad (20)$$

R 表示所有正样本中被预测为正样本的比例,公式为:

$$R = \frac{TP}{TP + FN} \quad (21)$$

$F_1$  分数是 Pre 和 Rec 的调和平均值,公式为:

$$F_1 = 2 \times \frac{P \times R}{P + R} \quad (22)$$

式中:  $TP$  是正确分类为属于该类的样本数;  $TN$  是正确分类为不属于该类的样本数;  $FP$  是错误分类为属于该类的样本数;  $FN$  是错误分类为不属于该类的样本数。

### 3.4 不同改进方案对比分析

为确定本文模型改进中所采用的 Stem 模块卷积核尺寸调整、融合非对称卷积和高效通道注意力机制操作对干香菇等级识别能力的具体提升效果,对 Stem 模块、残差块、ECA-Net 进行组合设计测试试验。试验组合如表 2 所示,其中方案 5 为本文提出的 ECA-AC-ResNet18 模型。

表 2 测试设计方案

方案 编号	Stem 模块		残差块		ECA-Net	
	$7 \times 7$	$3 \times 3 \times 3$	普通 卷积	融合非对 称卷积	不使用	使用
1	✓		✓		✓	
2	✓			✓	✓	
3		✓	✓		✓	
4		✓		✓	✓	
5		✓		✓		✓

测试结果如表 3 所示,由试验结果可知,对比方案 1 和方案 3,使用 3 个  $3 \times 3$  卷积层串联替换  $7 \times 7$  卷积层模型识别准确率提升了 1.03%,这主要是因为 3 个  $3 \times 3$  卷积层串联增加了网络的深度。深层网络通常具有更强的

特征提取能力,可以更好地捕捉干香菇图像中的细节和高级特征,从而提高模型的表达能力。同时这可以在一定程度上降低网络的计算复杂性,加快模型训练和推理的速度。

表3 各试验方案测试结果

方案 编号	准确率 /%	召回率 /%	$F_1$ 分数	内存大小 /MB
1	92.22	92.05	92.11	42.7
2	93.61	93.42	93.47	43.0
3	93.25	93.03	93.24	42.9
4	95.72	95.56	95.75	43.2
5	97.04	97.19	97.05	43.2

对比方案1和方案2,使用融合非对称卷积替换原本的普通卷积块模型识别准确率提升了1.39%,一方面,这主要是因为融合非对称卷积使用了不同尺寸的卷积核,从而增加了卷积操作的感受野,感受野的增加使得网络可以更好地理解输入干香菇图像的全局结构和上下文信息,有助于提高对干香菇图像细节的特征抽取能力;另一方面,融合非对称卷积块的引入增加了模型的复杂性,使其能够进行更深层次的特征学习。这有助于提高网络的特征提取能力和判别能力,从而提高识别准确率。

对比方案4和方案5,ECANet的引入使模型识别准确率提升了1.32%,这主要是因为ECANet可以对每个通道的特征进行加权,增强重要特征的表达,使得模型能够更好地捕捉干香菇图像中的有用信息。这有助于提高特征的多样性和丰富性,进而提高了识别准确率。

综上所述,当同时使用上述改进后,方案5即本文提出的ECA-AC-ResNet18模型的干香菇等级识别结果最佳,准确率达到97.04%。相比最初的ResNet18网络模型,准确率提升了4.82%。模型占用内存大小43.2 MB,在内存增加较少的情况下大幅度的提升了干香菇等级识别的准确率。

### 3.5 模型性能分析

为了更清楚地展示ECA-AC-ResNet18模型的识别精度和分级结果,对测试集上的模型分级结果进行统计,得到一个针对不同等级干香菇识别详情的混淆矩阵,如图11所示。结合混淆矩阵分析,本文提出的方法适用于干花菇特级、干花菇一级、干花菇二级、干厚菇特级、干厚菇一级、干薄菇特级6种不同等级的干香菇。在识别不同干香菇等级类型时,每种干香菇的准确性和召回率是不同的,ECA-AC-ResNet18模型网络针对不同等级干香菇的具体性能如表4所示。可以看出,各个等级干香菇的识别精度保持在90.86%~99.21%,平均为97.01%;召回率保持在94.46%~99.06%,平均为97.19%。 $F_1$ 分数为

94.52%~98.90%,平均为97.05%。上述结果表明,所提出的方法在既定的数据集中表现良好,可以应用于不同干香菇等级在线识别。

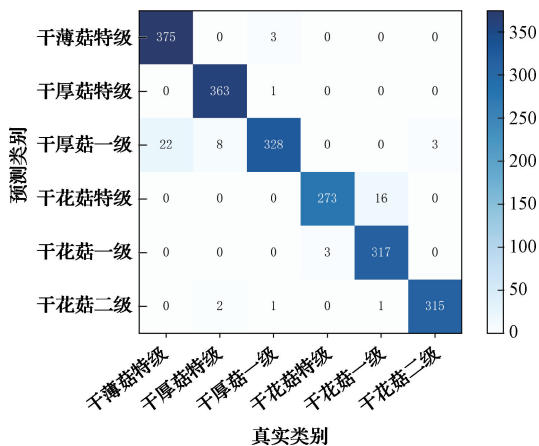


图11 使用ECA-AC-ResNet18的不同等级干香菇识别混淆矩阵

表4 ECA-AC-ResNet18网络在不同等级干香菇上的性能表现 (%)

干香菇种类	精确率	召回率	$F_1$ 分数
干薄菇特级	99.21	94.46	96.77
干厚菇特级	99.73	97.32	98.51
干厚菇一级	90.86	98.50	94.52
干花菇特级	94.46	98.91	96.64
干花菇一级	99.06	94.91	96.94
干花菇二级	98.75	99.06	98.90
平均值	97.01	97.19	97.05

### 3.6 不同算法模型对比

为了进一步验证ECA-AC-ResNet18模型的有效性,选用5种具有代表性的神经网络模型MobileNetV2、DenseNet121、VGG16、ResNet34、ResNet18进行对比实验,对比实验结果如表5所示。图12和13所示为不同模型在训练过程中的准确率对比和损失值对比。

表5 改进ResNet18网络与其他网络对比

模型	准确率 /%	召回率 /%	$F_1$ 分 数/%	耗时 /ms
MobileNetV2	84.09	83.82	84.15	24.13
DenseNet121	87.54	87.18	88.63	6.40
VGG16	88.97	88.87	88.96	97.49
ResNet34	90.89	90.60	90.82	32.50
ResNet18	92.22	92.05	92.11	4.43
ECA-AC-ResNet18	97.04	97.19	97.05	5.91

表5为6种不同网络的网络评估与耗时结果。由实



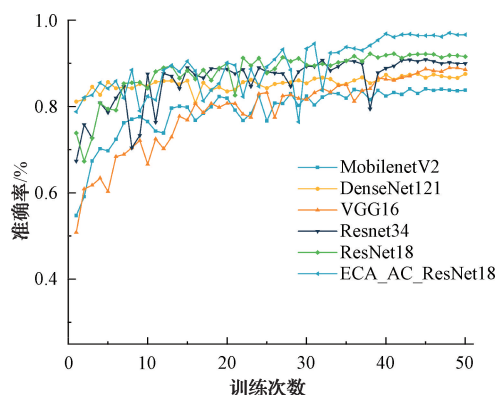


图12 不同模型准确率对比曲线

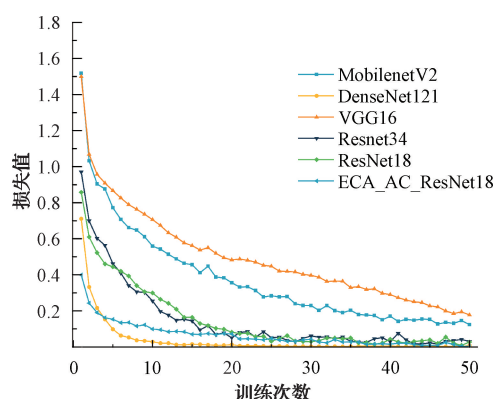


图13 不同模型损失值对比曲线

验结果可知,本文提出的 ECA-AC-ResNet18 模型的干香菇等级识别性能更加均衡,平均识别准确率为 97.04%,召回率为 96.84%, $F_1$  分数为 96.97%,平均识别单张干香菇图像时间为 5.91 ms。相比 MobileNetV2、DenseNet121、VGG16、ResNet34、ResNet18 这些网络,准确率分别提升了 12.95%、9.50%、8.07%、6.15%、4.82%,同时识别速度相对较快。对比实验结果表明,本文提出的 ECA-AC-ResNet18 模型在不同等级干香菇识别中具有更加优秀的性能,特别是在识别准确率这一指标上,具有明显优势。

#### 4 结 论

为了解决干香菇等级识别中的一系列问题,本文提供了干香菇图像拍摄、图像预处理、建立数据集和等级识别的整套方案。提出了一种基于 ECA-AC-ResNet18 网络的干香菇等级识别方法。首先,调整 ResNet18 网络中 Stem 模块卷积尺寸,用 3 个  $3 \times 3$  的卷积核替换原来的  $7 \times 7$  卷积核,在保证感受野不变的情况下大大减少了参数量。其次,对 ResNet18 网络中残差模块进行了改进,在残差块中引入融合非对称卷积并使用 h-swish 激活函数替代原有的 ReLU 激活函数,提升了模型对不同等级干香菇细节纹理的提取能力,降低模型对空间信息的依赖性,

提高干香菇等级识别精度。最后,在 ResNet18 网络中引入高效通道注意力策略,通过自适应选择一维卷积核来确定跨通道覆盖率,在保持高效通道注意力性能的同时降低了计算成本,加强模型提取特征的能力。最终,ECA-AC-ResNet18 网络识别准确率为 97.04%,损失值为 0.002 32,平均  $F_1$  分数为 97.05%,识别单张干香菇图像时间为 5.91 ms。试验结果表明,ECA-AC-ResNet18 网络准确率相比典型网络 MobileNetV2、DenseNet121、VGG16、ResNet34、ResNet18 分别提高 12.94%、9.49%、8.06%、6.14%、4.81%,模型性能提升显著。

本文提出的网络模型能够有效地对不同等级的干香菇进行识别,然而研究也存在一些局限性。首先,本文所进行的研究是对实际现场中部分采样进行训练,并未在实际工业环境中进行实验。其次,模型训练的过程中没有利用不同等级干香菇的尺寸信息,因此遗失了与干香菇尺寸相关的特征。未来的研究方向应考虑如下两点。1)设计的图像采集装置需要进一步优化,以获得更高质量的图像,同时考虑如何将其应用到实际工业环境中,并就本文所提出的方法和模型进行实际验证。2)探究不同等级干香菇自身的几何特征,如长度、宽度、面积和纵横比等对干香菇等级识别的影响,将这些特征与图像信息共同输入网络。

#### 参 考 文 献

- [1] 曹贤,邹明,高俊峰,等.基于氨基酸含量分析 14 类香菇的品质特性[J].中国瓜菜,2023,36(8):48-55.
- [2] RATHORE H, PRASAD S, SHARMA S. Mushroom nutraceuticals for improved nutrition and better human health: A review[J]. Pharma Nutrition, 2017, 5(2): 35-46.
- [3] 周伟,凌亮,郭尚.香菇食药价值综述[J].食药菌,2020,28(6):461-465,469.
- [4] 苏静雷,王红军,王政博,等.多通道卷积神经网络和迁移学习的燃气轮机转子故障诊断方法[J].电子测量与仪器学报,2023,37(3):132-140.
- [5] 王芳,崔丹丹,李林.基于深度学习的采摘机器人目标识别定位算法[J].电子测量技术,2021,44(20):162-167.
- [6] 汪威,刘亚川,吕斌,等.一种去柄鲜香菇视觉分级系统设计[J].食品与机械,2021,37(3):105-111.
- [7] 陈红,夏青,左婷,等.基于纹理分析的香菇品质分选方法[J].农业工程学报,2014,30(3):285-292.
- [8] 时宇,孙冬,高清维,等.基于纹理特征的干香菇分类方法[J].国外电子测量技术,2022,41(1):158-162.
- [9] 孟志超,贺磊盈,杜小强,等.基于 EnhancedVGG16 的油茶品种分类[J].农业工程学报,2022,38(10):176-181.
- [10] 张鹏程,余勇华,陈传武,等.基于改进 MobileNetV2

- 的柑橘害虫分类识别方法[J]. 华中农业大学学报, 2023, 42(3): 161-168.
- [11] 林志玮, 丁启禄, 黄嘉航, 等. 基于 DenseNet 的无人机光学图像树种分类研究[J]. 遥感技术与应用, 2019, 34(4): 704-711.
- [12] ZHUANG Q, GAN S, ZHANG L. Human-computer interaction based health diagnostics using ResNet34 for tongue image classification [J]. Computer Methods and Programs in Biomedicine, 2022, 226: 107096.
- [13] 张文景, 蒋泽中, 秦立峰. 基于弱监督下改进的 CBAM-ResNet18 模型识别苹果多种叶部病害[J]. 智慧农业(中英文), 2023, 5(1): 111-121.
- [14] 郭皓然, 邵伟, 周阿维, 等. 全局阈值自适应的高亮金属表面缺陷识别新方法[J]. 仪器仪表学报, 2017, 38(11): 2797-2804.
- [15] 王文亮, 杨晓迪, 张博雅, 等. 轻量化卷积神经网络在船舶分类中的应用[J]. 激光与光电子学进展, 2023, 60(6): 73-80.
- [16] 王红霞, 周家奇, 辜承昊, 等. 用于图像分类的卷积神经网络中激活函数的设计[J]. 浙江大学学报(工学版), 2019, 53(7): 1363-1373.
- [17] 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.
- [18] WANG Q, WU B, ZHU P, et al. ECA-Net: Efficient channel attention for deep convolutional neural networks[C]. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2020: 11531-11539.
- [19] 杨丽娟, 李利. 基于双线性插值的内容感知图像缩放算法仿真[J]. 计算机仿真, 2019, 36(12): 244-248.
- [20] 冯晓, 李丹丹, 王文君, 等. 基于轻量级卷积神经网络和迁移学习的小麦叶部病害图像识别[J]. 河南农业科学, 2021, 50(4): 174-180.
- [21] KANG J, GWAK J. Ensemble of multi-task deep convolutional neural networks using transfer learning for fruit freshness classification[J]. Multimedia Tools and Applications, 2022, 81(16): 22355-22377.

#### 作者简介

王莉, 博士, 教授, 主要研究方向为图像处理。

E-mail: hautwangli@haut.edu.cn

董鹏豪, 硕士研究生, 主要研究方向为机器视觉与深度学习。

E-mail: dphcontrol@163.com

牛群峰(通信作者), 博士, 副教授, 主要研究方向为机器学习。

E-mail: niuqunfeng@gmail.com