

# 融合双线性网络和注意力机制的油橄榄品种识别

朱学岩<sup>1,2</sup>, 陈锋军<sup>1,2,3\*</sup>, 郑一力<sup>1,2,4</sup>, 李志强<sup>1,4</sup>, 张新伟<sup>1,5</sup>

(1. 北京林业大学工学院, 北京 100083; 2. 林木资源高效生产全国重点实验室, 北京 100083; 3. 城乡生态环境北京实验室, 北京 100083; 4. 林业装备与自动化国家林业局重点实验室, 北京 100083; 5. 智慧林业研究中心, 北京 100083)

**摘要:** 为解决自然条件下的油橄榄品种识别问题, 该研究以油橄榄品种佛奥、莱星、皮削利和鄂植 8 号为研究对象, 融合双线性网络与注意力机制, 提出双线性注意力 EfficientNet 模型。针对不同品种油橄榄表型差异很小的特点, 搭建双线性网络以充分提取油橄榄图像中的特征信息。在此基础上, 选用兼顾了速度和精度的 EfficientNet-B0 网络为特征提取网络。针对自然条件下油橄榄品种识别易受复杂背景干扰的问题, 将 CBAM (convolutional block attention module, CBAM) 注意力与双线性网络结合, 使模型在提取油橄榄图像特征时, 能够聚焦到对油橄榄品种识别起关键作用的特征上。经测试, 所提模型对 4 个油橄榄品种识别的总体准确率达到 90.28%, 推理时间为 9.15 ms。Grad-CAM (gradient-weighted class activation mapping, Grad-CAM) 热力图可视化结果也表明, 所提模型在识别油橄榄品种时重点关注了果实以及部分叶子区域。消融试验结果表明, 在 EfficientNet 模型中引入 CBAM 注意力和搭建双线性网络后, 总体准确率分别提高了 5.00 和 10.97 个百分点。并且, 对比试验结果表明, 与双线性 ResNet34、EfficientNet-SE 注意力、双线性 ResNet18、双线性 VGG16 和双线性 GoogLeNet 等模型相比, 所提模型的总体识别准确率分别高 12.78、11.53、11.11、10.70 和 5.00 个百分点。该研究为解决自然条件下的油橄榄品种识别提供了依据, 同时也可对其他作物的品种识别提供参考。

**关键词:** 图像处理; 模型; 品种识别; 油橄榄; EfficientNet-B0; CBAM 注意力; Grad-CAM

doi: 10.11975/j.issn.1002-6819.202303195

中图分类号: TP391.4; S24

文献标志码: A

文章编号: 1002-6819(2023)10-0183-10

朱学岩, 陈锋军, 郑一力, 等. 融合双线性网络和注意力机制的油橄榄品种识别[J]. 农业工程学报, 2023, 39(10): 183-192. doi: 10.11975/j.issn.1002-6819.202303195 <http://www.tcsae.org>

ZHU Xueyan, CHEN Fengjun, ZHENG Yili, et al. Identification of olive cultivars using bilinear networks and attention mechanisms[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2023, 39(10): 183-192. (in Chinese with English abstract) doi: 10.11975/j.issn.1002-6819.202303195 <http://www.tcsae.org>

## 0 引言

油橄榄是四大木本油料作物之一, 在中国的甘肃陇南、四川达州、云南丽江和四川阆中等地广泛种植。油橄榄果实可以用来榨取富含不饱和脂肪酸的橄榄油, 具有极高的经济价值<sup>[1-2]</sup>。油橄榄品种多且不同品种油橄榄在表型上差异很小, 而在含油率和脂肪酸组成等品质上的差异却较大, 这也导致在市场上以次充好的现象时有发生<sup>[3-4]</sup>。随着油橄榄选育品种的增多, 依靠人工目视的油橄榄品种识别方法已无法满足实际需求。为此, 有必要对自然条件下的油橄榄品种识别展开研究, 为油橄榄种植和生产解决关键的品种识别问题。

为实现品种识别, 国内外研究者主要采用传统图像处理和深度学习技术展开研究。其中, 传统图像处理技术是依靠人工提取颜色<sup>[5]</sup>、纹理和几何特征<sup>[6]</sup>后输入决

策树、支持向量机<sup>[7]</sup>、人工神经网络<sup>[8]</sup>和 BP 神经网络<sup>[9]</sup>等进行品种识别<sup>[10-11]</sup>。在油橄榄<sup>[12]</sup>、小麦<sup>[13]</sup>、玉米<sup>[14-15]</sup>和辣椒<sup>[16]</sup>的品种识别研究中, 广泛使用了基于传统图像处理的品种识别方法。MARTÍNEZ 等<sup>[17]</sup>尝试在提取油橄榄果实的果形指数和最大横径等特征后, 采用偏小二乘判别分析法识别油橄榄品种。经测试, 该方法对 5 个主栽油橄榄品种 of 的识别准确率为 89%。但是, 由于人工手动提取的特征表达能力有限, 导致基于传统图像处理的品种识别方法大多只能应用于实验室环境。在解决自然条件下的品种识别问题时, 基于传统图像识别的品种识别方法会面临复杂光照、复杂背景以及环境噪声等挑战, 难以取得满意效果。

为应对自然条件下品种识别的挑战, 有研究者尝试采用基于数据驱动的深度学习技术对自然条件下的作物品种识别展开研究<sup>[18]</sup>。因为, 深度学习技术可以提取作物图像中不受复杂光照、复杂背景和环境噪声干扰的更稳定的图像特征。目前, 基于深度学习的品种识别方法已被广泛应用在水稻<sup>[19]</sup>、榛子<sup>[20]</sup>、菜豆<sup>[21]</sup>、油菜<sup>[22]</sup>和菊花<sup>[23]</sup>的品种识别中。为识别油橄榄品种, MIHO 等<sup>[24]</sup>融合 DenseNet 和专家经验设计了 OliVaR 模型。经测试, 该模型 of 的油橄榄品种识别准确率超过 86%。OSAKO

收稿日期: 2023-03-27 修订日期: 2023-04-07

基金项目: 国家重点研发计划 (2019YFD1002401); 中央高校基本科研业务费专项资金 (2021ZY74); 北京市共建项目联合资助。

作者简介: 朱学岩, 硕士生, 研究方向为林业信息检测。Email: xueyan0111@bjfu.edu.cn。

\*通信作者: 陈锋军, 博士, 副教授, 研究方向为林业信息检测与智能处理。Email: chenjf227@bjfu.edu.cn。

等<sup>[25]</sup>利用 VGG16 模型对 4 个荔枝品种进行识别, 准确率达到 98.33%。孟志超等<sup>[26]</sup>提出一种 Enhanced VGG16 模型用于识别油茶品种, 准确率为 98.44%。LIU 等<sup>[27]</sup>采用 VGG16 和 ResNet50 模型识别菊花品种, 准确率分别为 89.43% 和 95.39%。先前的研究已经表明, 基于深度学习的图像识别技术在作物品种识别中展现出了强大的潜力。

与油茶和荔枝的品种识别相比, 自然条件下的油橄榄品种识别存在的挑战更多且更复杂。这是因为油橄榄果实更小且不同品种的油橄榄在表型上差异很小<sup>[28]</sup>。因此, 这就要求油橄榄品种识别模型不仅能充分提取到油橄榄图像的特征信息, 而且能聚焦对品种识别起关键作用的图像特征。充分提取输入图像中的特征, 双线性网络的优势非常显著<sup>[29]</sup>。双线性网络采用两个并列的卷积神经网络作为特征提取器以充分提取输入图像的深度卷积特征。在此基础上, 双线性网络中的双线性池化还能将两个卷积神经网络提取的深度卷积特征相互关联, 增强图像特征的表达能力, 进而获得更准确的识别结果<sup>[30]</sup>。聚焦对图像识别起关键作用的图像特征, CBAM (convolutional block attention module, CBAM) 注意力机制的作用显著<sup>[31]</sup>。在油茶<sup>[32]</sup>和玉米<sup>[33]</sup>等的品种识别和农作物病虫害<sup>[34]</sup>识别研究中, CBAM 注意力机制均被使用。

为准确识别自然条件下的油橄榄品种, 本研究将双线性网络和注意力机制融合。以在图像分类中表现良好的 EfficientNet-B0 为特征提取网络<sup>[35]</sup>, 结合 CBAM 注意力的关键特征聚焦能力和双线性网络的特征提取能力, 搭建了用于自然条件下油橄榄品种识别的双线性注意力 EfficientNet 模型。采用 Grad-CAM (gradient-weighted class activation mapping, Grad-CAM) 热力图可视化技术解释双线性注意力 EfficientNet 模型油橄榄品种识别结果。本研究的主要贡献是提出了一种用于自然条件下油橄榄品种识别的双线性注意力 EfficientNet 模型, 并通过试验验证了该模型的有效性和先进性。本研究可为自然条件下的作物品种识别提供参考。

## 1 数据采集

### 1.1 研究区域与图像采集

甘肃省陇南市的油橄榄种植基地 (100°15'34"E, 27°18'24"N, 海拔 1 725 m) 为本研究的试验区域, 基地占地 690.2 hm<sup>2</sup>。以高产油橄榄品种佛奥、莱星、皮削利和鄂植 8 号为研究对象, 于 2020 年 8 月至 12 月份使用小米 5 手机每隔五天采集一次自然条件下的油橄榄图像。为保证图像中油橄榄清晰且完整, 手机与油橄榄间的距离范围在 10~30 cm。考虑到光照对油橄榄成像的影响, 分别于每天的 8:00—9:00、11:00—12:00 和 15:00—16:00 采集自然条件下的油橄榄图像。最终, 共采集 4 个主栽油橄榄品种的 2 400 幅图像 (每个品种 600 幅, 3 456×4 608 像素), 以 JPEG 格式保存。不同品种油橄榄的示例图像如图 1 所示。

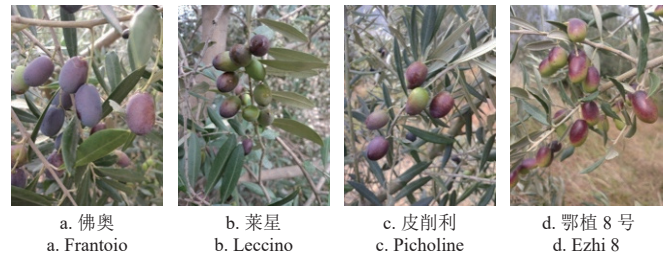


图 1 不同品种油橄榄示例图像

Fig.1 Examples of images of different olive cultivars

由图 1 可知, 不同品种油橄榄在果实和叶片上存在细微的差异。为充分描述不同品种油橄榄间的细微差异, 本研究对佛奥、莱星、皮削利和鄂植 8 号的果实和叶片特征进行总结, 如表 1 所示。

表 1 不同品种油橄榄性状特征

Table 1 Characteristics of olive in different cultivars

品种 Cultivar	果实特征 Fruit characteristics		叶片特征 Leaf characteristics	
	果实大小 Fruit size	果实形状 Fruit shape	叶片形状 Leaf shape	叶片颜色 Leaf color
佛奥 Frantoio	中等	椭圆形	长椭圆披针形	深绿且光亮
莱星 Leccino	中等	椭圆形	长椭圆形	深绿色
皮削利 Picholine	较大	卵圆形	长椭圆披针形	深灰绿色
鄂植 8 号 Ezhi 8	较大	长椭圆形	菱形	深绿色

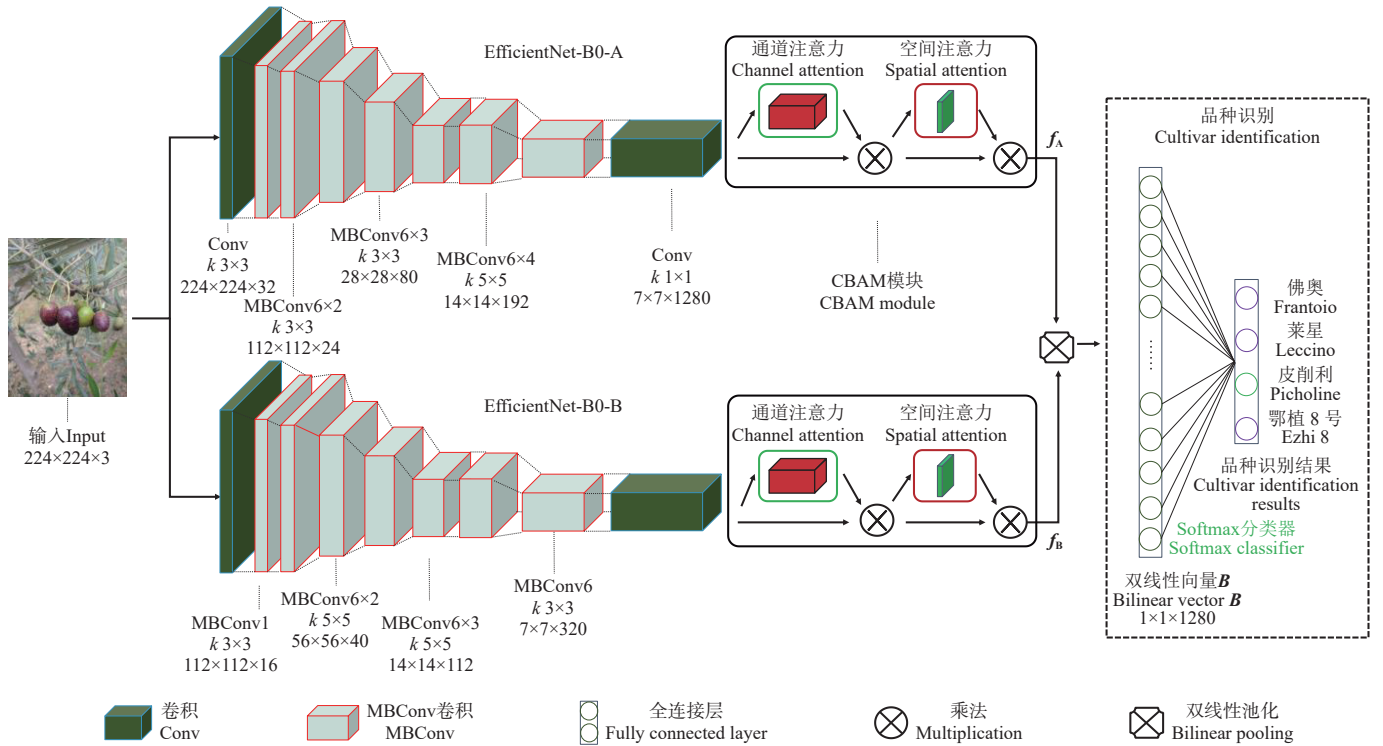
### 1.2 图像预处理与数据集构建

以采集的 2 400 幅自然条件下的油橄榄图像为试验数据, 搭建油橄榄品种识别数据集。在此基础上, 按照 7:3 的比例将 2 400 幅油橄榄图像划分为训练集 (1 680 幅, 每个品种 420 幅) 和测试集 (720 幅, 每个品种 180 幅)。本研究采用了亮度调整、图像翻转、图像旋转以及添加椒盐噪声 4 种方式对采集的 1 680 幅油橄榄果实图像进行扩充, 以提高油橄榄品种识别模型对光照、拍摄角度以及环境噪声的鲁棒性。经扩充后, 训练集的 1 680 幅油橄榄果实图像被扩充至 6 000 幅 (每个品种 1 500 幅)。

## 2 研究方法

### 2.1 双线性注意力 EfficientNet 模型

自然条件下不同品种的油橄榄仅在果实的颜色和形状以及叶子形状等方面存在细微的差异, 因此, 油橄榄的品种识别属于典型的细粒度图像分类问题<sup>[36]</sup>。对于油橄榄品种识别这种细粒度图像分类问题, 采用传统的图像识别网络往往难以取得满意的结果<sup>[32]</sup>。为解决自然条件下的油橄榄品种识别问题, 本文将双线性网络与注意力机制融合, 设计了用于油橄榄品种识别的双线性注意力 EfficientNet 模型。双线性注意力 EfficientNet 模型主要包括 2 个 EfficientNet-B0 网络、2 个 CBAM 注意力模块和 1 个双线性池化, 其基本网络结构如图 2 所示。



注: Conv 表示卷积, MBConv 表示移动翻转瓶颈卷积,  $f_A$  表示 EfficientNet-B0-A 提取的油橄榄图像特征,  $f_B$  表示 EfficientNet-B0-B 提取的油橄榄图像特征,  $k \ 1 \times 1$ 、 $k \ 3 \times 3$  和  $k \ 5 \times 5$  表示卷积核的大小分别为  $1 \times 1$ 、 $3 \times 3$  和  $5 \times 5$ 。  
 Note: Conv denotes convolution, MBConv denotes mobile inverted bottleneck convolution,  $f_A$  denotes olive image feature extracted by EfficientNet-B0-A,  $f_B$  denotes olive image feature extracted by EfficientNet-B0-B,  $k \ 1 \times 1$ ,  $k \ 3 \times 3$  and  $k \ 5 \times 5$  denotes convolution kernels of sizes  $1 \times 1$ ,  $3 \times 3$  and  $5 \times 5$ , respectively.

图 2 双线性注意力 EfficientNet 模型网络结构  
 Fig.2 Network structure of bilinear attention EfficientNet

2.1.1 特征提取网络

充分提取油橄榄图像中的特征信息是实现油橄榄品种准确识别的关键。本研究采用 2 个 EfficientNet-B0 模型作为油橄榄品种识别模型的特征提取网络,充分提取图像中油橄榄的特征信息。因为,先前的研究已经表明,与 VGG16、ResNet 和 GoogLeNet 等模型相比, EfficientNet-B0 模型具有参数量小、识别精度高的特点<sup>[37]</sup>。这主要得益于 EfficientNet-B0 模型采用的模型复合缩放策略,即通过复合系数 $\varphi$ 同时调整模型的网络深度、宽度以及图像大小。EfficientNet-B0 模型的复合缩放计算如式 (1)。

$$d = \alpha^\varphi, w = \beta^\varphi, r = \gamma^\varphi \quad \text{s.t.} \quad \begin{cases} \alpha \cdot \beta^2 \cdot \gamma^2 \approx 2 \\ \alpha \geq 1, \beta \geq 1, \gamma \geq 1 \end{cases} \quad (1)$$

式中  $d$ 、 $w$  和  $r$  分别表示模型在深度、宽度和图像大小上的放缩系数,  $\alpha$ 、 $\beta$  和  $\gamma$  为通过模型搜索确定的常数。

EfficientNet-B0 模型主要包括卷积层、移动翻转瓶颈卷积 (mobile inverted bottleneck convolution, MBConv)<sup>[38]</sup>、池化层和全连接层,基本网络结构如表 2 所示。

由表 2 可知, EfficientNet-B0 模型的核心是移动翻转瓶颈卷积模块,其结构示意图如图 3 所示。对于输入的特征图,移动翻转瓶颈卷积首先通过逐点卷积改变特征图的维度。其次,将降维后的特征图输入深度卷积以减少参数量。然后,再通过压缩和激励 (squeeze-and-excitation, SE) 模块执行通道维度上的注意力操作,以更好的捕获对油橄榄品种识别起关键作用的图像特征<sup>[39]</sup>。接

着,通过逐点卷积将输入特征图恢复到初始维度。最后,执行失活连接和跳跃连接操作。

表 2 EfficientNet-B0 模型网络结构  
 Table 2 Network structure of EfficientNet-B0 model

阶段 Stage	类型 Type	分辨率 Resolution	通道 Channels	层 Layers
1	Conv, $k \ 3 \times 3$	224×224	32	1
2	MBConv 1, $k \ 3 \times 3$	112×112	16	1
3	MBConv 6, $k \ 3 \times 3$	112×112	24	2
4	MBConv 6, $k \ 5 \times 5$	56×56	40	2
5	MBConv 6, $k \ 3 \times 3$	28×28	80	3
6	MBConv 6, $k \ 5 \times 5$	14×14	112	3
7	MBConv 6, $k \ 5 \times 5$	14×14	192	4
8	MBConv 6, $k \ 3 \times 3$	7×7	320	1
9	Conv, $k \ 1 \times 1$ , Pooling, FC	7×7	1280	1

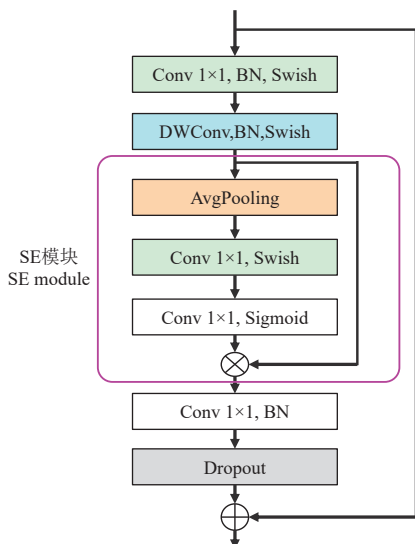
注: Pooling 表示池化层, FC 表示全连接层。  
 Note: Pooling denotes pooling layer, FC denotes fully connected layer.

2.1.2 CBAM 注意力机制

自然条件下采集的油橄榄图像中包含杂草、天空和地面等复杂背景,这些复杂背景会对品种识别产生负向影响。为使双线性注意力 EfficientNet 模型识别油橄榄品种时将更多的注意力聚焦到对品种识别起关键作用的特征上,本研究在 EfficientNet-B0 后引入了 CBAM 注意力模块<sup>[40]</sup>。CBAM 注意力模块结构示意图如图 4 所示。

如图 4 所示,对于输入的油橄榄图像特征  $F$ ,CBAM 注意力模块会先后进行通道注意力和空间注意力的提取。其中,通道注意力首先通过平均池化和最大池化对输入特征图  $F$  进行处理,生成两个不同的通道信息

描述  $F_{AvgPool}^c$  和  $F_{MaxPool}^c$ 。接着, 将  $F_{AvgPool}^c$  和  $F_{MaxPool}^c$  输入共享多层感知机处理, 获得两个特征描述图。在此基础上, 通过逐元素求和以及 Sigmoid 函数处理两个特征描述图, 获得通道注意力权重  $M_c$ 。最后, 将通道注意力权重  $M_c$  和输入特征图  $F$  相乘, 获得精炼特征  $F'$ 。通道注意力权重  $M_c$  和精炼特征  $F'$  的计算过程如式 (1) ~ (3)。



注: BN 为批规范化, Swish 为 Swish 激活函数, DWConv 为深度卷积, AvgPooling 为平均池化, Sigmoid 为 Sigmoid 激活函数, Dropout 为随机失活层, SE 为压缩和激励。

Note: BN represents batch normalization, Swish represents Swish activation function, DWConv represents deep convolution, AvgPooling denotes average pooling, Sigmoid represents Sigmoid activation function, Dropout denotes random deactivation layer, and SE denotes squeeze and excitation.

图3 MBConv 结构示意图

Fig.3 Structure diagram of MBConv

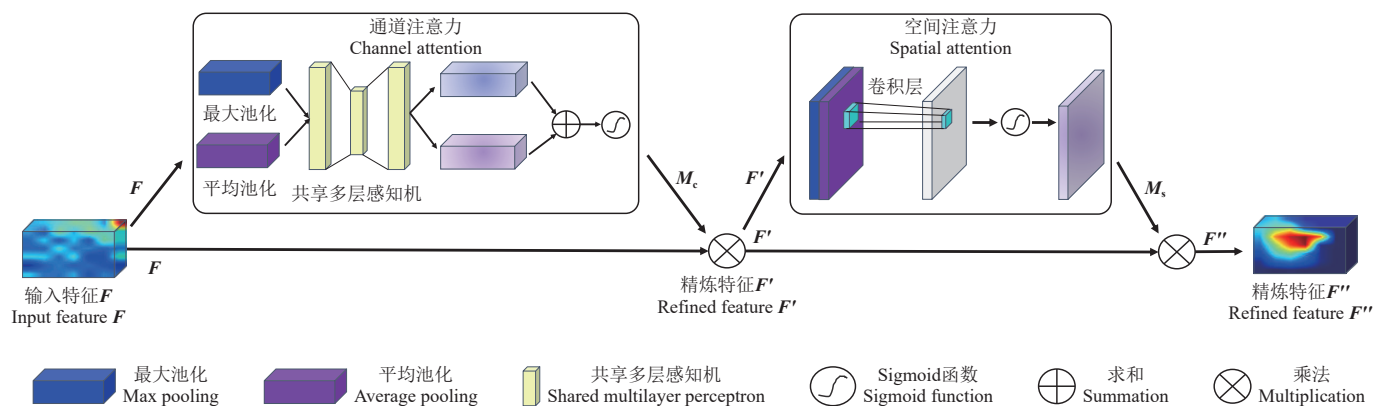


图4 CBAM 注意力模块结构示意图

Fig.4 Structure diagram of CBAM attention module

式中  $f_A(L)$  表示 EfficientNet-B0-A 提取的油橄榄图像中位置  $L$  处的特征,  $f_B(L)$  表示 EfficientNet-B0-B 提取的油橄榄图像中位置  $L$  处的特征,  $B(L, f_A, f_B)$  表示位置  $L$  处的双线性特征,  $x$  表示双线性特征,  $y$  为一维双线性融合特征,  $z$  表示归一化后的双线性融合特征。

最后, 将归一化后的融合双线性特征向量  $z$  输入 Softmax 分类器, 即可完成油橄榄图像的品种识别, 如式 (7)。

$$M_c = S_m(M_{lp}(P_{max}(F))) + S_m(M_{lp}(P_{avg}(F))) \quad (2)$$

$$F' = M_c F \quad (3)$$

式中  $S_m$  为 Sigmoid 激活函数,  $M_{lp}$  为多层感知机,  $P_{max}$  为最大池化,  $P_{avg}$  为平均池化。

在完成通道注意力的提取后, CBAM 注意力模块将对通道注意力模块输出的精炼特征  $F'$  执行空间注意力操作。首先, 对通道注意力模块输出的精炼特征执行平均池化和最大池化操作, 生成两个不同的空间信息描述  $F_{AvgPool}^s$  和  $F_{MaxPool}^s$  后按通道拼接。  $F_{AvgPool}^s$  和  $F_{MaxPool}^s$  分别表示空间中的平均池化特征和最大池化特征。然后, 通过大小为  $7 \times 7$  的卷积处理拼接后的特征。在此基础上, 通过 Sigmoid 函数处理得到空间注意力权重  $M_s$ 。最后, 将精炼特征  $F'$  与空间注意力权重  $M_s$  相乘, 获得最终融合通道和空间注意力的精炼特征  $F''$ 。空间注意力的权重系数  $M_s$  和精炼特征  $F''$  的计算过程如下:

$$M_s = S_m(f_{7 \times 7}[P_{max}(F'); P_{avg}(F')]) \quad (4)$$

$$F'' = M_s F' \quad (5)$$

式中  $f_{7 \times 7}$  为大小为  $7 \times 7$  的卷积。

### 2.1.3 特征融合和品种识别

为充分融合 EfficientNet-B0-A 和 EfficientNet-B0-B 提取的油橄榄图像特征  $f_A$  和  $f_B$ , 采用双线性池化对其进行处理, 如式 (6)。

$$\begin{cases} B(L, f_A, f_B) = f_A^T(L) \cdot f_B(L) \\ x = \sum B(L, f_A, f_B) \\ y = \text{sign}(x) \sqrt{|x|} \\ z = y / \|y\|_2 \end{cases} \quad (6)$$

$$\sigma(z_j) = \frac{e^{z_j}}{\sum_{m=1}^M e^{z_m}} \quad \text{s.t. } j \in [1, M] \quad (7)$$

式中  $j$  表示输出节点的编号,  $M$  表示油橄榄品种的总数,  $m$  表示第  $m$  个油橄榄品种。

## 2.2 评价指标

对于自然条件下的油橄榄品种识别问题, 模型精度和速度是需要重点关注的指标。为此, 本研究以精确率 (precision,  $P$ )、召回率 (recall,  $R$ )、F1 分数 (F1 score,

$F_1$ )、推理时间 (inference time,  $I_t$ ) 和总体准确率 (overall accuracy,  $O_a$ ) 为评价指标, 对油橄榄品种识别模型进行定量评价<sup>[41]</sup>。精确率、召回率、 $F1$  分数和总体准确率的具体计算过程如下:

$$P = \frac{T_P}{T_P + F_P} \quad (8)$$

$$R = \frac{T_P}{T_P + F_N} \quad (9)$$

$$F1 = \frac{2 \cdot P \cdot R}{P + R} \quad (10)$$

$$O_a = \frac{1}{N} \sum_{m=1}^M x_m \quad (11)$$

式中  $T_P$  表示测试图像中油橄榄品种被正确识别为对应品种的图像数量;  $F_P$  表示测试图像中其他油橄榄品种被错误识别为该品种的图像数量,  $F_N$  表示测试图像中该油橄榄品种被错误识别为其他品种的图像数量。  $N$  为测试集中不同品种油橄榄的图像总数,  $x_m$  表示第  $m$  个品种的油橄榄测试图像中被正确识别的图像数量。

### 2.3 试验设置

油橄榄品种识别试验在深度学习服务器上进行, 该服务器搭载 Windows 操作系统, 硬件环境为 Intel Core i7-8700K @ 3.7 GHz CPU, GeForce RTX 2070 SUPER 8G GPU, 16G 运行内存, 500G 固态硬盘和 1T 机械硬盘。该服务器的软件环境为 Python 3.6, TensorFlow 1.12, Keras 2.2.4。油橄榄品种识别算法的模型训练过程使用 Adam 优化器更新参数, 设置初始学习率为 0.001, 权值衰减设置为 0.0005, 批处理量为 16, 丢弃比为 0.5。基于油橄榄品种识别算法试验的硬件条件, 在综合考虑油橄榄品种识别的精度和速度后, 设置双线性注意力 EfficientNet 模型的输入图像大小为 224×224。

## 3 结果与分析

### 3.1 品种识别结果

以测试集中 4 个油橄榄品种的 720 幅图像为试验数据, 对提出的油橄榄品种识别模型进行测试。油橄榄品种识别结果以混淆矩阵的形式呈现, 如图 5 所示。

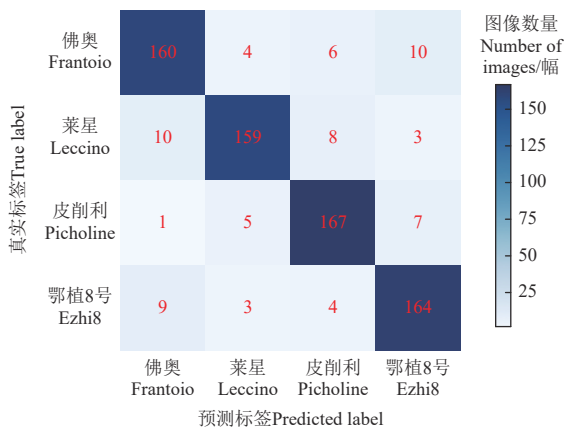


图 5 品种识别结果的混淆矩阵

Fig.5 Confusion matrix of cultivar identification results

由图 5 可知, 所提模型对测试集中的大部分油橄榄图像实现了准确的品种识别。从整体看, 对于测试集包

含的 4 个油橄榄品种的 720 幅图像, 所提模型准确识别了其中的 650 幅。从单个油橄榄品种的识别来看, 以测试集中的 180 幅品种为皮削利的油橄榄图像为例, 该模型准确识别了其中 167 幅, 仅有 13 幅品种为皮削利的油橄榄图像被错误识别。对于测试集中 180 幅品种为莱星的油橄榄果实图像, 该模型准确识别了其中的 159 幅。结果表明, 所提模型对单个油橄榄品种的识别, 也取得了较好的效果。

以精确率、召回率、 $F1$  分数、推理时间和总体准确率对双线性注意力 EfficientNet 模型油橄榄品种识别进行定量评价, 结果如表 3 所示。

表 3 油橄榄品种识别定量评价结果  
Table 3 Quantitative evaluation results of olive cultivar identification

品种 Cultivar	精确率 Precision P/%	召回率 Recall R/%	F1 分数 F1 score /%	推理时间 Inference time $I_t$ /ms	总体准确率 Overall accuracy $O_a$ /%
佛奥 Frantoio	88.89	88.89	88.89		
莱星 Leccino	92.98	88.33	90.60	9.15	90.28
皮削利 Picholine	90.27	92.78	91.51		
鄂植 8 号 Ezhi 8	89.13	91.11	90.11		

由表 3 可知, 所提模型识别 4 个油橄榄品种的总体准确率达到 90.28%, 推理时间仅为 9.15 ms。在测试试验的 4 个油橄榄品种中, 所提模型对皮削利的识别效果最好, 其精确率、召回率和  $F1$  分数分别达到了 90.27%、92.78% 和 91.51%。经测试, 所提模型对测试集中莱星识别的精确率、召回率和  $F1$  分数分别为 92.98%、88.33% 和 90.60%; 鄂植 8 号的识别精确率、召回率和  $F1$  分数分别为 89.13%、91.11% 和 90.11%。与皮削利、莱星和鄂植 8 号相比, 所提模型对佛奥的品种识别虽然效果略差, 但其精确率、召回率和  $F1$  分数均达到了 88.89%。

### 3.2 消融试验结果

设计消融试验验证双线性网络和 CBAM 注意力对油橄榄品种识别的有效性, 结果如表 4 所示。可以发现, 在 EfficientNet 模型中引入 CBAM 注意力和搭建双线性网络提高了油橄榄品种识别效果。

表 4 双线性注意力 EfficientNet 模型的消融试验结果  
Table 4 Ablation experiment results of bilinear attention EfficientNet model

模型 Model	品种 Cultivar	P/%	R/%	F1/%	$I_t$ /ms	$O_a$ /%
EfficientNet-B0	佛奥	72.16	70.56	71.35		
	莱星	76.19	80.00	78.05		
	皮削利	81.60	73.89	77.55	8.93	75.42
	鄂植 8 号	72.40	77.22	74.73		
EfficientNet-CBAM	佛奥	80.46	77.78	79.10		
	莱星	79.89	81.67	80.77		
	皮削利	84.52	78.89	81.61	9.06	80.42
	鄂植 8 号	77.32	83.33	80.21		
双线性 EfficientNet Bilinear EfficientNet	佛奥	86.71	83.33	84.99		
	莱星	85.56	85.56	85.56		
	皮削利	89.67	91.65	90.66	9.14	86.39
	鄂植 8 号	83.61	85.00	84.30		
双线性注意力 EfficientNet Bilinear attention EfficientNet	佛奥	88.89	88.89	88.89		
	莱星	92.98	88.33	90.60		
	皮削利	90.27	92.78	91.51	9.15	90.28
	鄂植 8 号	89.13	91.11	90.11		

从整体看,在 EfficientNet 模型中引入 CBAM 注意力搭建 EfficientNet-CBAM 模型后,油橄榄品种识别的总体准确率提高 5 个百分点,推理时间增加了 0.13 ms。以 EfficientNet 模型为基础搭建双线性 EfficientNet 模型后,总体准确率提高 10.97 个百分点,推理时间增加了 0.21 ms。当向 EfficientNet 模型中引入 CBAM 注意力并以此为基础搭建双线性网络后,总体准确率提升 14.86 个百分点,推理时间增加了 0.22 ms。从单个品种看,不论是向 EfficientNet 模型中引入 CBAM 注意力还是搭建

双线性网络均起到了提高单个油橄榄品种识别效果的作用。并且,当向 EfficientNet 模型中同时引入 CBAM 注意力和搭建双线性网络后,油橄榄的品种识别效果提升最显著。

### 3.3 不同模型的对比试验结果

对于油橄榄品种识别这种细粒度图像分类问题,主流的方法是搭建双线性网络或者结合注意力机制。不同方法的品种识别结果以混淆矩阵的形式呈现,如图 6 所示。

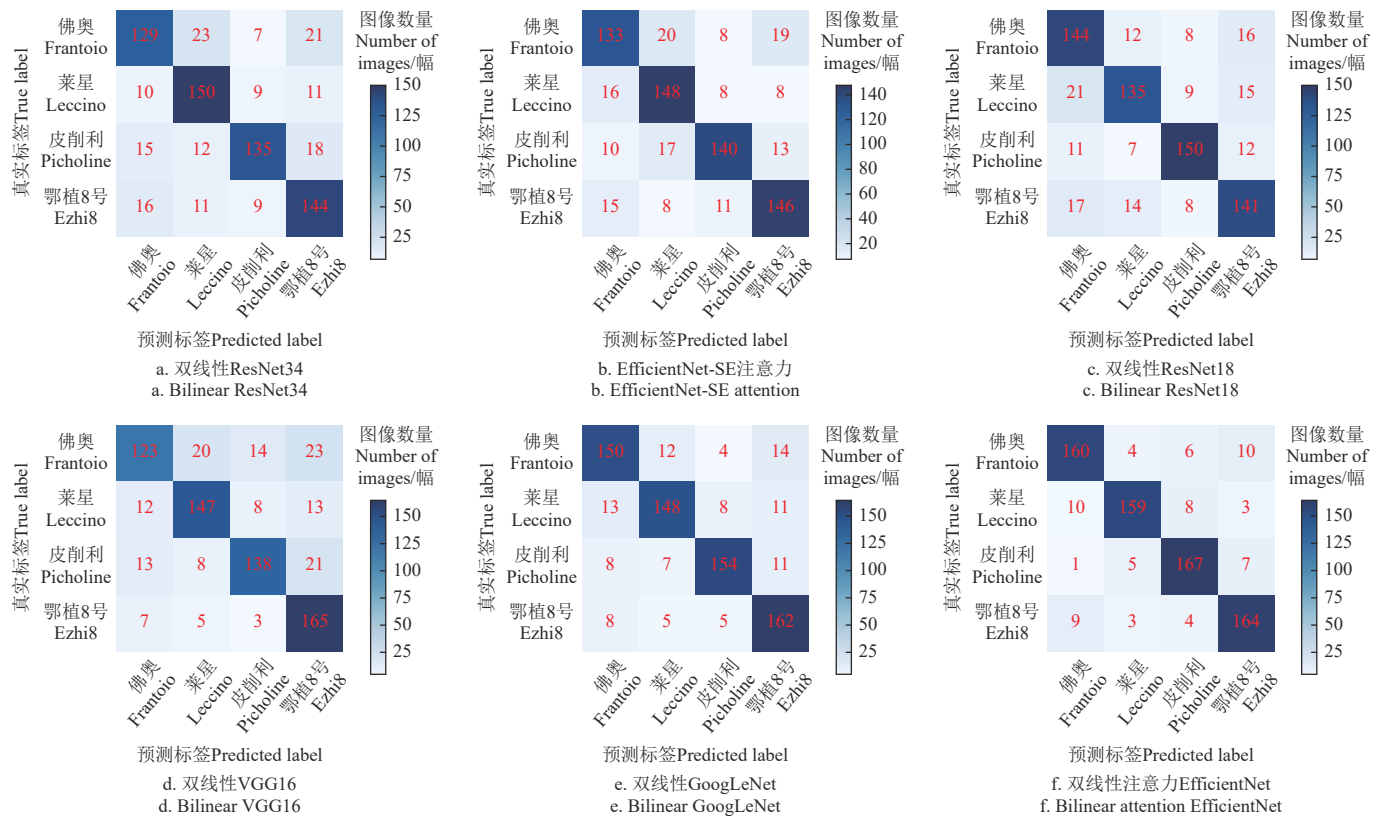


图 6 不同方法识别油橄榄品种的混淆矩阵

Fig.6 Confusion matrix of different methods to identify olive cultivars

双线性 ResNet34 模型、EfficientNet-SE 注意力模型、双线性 ResNet18 模型和双线性 VGG16 模型的识别效果较差(图 6a, 6b、6c 和 6d)。从 4 个油橄榄品种识别的整体结果分析,对于测试集的 720 幅油橄榄图像,双线性 ResNet34 模型、EfficientNet-SE 注意力模型、双线性 ResNet18 模型和双线性 VGG16 模型仅分别准确识别了其中的 558、567、570 和 573 幅。以佛奥为例从单个油橄榄品种识别角度分析,对于测试集中佛奥的 180 幅油橄榄图像,双线性 ResNet34 模型、EfficientNet-SE 注意力模型、双线性 ResNet18 模型和双线性 VGG16 模型仅准确识别了其中的 129、133、144 和 123 幅。可以发现,不论是对于单个油橄榄品种的识别还是 4 个油橄榄品种的识别,双线性 ResNet34 模型、EfficientNet-SE 注意力模型、双线性 ResNet18 模型和双线性 VGG16 模型均出现了较多的识别错误。

与双线性 ResNet34 模型、EfficientNet-SE 注意力模型、双线性 ResNet18 模型和双线性 VGG16 模型相比,

双线性 GoogLeNet 模型的油橄榄品种识别效果更好(图 6e)。测试集中 4 个油橄榄品种的 720 幅图像,双线性 GoogLeNet 模型准确识别了其中的 614 幅。测试集中品种为皮削利的 180 幅图像,双线性 GoogLeNet 模型准确识别了其中的 154 幅。

所提模型的油橄榄品种识别效果最好(图 6f)。对于测试集 4 个油橄榄品种的 720 幅图像,所提模型准确识别了其中的 650 幅,优于对比试验中的其他方法。并且,对于单个油橄榄品种的识别,所提模型也取得了不错的效果。测试集中佛奥、莱星、皮削利和鄂植 8 号各 180 幅,所提模型分别准确识别了其中的 160、159、167 和 164 幅。

由表 5 可知,所提模型对测试集中 4 个油橄榄品种识别的总体准确率达到 90.28%,分别比双线性 ResNet34、EfficientNet-SE 注意力、双线性 ResNet18、双线性 VGG16、双线性 GoogLeNet 高 12.78、11.53、11.11、

10.70 和 5.00 个百分点。从推理时间的角度分析，所提模型的推理时间为 9.15 ms，与 EfficientNet-SE 注意力和双线性 ResNet18 接近且优于双线性 ResNet34、双线性 VGG16 和双线性 GoogLeNet。

表 5 不同品种识别方法的定量评价结果

Table 5 Quantitative evaluation results of different cultivar identification methods

模型 Model	品种 Cultivar	P/%	R/%	F1/%	I/ms	O <sub>d</sub> /%
双线性 ResNet34 Bilinear ResNet34	佛奥	75.88	71.67	73.71	13.16	77.50
	莱星	76.53	83.33	79.79		
	皮削利	84.38	75.00	79.41		
	鄂植 8 号	74.23	80.00	77.01		
EfficientNet-SE 注意力 EfficientNet-SE attention	佛奥	76.44	73.89	75.14	9.04	78.75
	莱星	76.68	82.22	79.36		
	皮削利	83.83	77.78	80.69		
	鄂植 8 号	78.49	81.11	79.78		
双线性 ResNet18 Bilinear ResNet18	佛奥	74.61	80.00	77.21	8.72	79.17
	莱星	80.36	75.00	77.59		
	皮削利	85.71	83.33	84.51		
	鄂植 8 号	76.63	78.33	77.47		
双线性 VGG16 Bilinear VGG16	佛奥	79.35	68.33	73.43	9.32	79.58
	莱星	81.67	81.67	81.67		
	皮削利	84.66	76.67	80.47		
	鄂植 8 号	74.32	91.67	82.09		
双线性 GoogLeNet Bilinear GoogLeNet	佛奥	83.80	83.33	83.57	21.03	85.28
	莱星	86.05	82.22	84.09		
	皮削利	90.06	85.56	87.75		
	鄂植 8 号	81.82	90.00	85.71		
双线性注意力 EfficientNet Bilinear attention EfficientNet	佛奥	88.89	88.89	88.89	9.15	90.28
	莱星	92.98	88.33	90.60		
	皮削利	90.27	92.78	91.51		
	鄂植 8 号	89.13	91.11	90.11		

以皮削利为例，所提模型精确率、召回率和  $F1$  分数分别达到 90.27%、92.78% 和 91.51%。与双线性 ResNet34、EfficientNet-SE 注意力、双线性 ResNet18、双线性 VGG16、双线性 GoogLeNet 相比，所提模型的精确率、召回率和  $F1$  分数均更优。

### 3.4 品种识别热力图可视化

不同品种油橄榄往往在果实颜色、形状以及叶子形状等表型信息上存在差异。为验证所提模型在识别油橄榄品种时是否也关注了果实和叶子区域，本研究采用 Grad-CAM 技术对双线性注意力 EfficientNet 模型的油橄榄品种识别进行可视化解释，以从视觉上评价所提出的油橄榄品种识别方法<sup>[42]</sup>。双线性注意力 EfficientNet 模型识别 4 个油橄榄品种的热力图如图 7 所示。

由图 7 可知，所提模型在识别油橄榄品种时将注意力主要聚焦到了图像中油橄榄的果实区域以及部分叶子区域。油橄榄品种识别的热力图可视化试验证明了所提模型在识别油橄榄品种时所依据的图像信息是可信的。

本研究提出的双线性注意力 EfficientNet 模型在对测试集中油橄榄进行品种识别时，也出现了少数错误识别的情况。为解释所提模型错误识别的原因，本研究对发生错误识别的油橄榄图像进行可视化解释，如图 8 所示。

在对图 8 中的油橄榄图像进行品种识别时，所提模型并未将注意力聚焦到果实和叶子区域，而是聚焦到了对品种识别有干扰的背景区域。由此可以推断，不能有

效聚焦到油橄榄图像中对品种识别起关键作用的果实和叶子区域是发生错误识别的主要原因。

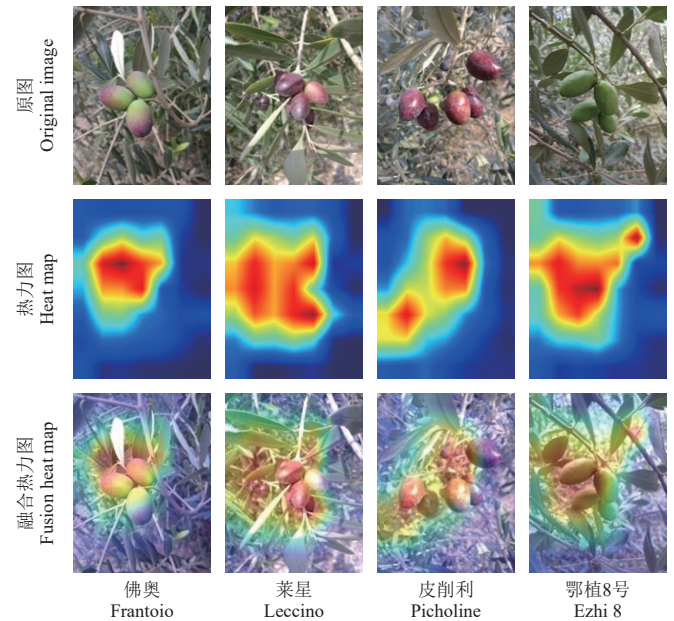


图 7 油橄榄品种识别的可视化解释

Fig.7 Visual explanation of olive cultivar identification

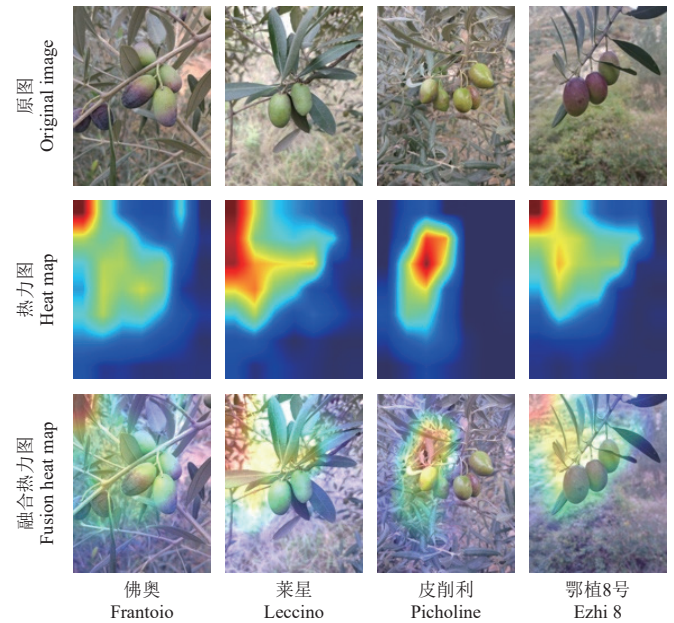


图 8 发生品种误识别的油橄榄图像的可视化解释

Fig.8 Visual explanation of olive images with cultivar misidentification

## 4 结论

为实现自然条件下油橄榄品种的识别，本研究以甘肃主栽油橄榄品种佛奥、莱星、皮削利和鄂植 8 号为研究对象，融合深度学习中的双线性网络与注意力机制，提出基于双线性注意力 EfficientNet 模型的油橄榄品种识别方法。试验评价双线性注意力 EfficientNet 模型的油橄榄品种识别结果，得到如下结论：

1) 双线性注意力 EfficientNet 模型可以快速准确识

别油橄榄品种, 总体准确率达到 90.28%, 推理时间为 9.15 ms。并且, 双线性注意力 EfficientNet 模型的热力图可视化结果显示, 该模型在识别油橄榄品种时将注意力聚焦到了油橄榄图像的果实和叶子区域。油橄榄品种识别试验和消融试验结果充分验证了双线性注意力 EfficientNet 模型识别油橄榄品种的有效性。

2) 与主流品种识别方法相比, 本文提出的双线性注意力 EfficientNet 模型在油橄榄品种识别精度上具有明显优势。双线性注意力 EfficientNet 模型对测试集油橄榄品种识别的总体准确率达到 90.28%, 分别比双线性 ResNet34、EfficientNet-SE 注意力、双线性 ResNet18、双线性 VGG16、双线性 GoogLeNet 高 12.78、11.53、11.11、10.70、5.00 个百分点。并且, 对于单个油橄榄品种的识别, 双线性注意力 EfficientNet 模型也优于对比试验的其他方法。

#### [参 考 文 献]

- [1] ZHANG Y, ZHANG C, XU C, et al. Effect of geographical location and soil fertility on main phenolic compounds and fatty acids compositions of virgin olive oil from Leccino cultivar in China[J]. *Food Research International*, 2022, 157: 111207.
- [2] ZHU Z, LI X, ZHANG Y, et al. Profiling of phenolic compounds in domestic and imported extra virgin olive oils in China by high performance liquid chromatography-electrochemical detection[J]. *LWT-Food Science and Technology*, 2023, 174: 114424.
- [3] LECHHAB T, LECHHAB W, CACCIOLA F, et al. Sets of internal and external factors influencing olive oil (*Olea europaea L.*) composition: A review[J]. *European Food Research and Technology*, 2022, 248: 1069-1088.
- [4] 朱万泽, 范建容, 彭建国, 等. 四川省油橄榄引种品种果实含油率及其脂肪酸分析[J]. *林业科学*, 2010, 46(8): 91-100.  
ZHU Wanze, FAN Jianrong, PENG Jianguo, et al. Analysis of the oil content and its fatty acid composition of fruits for introduced olive cultivars in Sichuan province[J]. *Scientia Silvae Sinicae*, 2010, 46(8): 91-100. (in Chinese with English abstract)
- [5] 王玉亮, 刘贤喜, 苏庆堂, 等. 多对象特征提取和优化神经网络的玉米种子品种识别[J]. *农业工程学报*, 2010, 26(6): 199-204.  
WANG Yuliang, LIU Xianxi, SU Qingtang, et al. Maize seeds varieties identification based on multi-object feature extraction and optimized neural network[J]. *Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE)*, 2010, 26(6): 199-204. (in Chinese with English abstract)
- [6] WANG B, GAO Y, YUAN X, et al. From species to cultivar: Soybean cultivar recognition using joint leaf image patterns by multiscale sliding chord matching[J]. *Biosystems Engineering*, 2020, 194: 99-111.
- [7] TRIPATHI M K and MAKTEDAR D. A role of computer vision in fruits and vegetables among various horticulture products of agriculture fields: A survey[J]. *Information Processing in Agriculture*, 2020, 7(2): 183-203.
- [8] ABBASPOUR-GILANDEH Y, MOLAEI A, SABZI S, et al. A combined method of image processing and artificial neural network for the identification of 13 Iranian rice cultivars[J]. *Agronomy*, 2020, 10: 117.
- [9] 李景彬, 陈兵旗, 邵鲁浩, 等. 基于 BP 神经网络的脱绒棉种品种识别[J]. *农业工程学报*, 2012, 28(26): 265-269.  
LI Jingbin, CHEN Bingqi, SHAO Luhao, et al. Variety identification of delinted cottonseeds based on BP neural network[J]. *Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE)*, 2012, 28(26): 265-269. (in Chinese with English abstract)
- [10] 李浩光, 李卫军, 覃鸿, 等. 基于栈式自编码神经网络的包衣单籽粒玉米品种识别[J]. *农业机械学报*, 2017, 48(s1): 422-428.  
LI Haoguang, LI Weijun, QIN Hong, et al. Varietal identification for single maize seed based on stacked auto encoder neural network[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2017, 48(s1): 422-428. (in Chinese with English abstract)
- [11] AVOLA G, DI GENNARO S F, CANTINI C, et al. Remotely sensed vegetation indices to discriminate field-grown olive cultivars[J]. *Remote Sensing*, 2019, 11: 1242.
- [12] BEYAZ A, ÖZKAYA M T, İÇEN D. Identification of some spanish olive cultivars using image processing techniques[J]. *Scientia Horticulturae*, 2017, 225: 286-292.
- [13] 黄敏, 夏超, 朱启兵, 等. 融合高光谱图像技术与 MS-3DCNN 的小麦种子品种识别模型[J]. *农业工程学报*, 2021, 37(18): 153-160.  
HUANG Min, XIA Chao, ZHU Qibing, et al. Recognizing wheat seed varieties using hyperspectral imaging technology combined with multi-scale 3D convolution neural network[J]. *Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE)*, 2021, 37(18): 153-160. (in Chinese with English abstract)
- [14] ZHANG J, DAI L, CHENG F. Corn seed variety classification based on hyperspectral reflectance imaging and deep convolutional neural network[J]. *Food Measure*, 2021, 15: 484-494.
- [15] LIU Q, WANG Z, LONG Y, et al. Variety classification of coated maize seeds based on Raman hyperspectral imaging[J]. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 2022, 270: 120772.
- [16] LI X, FAN X, ZHAO L, et al. Discrimination of pepper seed varieties by multispectral imaging combined with machine learning[J]. *Applied Engineering in Agriculture*, 2020, 36(5): 743-749.

- [17] MARTÍNEZ S S, GILA D M, BEYAZ A, et al. A computer vision approach based on endocarp features for the identification of olive cultivars[J]. *Computers and Electronics in Agriculture*, 2018, 154: 341-346.
- [18] FRANCZYK B, HERNES M, KOZIERKIEWICZ A, et al. Deep learning for grape variety recognition[J]. *Procedia Computer Science*, 2020, 176: 1211-1220.
- [19] HOANG V T, HOAI D P V, SURINWARANGKON T, et al. Comparative study of rice variety classification based on deep learning and hand-crafted features[J]. *ECTI Transactions on Computer and Information Technology*, 2020, 14(1): 1-10.
- [20] TANER A, ÖZTEKIN Y B, DURAN H. Performance analysis of deep learning CNN models for variety classification in hazelnut[J]. *Sustainability*, 2021, 13: 6527.
- [21] TAVAKOLI H, ALIREZAZADEH P, HEDAYATIPOUR A, et al. Leaf image-based classification of some common bean cultivars using discriminative convolutional neural networks[J]. *Computers and Electronics in Agriculture*, 2021, 181: 105935.
- [22] JUNG M, SONG J S, HONG S, et al. Deep learning algorithms correctly classify brassica rapa varieties using digital images[J]. *Frontiers in Plant Science*, 2021, 12: 738685.
- [23] 袁培森, 黎薇, 任守纲, 等. 基于卷积神经网络的菊花花型和品种识别[J]. *农业工程学报*, 2018, 34(5): 152-158.
- YUAN Peisen, LI Wei, REN Shougang, et al. Recognition for flower type and variety of chrysanthemum with convolutional neural network[J]. *Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE)*, 2018, 34(5): 152-158. (in Chinese with English abstract)
- [24] MIHO H, PAGNOTTA G, HITAJ D, et al. OliVaR: Improving olive variety recognition using deep neural networks[EB/OL]. (2023-03-01) [2023-03-27]<https://arxiv.org/abs/2303.00431>.
- [25] OSAKO Y, YAMANE H, LIN S Y, et al. Cultivar discrimination of litchi fruit images using deep learning[J]. *Scientia Horticulturae*, 2020, 269: 109360.
- [26] 孟志超, 贺磊盈, 杜小强, 等. 基于 Enhanced VGG16 的油茶品种分类[J]. *农业工程学报*, 2022, 38(10): 176-181.
- MENG Zhichao, HE Leiyong, DU Xiaoqiang, et al. Classification of *Camellia oleifera* based on Enhanced VGG16 network[J]. *Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE)*, 2022, 38(10): 176-181. (in Chinese with English abstract)
- [27] LIU Z, WANG J, TIAN Y, et al. Deep learning for image-based large-flowered chrysanthemum cultivar recognition[J]. *Plant Methods*, 2019, 15: 146.
- [28] MAMDOUH N and KHATTAB A. YOLO-Based deep learning framework for olive fruit fly detection and counting[J]. *IEEE Access*, 2021, 9: 84252-84262.
- [29] LIN T, ROYCHOWDHURY A, MAJI S. Bilinear CNN models for fine-grained visual recognition[C]//International Conference on Computer Vision (ICCV), IEEE, Santiago, Chile, 2015: 1449-1457.
- [30] GAO Y, BEIJBOM O, ZHANG N, et al. Compact bilinear pooling[C]//IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, Las Vegas, Nevada, USA, 2016: 317-326.
- [31] DU X, SI L, JIN X, et al. Classification of plug seedling quality by improved convolutional neural network with an attention mechanism[J]. *Frontiers in Plant Science*, 2022, 13: 967706.
- [32] ZHU X, ZHANG X, SUN Z, et al. Identification of oil tea (*Camellia oleifera* C. Abel) cultivars using EfficientNet-B4 CNN model with attention mechanism[J]. *Forests*, 2022, 13: 1.
- [33] MA R, WANG J, ZHAO W, et al. Identification of maize seed varieties using MobileNetV2 with improved attention mechanism CBAM[J]. *Agriculture*, 2023, 13(1): 11.
- [34] 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 with English abstract)
- [35] TAN M and LE Q. Efficientnet: Rethinking model scaling for convolutional neural networks[EB/OL]. (2019-05-28) [2023-03-27]<https://arxiv.org/abs/1905.11946>.
- [36] MENG X, YUAN Y, TENG G, et al. Deep learning for fine-grained classification of jujube fruit in the natural environment[J]. *Food Measure*, 2021, 15: 4150-4165.
- [37] ATILA Ü, UÇAR M, AKYOL K, et al. Plant leaf disease classification using EfficientNet deep learning model[J]. *Ecological Informatics*, 2021, 61: 101182.
- [38] SANDLER M, HOWARD A, ZHU M, et al. MobileNetV2: Inverted residuals and linear bottlenecks[C]//2018 IEEE/CVF Conference of Computer Vision and Pattern Recognition (CVPR), IEEE, Salt Lake City, UT, USA, 2018: 4510-4520.
- [39] HU J, SHEN L, ALBANE S, et al. Squeeze-and-Excitation networks[EB/OL]. (2019-05-16) [2023-03-27]<https://arxiv.org/abs/1709.01507>.
- [40] WOO S, PARK J, LEE J, et al. CBAM: Convolutional block attention module[C]//European Conference on Computer Vision (ECCV), Munich, Germany, Springer, 2018: 3-19.
- [41] 陈锋军, 朱学岩, 周文静, 等. 基于无人机航拍与改进 YOLOv3 模型的云杉计数的云杉计数[J]. *农业工程学报*, 2020, 36(22): 22-30.
- CHEN Fengjun, ZHU Xueyan, ZHOU Wenjing, et al. Spruce counting method based on improved YOLOv3 model in UVA images[J]. *Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE)*, 2020, 36(22): 22-30. (in Chinese with English abstract)
- [42] SELVARAJU R, COGSWELL M, DAS A, et al. Grad-CAM: Visual explanations from deep networks via gradient-based localization[J]. *International Journal of Computer Vision*, 2020, 128: 336-359.

## Identification of olive cultivars using bilinear networks and attention mechanisms

ZHU Xueyan<sup>1,2</sup>, CHEN Fengjun<sup>1,2,3\*</sup>, ZHENG Yili<sup>1,2,4</sup>, LI Zhiqiang<sup>1,4</sup>, ZHANG Xinwei<sup>1,5</sup>

(1. School of Technology, Beijing Forestry University, Beijing 100083, China; 2. National Key Laboratory-Forest Resource Efficient Production, Beijing 100083, China; 3. Beijing Laboratory of Urban and Rural Ecological Environment, Beijing 100083, China; 4. Key Laboratory of State Forestry Administration for Forestry Equipment and Automation, Beijing, 100083, China; 5. Research Center for Intelligent Forestry, Beijing 100083, China)

**Abstract:** The extensive range of olive cultivars available in the market exhibit minor differences in their phenotypic traits. Nonetheless, their quality attributes, particularly the oil content and fatty acid composition, significantly differ among distinct cultivars, resulting in the emergence of use inferior products as superior products in the market. The accurate and quick identification of olive cultivars holds significant importance in enhancing the production and quality of olives. As such, delving into the study of olive cultivar identification is crucial for the advancement of the olive industry. This study presents a novel approach to address the challenge of identifying olive cultivars in natural conditions. Specifically, a bilinear attentional EfficientNet model is proposed, which incorporates the bilinear network design concept and attention mechanism. The model is trained and evaluated using four commonly planted olive cultivars (i.e., Frantoio, Leccino, Picholine, and Ezhi 8) in Longnan, Gansu. The experimental results demonstrate the effectiveness of the proposed model in accurately and quickly identifying different olive cultivars. A bilinear network has been suggested to comprehensively extract feature information from olive images, for the limited phenotypic differences across different olive cultivars. In light of this, the selection of a feature extraction network has been made with consideration for both speed and accuracy, leading to the selection of the EfficientNet-B0 network. To tackle the challenge of identifying olive cultivars under natural conditions which are prone to intricate background interferences, a novel approach combining convolutional block attention module (CBAM) with bilinear network has been proposed. This approach facilitates the model in selectively focusing on the salient features responsible for cultivar identification during the feature extraction process of olive images. Upon conducting experiments, the bilinear attention EfficientNet model presented in this study has exhibited an overall accuracy of 90.28% and an inference time of 9.15 ms in identifying four distinct olive cultivars. These experiment results demonstrate that the proposed model better achieved better rapid and accurate identification of olive cultivars under natural conditions. The present study proposes a bilinear attention EfficientNet model for identifying olive cultivars and utilizes gradient-weighted class activation mapping (Grad-CAM) to analyse its performance. The results demonstrate that the proposed model exhibits a greater attention towards fruit regions, as well as some leaf regions, within olive images. These findings are in agreement with the expert knowledge and experience of human practitioners. The analysis of heat maps generated from misidentified olive images revealed that inadequate focus on the fruit and leaf regions, which are pivotal for successful identification of cultivars, was the primary contributing factor to misidentification. The outcomes of the ablation experiments indicated that the bilinear network and the CBAM exhibited a positive impact on the precision of olive variety recognition. To ascertain the efficacy of the method elucidated in this manuscript, a set of comparative experiments has been formulated. The primary objective of these experiments is to juxtapose the proposed bilinear attention EfficientNet model against the conventional cultivar identification models, including bilinear ResNet34, EfficientNet-SE attention, bilinear ResNet18, bilinear VGG16, and bilinear GoogLeNet. The experimental results obtained from the comparison analysis provide that the proposed bilinear attention EfficientNet model exhibits superior performance in terms of overall accuracy for the identification of olive cultivars. The bilinear attention EfficientNet model's accuracy for the identification of olive cultivars exceeded that of bilinear ResNet34, EfficientNet-SE attention, bilinear ResNet18, bilinear VGG16, and bilinear GoogLeNet models by 12.78, 11.53, 11.11, 10.70, and 5.00 percentage points, respectively. The present study establishes a foundation for resolving the challenge of accurately identifying olive cultivars in natural conditions. Moreover, it offers a valuable point of reference for the identification of cultivars of diverse crops.

**Keywords:** image processing; models; cultivar identification; olive; EfficientNet-B0; CBAM attention; Grad-CAM