DOI:10.13652/j.spjx.1003.5788.2023.60037 [文章编号] 1003-5788(2023)08-0129-07

基于改进 Faster R-CNN 模型的水果分类识别

Fruit identification using improved Faster R-CNN model

贾艳平[1] 桑妍丽[2] 李月茹[3]

JIAN Yanping[1] *SANG Yanli*[2] *LI Yueru*[3]

(1. 山西旅游职业学院,山西 太原 030031;2. 山西大学,山西 太原 030006;
3. 山西农业大学,山西 太原 030031)

(1. *Shanxi Vocational College of Tourism*, *Taiyuan*, *Shanxi* 030031, *China*; 2. *Shanxi University*, *Taiyuan*, *Shanxi* 030006, *China*; 3. *Shanxi Agricultural University*, *Taiyuan*, *Shanxi* 030031, *China*)

摘要:目的:提高水果类别识别准确率。**方法:**基于改进 Faster R-CNN 模型建立水果识别方法。使用正则化方法对高维参数进行权重衰减,以有效解决训练过程中可能出现的过拟合问题;在 Faster R-CNN 框架中添加两个损失函数:一个似然函数和一个正则化函数,以优化卷积层和池化层;以最小二乘法求解水果识别的目标函数;利用准确率、回召率、精度和 F1 分数对训练好的水果识别方法进行水果识别效果评估。**结果:**所提出的方法对水果识别的准确率、精度和回召率达到 99.69%,0.996 8,0.994 8;与其他 8 种水果识别方法相比,所提出方法对水果识别的准确率、精度和回召率至少提高了 0.91%,1.32%,0.51%。**结论:**该方法可准确识别水果种类。

关键词:水果;识别方法;Faster R-CNN;损失函数

Abstract: Objective: To improve the identification accuracy of fruit categories. **Methods:** An identification method for fruit categories was established based on an improved Faster regional convolutional neural network (Faster R-CNN) model. Firstly, the regularization method was used to attenuate the weight of some high-dimensional parameters to effectively solve the over-fitting problem that may occur in the training process. Then, two loss functions, a likelihood function and a regularization function, were added to the Faster R-CNN framework to optimize the convolution layer and the pooling layer. Additionally, the least square method was utilized to solve the objective function of fruit recognition. Finally, the accuracy, recall rate, precision and F1 score were used to evaluate the fruit identification effect of the trained fruit identification method. **Results:** The accuracy, precision and recall rates of the proposed method for fruit identification reached 99.69%, 0.996 8 and 0.994 8, respectively. Compared with the other 10 fruit identification methods, the accuracy, precision and recall rate of the proposed method were at least 0.91%, 1.32% and 0.51% higher. **Conclusion:** The method can realize the accurate recognition of different categories of fruits.

Keywords: fruit; identification method; Faster R-CNN; loss function

水果分拣机器人通常用于商店、市场、仓库等场所进行自动化的水果识别和分类工作。准确识别水果类别对水果分拣机器人至关重要[1],如果不能准确识别水果的类别,分拣机器人就无法提供准确的水果分拣服务。例如,在购物过程中无法正确地为顾客提供所需的水果,或者在仓库中无法正确地将水果分类和存储。通过准确地识别水果的类别,分拣机器人可以自动、高效地进行水果的分类、包装和储存,提高生产效率,减少错误和成本;而且还可以提升用户体验,确保顾客得到他们所需的水果,增加顾客的满意度[2-3]。

随着深度学习技术的迅速发展,基于快速区域 CNN (Faster regional CNN,Faster R-CNN)的目标识别算法,成为了水果识别任务完成过程中的重要工具。Faster R-CNN 具有高精度和高效率目标识别的特点,Faster R-CNN 可通过在单一网络中进行目标检测和分类,完成端到端的目标识别过程[4-5]。刘春妹等[6]研究了基于稀疏化 CNN 网络的水果图像识别系统,该系统主要包含卷积层、池化层以及全连接层,为降低模型的复杂度,文中对卷积层的权重进行了稀疏化设计;系统测试结果表明,在稀疏化与量化过程中,所构建的 CNN 模型的复杂度有所降低,但同时也牺牲了一部分水果图像识别准确性。

然而,传统的 Faster R-CNN 方法在水果识别过程中仍然面临挑战[7]。首先,水果在外观上存在较大的不同,其次在不同光照、拍摄角度、环境条件下,水果的外观(形状、颜色、大小等)也会发生变化,这对水果识别的准确性

基金项目:山西省教育科学"十三五"规划课题(编号:GH-1921025)

作者简介:贾艳平(1978—),女,山西旅游职业学院讲师,硕士。E-mail: bchenx28@foxmail.com

收稿日期:2023-02-03 **改回日期:**2023-07-04

和鲁棒性提出了较高的要求[8]。另外，传统 Faster R-CNN 方法对于遮挡严重的水果区域识别效果较差，这在实际应用中可能会导致漏检或者误检的问题[9]。因此，为更好地完成水果识别任务，需要对 Faster R-CNN 方法进行改进，以提高其识别性能和应对水果识别中的挑战。Zhong 等[10]提出基于改进 Faster R-CNN 模型的辣椒识别方法，该方法识别辣椒的整体准确率和平均准确率分别达到 75.79%和 87.30%，与 VGG16 模型相比，分别提高了 8.7%和 1.3%。

现有基于 Faster R-CNN 模型的水果识别方法虽取得了良好的水果识别效果，但识别准确率和泛化能力仍有待进一步提高。例如，Wan 等[11]提出的 Faster R-CNN 模型只对一类水果进行识别；Zhang 等[12]提出的 8 层和 13 层的 CNN 模型仅考虑了 18 类水果的识别。为进一步提高对水果类别的识别准确率，研究对 Faster R-CNN 进行改进，提出基于改进 Faster R-CNN 模型的水果识别新方法。

1　改进的 Faster R-CNN 模型

基于改进 Faster R-CNN 模型的水果识别框架如图 1 所示，该框架使用训练集中的水果图像作为输入，输出识别结果和图像中水果的边界框。特征提取过程使用了 CNN 模型 VGG-19。该模型包括 16 个卷积层、16 个 ReLu 层、4 个池化层和 3 个全连接层。

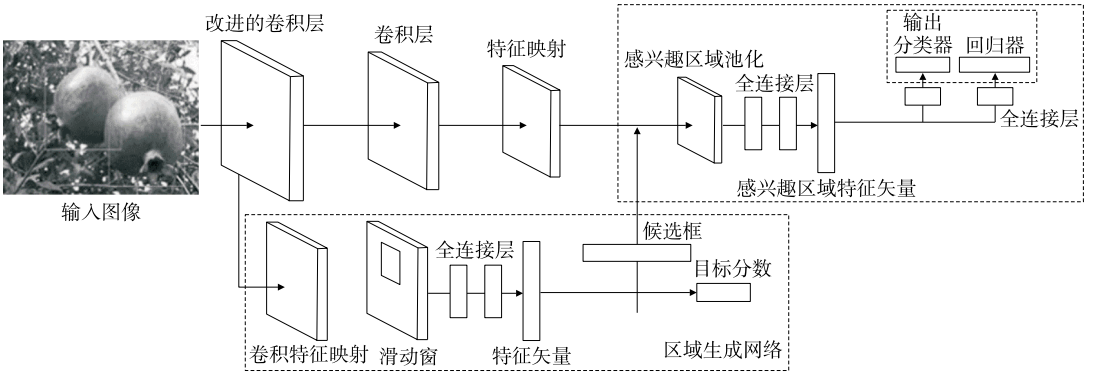


图 1　改进 Faster R-CNN 的架构

Figure 1　Improved architecture for Faster R-CNN

假设第 l 层为卷积层，该层的输出特征向量为：

$$x_j^l = f_l\left(\sum_{i\in M_j} x_i^{l-1} * k_{ij}^l + \delta_j^l\right)，\tag{1}$$

式中：

x_i^{l-1}——上一层（第 $l-1$ 层）的第 i 个输出；

k_{ij}^l——卷积核；

$*$——卷积乘法；

δ_j^l——第 l 层的偏置；

f_l——ReLu 函数。

假设第 M 层是池化层，该层的输出特征向量为：

$$x_j^m = f_2\left[\beta_j^m A(x_j^{m-1}) + \delta_j^m\right]，\tag{2}$$

式中：

β_j^m——连接系数；

x_j^{m-1}——第 M 层的输入；

A——第 M 层所有矩阵的总和；

δ_j^m——偏差；

f_2——softmax 函数。

训练过程中可能会出现过拟合问题，为了进一步提高水果识别的准确性，并平衡模型的复杂性和数据量，使用正则化方法对一些高维参数进行权重衰减[13]。为优化卷积层和池化层，添加了两个损失函数：一个似然函数[14]和一个正则化函数[15]。函数参数可以根据不同的水果图像拍摄角度自动调整，保证每个卷积层的大小和核参数都在合理的范围内。卷积层中的损失函数定义为：

$$L_{cv}(W,x^k) = \frac{1}{2m}\left[\sum_{i=1}^{m}\left(y_k - W * x_{ik}^2 + \sigma\sum_{j=1}^{n} w_j^2\right)\right]，\tag{3}$$

式中：

W——内核参数，$W=[\omega_1,\omega_2,\cdots,\omega_n]^{\mathrm{T}}$；

n——内核数量；

$x^{(k)}$——第 k 个输入，$x^{(k)}=[x_{1k},x_{2k},\cdots,x_{mk}]^{\mathrm{T}}$；

m——数据库维度；

y_k——第 k 个样本的实际标签；

σ——正则化惩罚因子。

通过添加的两个损失函数，确保输出水果识别结果与输入水果类别一致。池化层中的损失函数定义为：

$$L_{pl}(W,x^k) = \frac{1}{2m}\left[\sum_{i=1}^{m}\left(y_k - U * x_{ik}^2 + \sigma\sum_{j=1}^{n}\mu_j^2\right)\right]，\tag{4}$$

式中：

U——池化参数，$U=[\mu_1,\mu_2,\cdots,\mu_t]$；

t——参数总数。

σ 通常由专家知识决定，并将根据测试结果进行选择，然后得到每个待检测图像的卷积特征映射。区域生成网络以任意大小的待识别图像作为输入，并提供一组

矩形候选区域作为输出。

包括 4×4、2×2 和 1×1 向量的特定转换向量被用于池化过程。首先,使用 4×4 向量获得 16 个区域,每个区域都有最大池;接下来,应用 2×2 和 1×1 向量进行相同的操作。因此,任何具有不同大小的区域的缩减维数都设置为 21。然后,感兴趣区域池化层被馈入两个完全连接层——一个盒分类层和一个盒回归层。分类器和回归器的工作方式:

(1) 分类器:在感兴趣区域池化层计算候选区域包含对象的概率。在特征映射完全遍历后,使用 softmax 函数计算特征映射 i 中每个元素包含目标对象的概率 P_i。P_i 总排名前 300 的区域为候选区域[16]。

(2) 回归器:在回归器中,并集上的交集(Intersection over union,IoU)被用作衡量边界框的准确性的指标[17],并检测锚框中心点的坐标(x,y)以及该框的宽度 w 和长度 h。可通过式(5)计算回归器衡量边界框准确性的指标。

$$I_{\mathrm{IOU}}=\frac{A\cap B}{A\cup B}, \tag{5}$$

式中:

I_{IOU}——衡量边界框准确性的指标;

A、B——两个候选区域的面积。

假设(x,y)是候选区域中心点的坐标,而(w,h)是该区域的宽度和高度。那么,可以使用四维索引(x,y,w,h)来描述候选区域。(P_x,P_y,P_w,P_h)、(G'_x,G'_y,G'_w,G'_h)和(G_x,G_y,G_w,G_h)分别代表锚定框、预测框和实际框。IoU 的窗口可以在回归过程中调整原始候选区域的边缘。

原始候选区域和实际区域之间的误差为 $t_k=(t_x,t_y,t_w,t_h)$,其中:

$$t_x=\frac{G_x-P_x}{P_w}, \tag{6}$$

$$t_y=\frac{G_y-P_y}{P_h}, \tag{7}$$

$$t_w=\lg\left(\frac{G_w}{P_w}\right), \tag{8}$$

$$t_h=\lg\left(\frac{G_h}{P_h}\right)。 \tag{9}$$

因此,水果识别的目标函数 L_e 可以表示为:

$$L_e=\min\sum_{i=1}^{N}\left[t_K^i-\omega_K^{\mathrm{T}}\varphi(P_i)\right]^2, \tag{10}$$

式中:

$\varphi(P_i)$——输入特征向量;

ω_K^{T}——从训练中获得的系数;

N——样本总数。

目标函数可以用最小二乘法求解。因此,分类器输出 $2k$ 个分数来估计每个候选区域的对象或背景的概率,而回归器输出 k 个输出用以编码 $4k$ 个框的坐标。k 个候选区域与 k 个参考框相关联,表示为锚框。区域生成网络可以使用相同的 CNN 模型来获得,而无需训练单独的模型,这大大降低了处理速度。

2　数据集和性能指标

2.1　自建水果图像数据库

采用地面车辆自主牵引的水果图像采集装置(图 2)采集 30 个果园中的水果图像,该水果图像采集装置使用高质量的数码相机在行间果树旁拍摄水果图像。立体相机放置在距地面 1.83 m 的位置,4 个 LED 频闪灯距离地面的高度分别为 0.91,1.52,2.13,2.74 m。整个系统安装在一辆改装的四轮平板车上。用于收集数据时,其行驶速度约为 2 km/h。

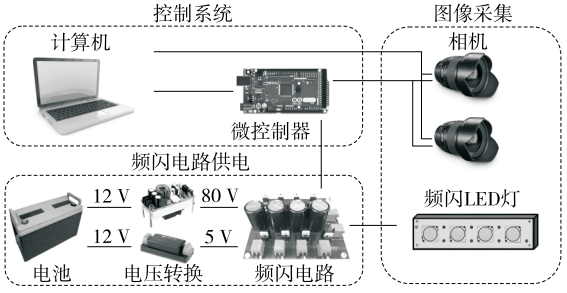


图 2　水果图像采集装置及其控制系统

Figure 2　Fruit image acquisition device and itscontrol system

2.1.1　相机传感器和镜头　相机传感器和镜头是根据薄镜头模型的近似值选择的。假设焦距固定,无限远对焦,焦距与镜头到像平面的距离相同,则相机传感器上苹果的大小与三维空间之间的关系可表示为:

$$W_{\mathrm{D}}=\frac{L_{\mathrm{F}}\times H_{\mathrm{O}}}{H_{\mathrm{SO}}}, \tag{11}$$

式中:

W_{D}——工作距离(相机到苹果的距离),mm;

L_{F}——相机镜头的焦距,mm;

H_{O}——三维苹果的高度,mm;

H_{SO}——相机传感器上苹果的高度,mm。

传感器高度、图像中苹果高度、图像高度、相机传感器上苹果高度满足以下关系:

$$\frac{H_{\mathrm{IO}}}{H_{\mathrm{SO}}}=\frac{H_{\mathrm{I}}}{H_{\mathrm{S}}}, \tag{12}$$

式中:

H_{IO}——图像上观察到的苹果高度,mm;

H_{I}——图像的高度,mm;

H_{S}——相机传感器的高度,mm。

结合式(3)和式(4),使用传感器上的苹果高度(H_{SO})作为公共变量,则:

$$\frac{W_{\mathrm{D}}}{H_{\mathrm{O}}}=\frac{L_{\mathrm{F}}\times H_{\mathrm{I}}}{H_{\mathrm{S}}\times H_{\mathrm{IO}}}。 \tag{13}$$

物体高度(H_O)和图像物体高度(H_{IO})均设置为 1，表示希望工作距离处三维坐标系中的 1 mm 图像至少由图像坐标系中的 1 个像素表示。图像高度(H_I)和传感器高度(H_S)基于根据确定的相机传感器类型进行选择。

2.1.2　人造光源　使用定制的有源光源来提高室外野外条件下的图像质量。图 2 中的频闪 LED 灯由一组板上芯片(Chip-on-board，COB)式 LED 组成，图像采集装置中共设置有 4 组频闪 LED 灯，它们距离地面的高度分别为 0.91，1.52，2.13，2.74 m。这些 LED 在短时间内过流驱动，产生强大的频闪效果。每个 LED 的额定功率为 173 W、光通量为 12 372。每个 LED 上的短脉冲功率增加到 1.6 kW，单个闪光灯的总输出功率为 25.6 kW。图 2 中的频闪电路使用 12 V、0.216 kW·h 密封铅酸电池供电，系统的其余部分使用单独的电池电源。

图像捕获期间的人工照明作用：① 减少阳光和阴影的影响，从而产生更均匀的亮度和彩色图像；② 允许缩短相机的快门持续时间，这有助于减少影响相机系统在现场移动的图像中的运动模糊。当摄像头系统在行间移动时，由于地面振动，图像容易受到垂直运动模糊的影响；然而，考虑到相机的快门持续时间较短，在整个图像采集时间段内任何数据收集试验中都没有在图像中观察到运动模糊效果。

LED 的过流驱动可能会因发热而损坏元器件；但在短时间和适当冷却的情况下，损坏概率较小。通过该水果图像采集装置共采集 85 种水果的 11 788 张图像，并以这些水果图像为基础，结合 google 图像搜索获得的另外 46 种水果的 4 600 张图像，自建水果图像数据库。

2.2　数据集

测试水果识别方法性能的水果图像数据中包含 131 种水果。数据来自 4 个数据库：Fruits 360 数据库[18]、自建数据库、ImageNet 数据库[19]和 VegFru 数据库[20]。所有图像均被调整为 100 像素×100 像素。这些水果图像均是围绕固定轴旋转水果进行拍摄，每个水果果实有 2～3 个不同的轴。为训练和测试水果识别方法，选择数据集中的一部分水果图像构建原始数据集，并对其进行水果标记。

每种水果的标记数据集被分成训练集和测试集，如表 1 所示。

表 1　原始数据集

Table 1　Raw data set

数据库	训练数据集中图像数	测试数据集中图像数	总水果图像数
Fruits 360	6 550	3 275	9 825
自建数据库	3 930	1 965	5 895
ImageNet	6 550	3 275	9 825
VegFru	6 550	3 275	9 825

2.3　数据增强

增强样本图像可以提高样本的质量和多样性，这有助于提高水果识别的准确率和精度[21]。例如，“Stern”可以增加有亮度变化的样本图像以提高目标识别精度。果园在自然光照下，尤其是强光下，部分覆盖水果或逆光拍摄形成的果实外影，使得正常光照下的区域与漫射区域的颜色差异较大；这将影响水果样本的图像质量，进而影响水果识别方法的识别准确率和精度。所改进快速 R-CNN 通过自适应直方图均衡技术增强水果图像，以减少光线对图像质量的影响，这相当于调整了图像的亮度，增加了样本照明条件的多样性。考虑到大部分水果果实以垂直悬挂状态挂在树上，部分果实由于枝干条件和果实的相互覆盖而处于不同角度的斜挂状态，将样本水平镜像旋转 10°和 20°，从中心截取旋转后的图像。如果旋转后图像边缘的目标有任何不完整甚至完全丢失，水果标记将被舍弃。

2.4　评估指标

利用准确率、回召率、精度和 F1 分数对训练好的水果识别方法进行水果识别效果评估。准确率、回召率、精度和 F1 分数的定义为

$$A_{\text{accuracy}}=\frac{N_{\text{TP}}+N_{\text{TN}}}{N_{\text{TP}}+N_{\text{TN}}+N_{\text{FP}}+N_{\text{FN}}}\times 100\%\ ,\tag{14}$$

$$P_{\text{precision}}=\frac{N_{\text{TP}}}{N_{\text{TP}}+N_{\text{FP}}}\ ,\tag{15}$$

$$R_{\text{recall}}=\frac{N_{\text{TP}}}{N_{\text{TP}}+N_{\text{FN}}}\ ,\tag{16}$$

$$F_{\text{F1score}}=\frac{2\times P_{\text{precision}}\times R_{\text{recall}}}{P_{\text{precision}}+R_{\text{recall}}}\ ,\tag{17}$$

式中：

N_{TP}——真实的阳性样本数；

N_{FP}——假阳性样本数；

N_{FN}——假阴性样本数；

N_{TN}——真阴性样本数；

C——水果类别数；

A_{accuracy}——准确度；

$P_{\text{precision}}$——精确率；

R_{recall}——回召率；

F_{F1score}——F1 分数。

3　仿真试验与结果分析

3.1　试验环境

所有训练和测试均在同一台笔记本电脑上进行，电脑主要配置为 Intel Core(TM) i7-7700HQ CPU @ 2.80 GHz、6 GB GPU GeForce GTX 1060 和 16 GB 随机存取内存。所有程序均采用 C++编写，调用 CUDA、OPENCV 库，在 Windows10 系统下运行。

3.2　水果识别方法性能测试

将研究提出的改进 Faster R-CNN 方法在测试数据

集上进行了水果识别测试试验。测试结果表明,研究提出的方法对水果识别的准确率、精度和回召率达到99.69%,0.996 8,0.994 8。以对 19 种常见水果的识别为例,研究提出的方法对 19 种常见水果的识别准确率混淆矩阵示例如图 3 所示。

根据生成的混淆矩阵,除 7 个水果类外,研究提出的方法对其余 12 种典型水果的分类准确率达到 100%。从混淆矩阵推断,对某些水果类别错误识别的可能原因是由于水果表面混合色块引起的网络层非常相似的激活。如图 4 所示,识别准确率最低的是橙子,容易与橘子、柚子、皇帝柑等混淆,识别准确率为 77.3%;其次是鸡心果,容易与一些苹果混淆,识别准确率为 92.13%。尽管研究提出的方法对一些水果的识别准确率仍有待提高,但其整体水果识别准确率达到了 99.69%,可以用于实际水果分类。

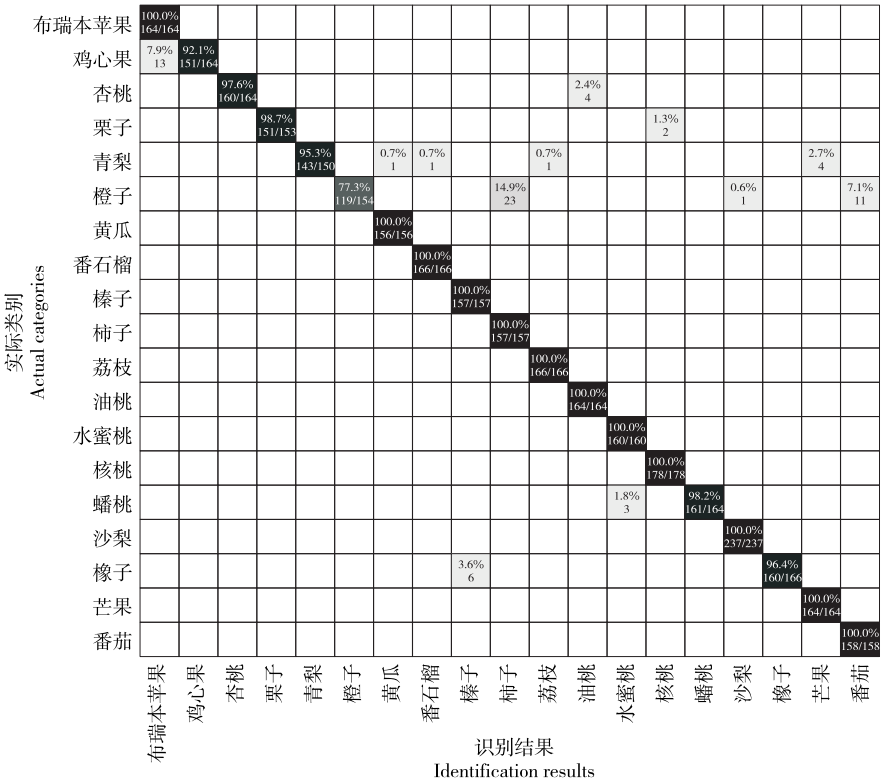
将取自 Fruits 360 数据库、自建数据库、ImageNet 数据库和 VegFru 数据库的水果图像用于所提出方法的水果识别性能测试时,水果识别精度与回召率的关系曲线如图 4 所示,无论水果图像来自哪个数据库,研究提出的方法均表现出良好的水果识别性能。

对 Fruits 360 数据库、自建数据库、ImageNet 数据库和 VegFru 数据库中水果进行识别的效果如表 2 所示,对于所有数据集,研究提出的方法都取得了良好的水果识别准确率、精度、回召率和 F1 分数,充分证明了研究提出的方法对不同水果识别条件和环境的强泛化能力。

3.3　水果识别性能比较

将研究提出的方法(基于改进 Faster R-CNN 模型的水果识别方法)的水果识别性能与 13 层 CNN 模型[12]、改进的 GoogleNet[22]、8 层 CNN 模型[23]、Mask R-CNN[24]、Faster R-CNN[11]、CNN-LSTM[25]、ResNet50[26] 和 DCDM[27]进行了对比,结果如表 3 所示,研究提出的方法对水果识别的准确率、精度和回召率比其他方法至少提高了 0.91%,1.32%,0.51%。

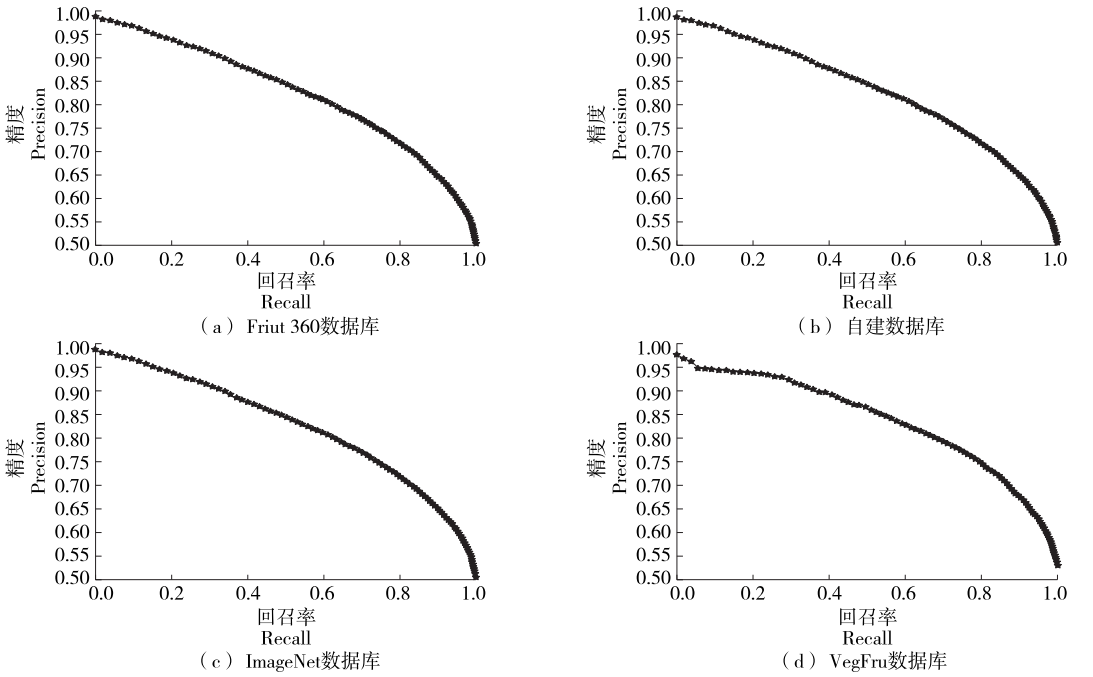
训练时间取决于水果识别过程中考虑的水果类别数量、使用的基本模型类型和实现的算法。例如,在文献[26]中,ResNet50 考虑了较少的迭代次数和仅三类水果的识别;在文献[11]中,Faster R-CNN 模型只对一类水果进行识别;在文献[12]中,8 层和 13 层的 CNN 模型仅考虑了 18 类水果的识别;在对研究提出的方法进行训练时,考虑了 120 种水果。因此,与其他水果分类方法相比,研究提出的改进 Faster R-CNN 模型的训练时间相对较长。



图中对角线表示正确识别样本的百分比,非对角线表示错误识别的百分比

图 3　研究提出的方法对 19 种常见水果的识别准确率混淆矩阵示例

Figure 3　An example of the confusion matrix for the recognition accuracy of 19 common fruits by the proposed method



（a）Friut 360数据库　（b）自建数据库　（c）ImageNet数据库　（d）VegFru数据库

图 4　对多个数据库中水果识别的精度和回召率关系图

Figure 4　Relationship maps between the accuracy and the recall rate of fruit identification for multiple databases

表 2　研究提出的方法对多个数据库中水果的识别性能

Table 2　Fruit identification performance for multiple databases of the proposed method

数据集	准确率/%	精度	回召率	F1 分数
Fruits 360	99.69	0.997	0.985	0.997
自建数据库	99.07	0.990	0.981	0.989
ImageNet	99.61	0.995	0.984	0.996
VegFru	98.44	0.983	0.976	0.984

表 3　各种水果识别方法的识别性能对比

Table 3　Comparison of fruit identification performances of various fruit identification methods

方法	准确率/%	训练时间/min	精度	回召率
13 层 CNN[12]	94.94	140.26	0.951	0.930
改进的 GoogLeNet[22]	92.23	85.72	0.921	0.909
8 层 CNN[23]	95.67	142.33	0.958	0.947
Mask R-CNN[24]	86.14	31.77	0.973	0.957
Faster R-CNN[11]	90.72	28.54	0.980	0.962
CNN-LSTM[25]	98.17	78.98	0.984	0.977
ReNet50[26]	95.86	96.10	0.957	0.944
DCDM[27]	98.78	63.00	0.984	0.980
研究提出的方法	99.69	261.60	0.997	0.985

以上对比结果表明，研究提出的基于改进 Faster R-CNN 模型的水果识别方法的水果识别准确率、精度和回召率均优于所对比的其他水果识别方法。该对比结果及研究提出方法的泛化能力的验证结果充分说明了所提方法的有效性和实用性。

4　结论

研究提出了一种基于改进 Faster R-CNN 模型的水果识别方法，对所提方法的测试结果表明，该法对 19 种水果的整体识别准确率达到了99.69%，对多个水果数据库中水果均表现出良好的识别性能，与其他 8 种水果识别方法相比，所提方法对水果识别的准确率、精度和回召率至少提高了 0.91%，1.32%，0.51%。与 13 层 CNN 等识别方法相比，该方法可提高水果识别准确率和泛化能力，但模型训练时间相对较长。后续将继续研究水果识别性能更好的方法，轻量化水果识别模型，在保持或提高水果识别准确率和泛化能力基础上，提升模型训练速度。

参考文献

[1] 胡国喜, 王超, 刘宇珩. 基于改进光滑滑模阻抗控制的水果分拣机械人夹持机构控制方法[J]. 食品与机械, 2021, 37(2): 130-134.
HU G X, WANG C, LIU Y H. A control technique for clamping mechanism of fruit sorting robot based on improved smooth sliding mode impedance control[J]. Food & Machinery, 2021, 37(2): 130-134.

[2] 王成军, 韦志文, 严晨. 基于机器视觉技术的分拣机器人研究综述[J]. 科学技术与工程, 2022, 22(3): 893-902.
WANG C J, WEI Z W, YAN C. Review on sorting robot based on

machine vision technology[J]. Science Technology and Engineering, 2022, 22(3): 893-902.

[3] 唐媛红, 刘月云. 基于 PLC 的水果分拣机器人动作控制优化[J]. 农机化研究, 2021, 43(12): 233-237.
TANG Y H, LIU Y Y. Optimization of fruit sorting robot action control based on PLC[J]. Journal of Agricultural Mechanization Research, 2021, 43(12): 233-237.

[4] 魏冉, 裴悦琨, 姜艳超, 等. 基于改进 Faster R-CNN 模型的樱桃缺陷检测[J]. 食品与机械, 2021, 37(10): 98-105, 201.
WEI R, PEI Y K, JIANG Y C, et al. Detection of cherry defects based on improved Faster R-CNN model[J]. Food & Machinery, 2021, 37(10): 98-105, 201.

[5] 金宗毅, 陶学恒, 金艳, 等. 海产品智能分拣装备与系统设计[J]. 食品与机械, 2022, 38(3): 75-79.
JIN Z Y, TAO X H, JIN Y, et al. Intelligent sorting equipment and system design of seafood[J]. Food & Machinery, 2022, 38(3): 75-79.

[6] 刘春妹, 高洪民, 王学田, 等. 基于深度学习的水果图像识别系统[J]. 微波学报, 2020, 36(S1): 427-430.
LIU C M, GAO H M, WANG X T, et al.Fruit image recognition system based on deep learning[J]. Journal of Microwaves, 2020, 36(S1): 427-430.

[7] INDIRA D N V S L S, GODDU J, INDRAJA B, et al. A review on fruit recognition and feature evaluation using CNN[J]. Materials Today: Proceedings, 2023 80: 3 438-3 443.

[8] 郑太雄, 江明哲, 冯明驰. 基于视觉的采摘机器人目标识别与定位方法研究综述[J]. 仪器仪表学报, 2021, 42(9): 28-51.
ZHENG T X, JIANG M Z, FENG M C. Vision based target recognition and location for picking robot: A review[J]. Chinese Journal of Scientific Instrument, 2021, 42(9): 28-51.

[9] NARANJO-TORRES J, MORA M, HERNÁNDEZ-GARCÍA R, et al. A review of convolutional neural network applied to fruit image processing[J]. Applied Sciences, 2020, 10(10): 3 443.

[10] ZHONG S H, XU W P, ZHANG T H, et al. Identification and depth localization of clustered pod pepper based on improved Faster R-CNN[J]. IEEE Access, 2022, 10: 93 615-93 625.

[11] WAN S H, GOUDOS S. Faster R-CNN for multi-class fruit detection using a robotic vision system[J]. Computer Networks, 2020, 168: 107036.

[12] ZHANG Y D, DONG Z, CHEN X, et al. Image based fruit category classification by 13-layer deep convolutional neural network and data augmentation[J]. Multimedia Tools and Applications, 2019, 78: 3 613-3 632.

[13] DU L, SUN Y Q, CHEN S, et al. A novel object detection model based on faster R-CNN for spodoptera frugiperda according to feeding trace of corn leaves[J]. Agriculture, 2022, 12(2): 248.

[14] MHALLA A, CHATEAU T, MAAMATOU H, et al. SMC faster R-CNN: Toward a scene-specialized multi-object detector[J]. Computer Vision and Image Understanding, 2017, 164: 3-15.

[15] LI L, HASSAN M A, YANG S R, et al. Development of image-based wheat spike counter through a Faster R-CNN algorithm and application for genetic studies[J]. The Crop Journal, 2022, 10(5): 1 303-1 311.

[16] KARAR M E, ALSUNAYDI F, ALBUSAYMI S, et al. A new mobile application of agricultural pests recognition using deep learning in cloud computing system[J]. Alexandria Engineering Journal, 2021, 60(5): 4 423-4 432.

[17] 黄晓红, 李铁锋, 刘祥鑫, 等. 基于改进 Faster R-CNN 算法的岩石裂隙发展方向跟踪预测研究[J]. 河南理工大学学报(自然科学版), 2022, 41(4): 134-141.
HUANG X H, LI T F, LIU X X, et al. Research on tracking and prediction of rock fissure development based on improved Faster R-CNN algorithm[J]. Journal of Henan Polytechnic University (Natural Science), 2022, 41(4): 134-141.

[18] KAGGLE. Fruits 360 dataset: A dataset of images containing fruits and vegetables[DB/OL]. [2023-04-20]. https://www.kaggle.com/moltean/fruit.

[19] MORID M A, BORJALI A, DEL FIOL G. A scoping review of transfer learning research on medical image analysis using ImageNet[J]. Computers in Biology and Medicine, 2021, 128: 104115.

[20] ZHAO Q, WANG X, LYU S C, et al. A feature consistency driven attention erasing network for fine-grained image retrieval[J]. Pattern Recognition, 2022, 128: 108618.

[21] 倪云峰, 叶健, 樊娇娇. 基于图像识别的水果分拣系统[J]. 江苏农业科学, 2021, 49(10): 170-176.
NI Y F, YE J, FAN J J. Fruit sorting system based on image recognition[J]. Jiangsu Agricultural Science, 2021, 49(10): 170-176.

[22] STEINBRENER J, POSCH K, LEITNER R. Hyperspectral fruit and vegetable classification using convolutional neural networks [J]. Computers and Electronics in Agriculture, 2019, 162: 364-372.

[23] WANG S H, CHEN Y. Fruit category classification via an eight-layer convolutional neural network with parametric rectified linear unit and dropout technique[J]. Multimedia Tools and Applications, 2020, 79: 15 117-15 133.

[24] JIA W K, TIAN Y Y, LUO R, et al. Detection and segmentation of overlapped fruits based on optimized mask R-CNN application in apple harvesting robot[J]. Computers and Electronics in Agriculture, 2020, 172: 105380.

[25] VASUMATHI M T, KAMARASAN M. An effective pomegranate fruit classification based on CNN-LSTM deep learning models[J]. Indian Journal of Science and Technology, 2021, 14(16): 1 310-1 319.

[26] KHAN A, NAWAZ U, UIHAQ A, et al. Real-time plant health assessment via implementing cloud-based scalable transfer learning on AWS DeepLens[J]. PLoS One, 2020, 15(12): e0243243.

[27] XUE G, LIU S F, MA Y C. A hybrid deep learning-based fruit classification using attention model and convolution autoencoder [J]. Complex & Intelligent Systems, 2023, 9: 2 209-2 219.