

顾伟,王巧华,李庆旭,等.基于改进 SSD 的棉种破损检测[J].华中农业大学学报,2021,40(3):278-285.  
DOI:10.13300/j.cnki.hnlkxb.2021.03.032

# 基于改进 SSD 的棉种破损检测

顾伟,王巧华,李庆旭,施行,张洪洲

华中农业大学工学院/农业农村部长江中下游农业装备重点实验室,武汉 430070

**摘要** 为实现群体棉籽的破损检测,以新路早-50#脱绒棉籽为研究对象,将群体棉籽随机摆放,使用 CCD 相机采集群体棉籽的图像,在经典的单步多框检测(single shot multibox detector, SSD)算法上进行改进。基于改进 SSD,利用 ResNet 50 网络代替经典 SSD 算法中的 VGG 网络,将 ResNet50 作为 SSD 的基础网络,用来快速提取群体棉籽图像的特征,最终对群体脱绒棉籽中的破损棉籽实现精准识别。试验结果表明:该方法建立的模型对群体棉籽的检测精度、召回率、漏检率分别达到 96.1%、97.3%、0%;高于经典 SSD 网络模型(检测精度、召回率、漏检率分别为 92.5%、96.4%、1.4%)。

**关键词** 脱绒棉种; 破损棉籽检测; 无损检测; ResNet50; SSD; 精准识别; 分选智能化; 棉籽精选

**中图分类号** S 561; TP 391    **文献标识码** A    **文章编号** 1000-2421(2021)03-0278-08

棉种在轧花、脱绒或磨光等过程中,都会有不同程度的破损,这将直接影响棉花的产量和品质<sup>[1]</sup>。新疆南疆地区是我国重要的棉花生产基地,棉籽精选对提高该地区的棉花产量具有重要意义。目前,国内外有关种子检测方法的研究较多,如王亚丽等<sup>[2]</sup>、Tigabu 等<sup>[3]</sup>使用近红外(NIR)光谱分别对玉米种子和松树种子进行分类;董翠翠等<sup>[4]</sup>、闫彬等<sup>[5]</sup>采用机器视觉对种子的颜色以及破损情况进行研究与分选;余淑华等<sup>[6-7]</sup>运用色选理论对棉籽的成熟度以及完整度进行识别;Golshan 等<sup>[8]</sup>、Brennen 等<sup>[9]</sup>利用不同的算法自动检测图像中的种子;Dean 等<sup>[10]</sup>利用种子的密度成功分选了候选突变体。以上研究大多都针对单粒种子的检测,单粒棉籽的检测效率低,实际应用很难满足生产需求,因此提高棉籽检测效率是该研究的主要方向。

基于传统的机器学习方法实现破损棉籽的鉴别,需要人工提取棉籽数据特征才能建立破损棉籽的判别模型,因此该方法的检测效率很难满足生产需要<sup>[11]</sup>。卷积神经网络(deep convolutional neural network)自动提取图像特征优势明显,且在农业检测领域应用广泛<sup>[12-13]</sup>。由于棉籽尺寸较小,在图像中属于小目标,手动提取特征较为困难。SSD-

VGG16 网络能够快速检测出图像中的棉籽,并识别其类别与位置信息,适用于多目标检测,但对于群体棉籽的品质检测还需要进一步的改进。

本研究在 SSD 网络<sup>[14]</sup>的基础上,提出一种基于 ResNet50 网络<sup>[15]</sup>改进的群体棉籽破损检测方法,利用 ResNet50 网络作为基础网络自动提取棉籽图像特征,并利用 SSD 网络对群体棉籽实现破损判别,旨在为检测群体破损棉籽信息提供新的思路,加速分选棉籽智能化,进一步提高棉籽检测效率,为后续研发相关自动化设备提供技术支持。

## 1 材料与方法

### 1.1 试验图像视觉采集系统

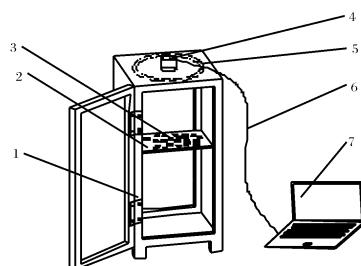
棉籽采集装置主要是由暗箱、光源、相机、镜头、背景板以及计算机组成(图 1)。相机使用 AD-080GE 工业相机,分辨率为  $1024 \times 768$  dpi, 帧率为 30 FPS, 品牌为 JAI; 镜头使用 Kowa 公司制造的 LM6NC3 型号。光源为 12 W 的正白光 LED 环形光源,采集计算机为 Lenovo B50-30,采集软件为 JAI SDK。棉籽在暗箱内的物距、焦距和光圈在采集过程中保持不变。

收稿日期: 2020-12-01

基金项目:国家自然科学基金项目(61701334);兵团南疆重点产业科技支撑项目(2018DB001);塔里木大学中国农业大学联合基金项目(TDZNLH201703)

顾伟, E-mail: 534893320@qq.com

通信作者:王巧华, E-mail: wqh@mail.hzau.edu.cn



1.暗箱 Black box; 2.背景板 Background plate; 3.棉籽 Cottonseed; 4.相机和镜头 Camera and lens; 5.环形光源 Ring light; 6.数据传输线 Data transmission line; 7.计算机 Computer.

图 1 图像采集系统

Fig.1 Image acquisition system



图 2 棉籽采集图像

Fig.2 Image collection of cottonseed

籽,其中训练、验证、测试集按 7:2:1 比例划分,10 000 粒棉籽图像具有足够的泛化代表性,以此来验证棉籽检测的准确率。

共采集 500 张群体棉籽图片,每张图片约有 20 粒棉籽,分辨率为  $1024 \times 768$  dpi,将图像分辨率调小至  $300 \times 300$  dpi,在使用 SSD 训练数据前,需要手动标记并制作 tfrecord 标签文件,图像标注时利用棉籽的最小外接矩形进行标注,保证每粒棉籽尽量在 1 个矩形框里。标注完成将会生成含有类别和位置信息的 xml 文件。标注完成后将 500 幅图像划分为训练集、测试集和验证集,其中训练集 350 幅图像、测试集 100 幅图像、验证集 50 幅图像。每幅图像含有约 20 个棉籽样本,500 幅图像中含边框标注样本个数为 10 000(其中破损棉籽和非破损棉籽的比例为 1:1),训练集中棉籽 7 000 粒,测试集中棉籽 2 000 粒,验证集中棉籽 1 000 粒,能够满足深度学习样本数量要求,具体数据如表 1 所示。

#### 1.4 基于改进 SSD 群体棉籽破损分类实现

1)改进前 SSD 模型结构。本研究主要目标是从采集的一幅群体棉籽图像中检测出棉籽目标并将其分类。SSD 是基于前向传播的 CNN 网络,经典 SSD 模型主网络是由 VGG16 中部分卷积层组成,并将最后 2 层的 Conv6 和 Conv7 换成全连接层,用来进行图像分类。训练模型时,检测框架的损失函

#### 1.2 试验原始图像数据

非破损棉籽原图如图 2 A 所示,破损棉籽图像如图 2 B 所示,采集得到的群体棉籽图像如图 2 C 所示,通过棉籽表面可以明显观察出破损棉籽与非破损棉籽的区别,破损棉籽表面坑洼,非破损棉籽表面光滑。

#### 1.3 数据集准备

取新疆新路早棉籽 10 000 粒,其中破损与非破损棉籽数量比例为 1:1。将破损与非破损棉籽按 1:1 比例均匀洒落在背景板上,利用 CCD 相机采集群体棉籽图像,之后用 LabelImg 软件对采集的图像进行人工标注,分别标注出破损棉籽与非破损棉

表 1 棉籽试验数据集

Table 1 Cottonseed test data sets

项目 Item	训练集 Training sets	测试集 Test sets	验证集 Validation sets	总数 Total
图片数/张 Number of pictures	350	100	50	500
棉籽数/粒 Number of cottonseed	7 000	2 000	1 000	10 000

数用位置损失( $L_{conf}$ )和置信损失( $L_{loc}$ )的加权和表示,如公式(1)所示。对于目标检测,基于区域的卷积神经网络有 RCNN<sup>[16]</sup>、Fast RCNN<sup>[17]</sup> 和 Faster RCNN<sup>[18]</sup>,核心思想是在卷积层后接入了感兴趣区域(region of interest, ROI),池化层使网络可以定位 ROI<sup>[19]</sup>,然后在当前区域内进行分类。SSD 的检测框架由 2 个部分组成(图 3),一部分是前面的卷积神经网络,经典模型中采用的是 VGG16,用于对图像特征的初步提取,第二部分是位于后端的多尺度特征检测网络,用于对前端网络产生的特征层进行不同尺度条件下的特征提取,利用多尺度特征检测网络当中的浅层特征图尺寸较大的特点,检测棉籽这一类尺寸较小的物体。图 3 所示 SSD 针对 Conv4\_3、FC7、Conv8\_2、Conv9\_2、Conv10\_2 和 Conv11\_2 特征层的每个单元按照不同长宽比分别提取 4~6 个默认框,最终获取 8 732

个默认框,训练阶段利用公式(1)为预测框做坐标偏移的回归分析,使其尽可能接近标注框。测试阶段每个默认框都和标注框进行重叠率匹

配,并按照匹配分数从高到低排序,利用非极大值抑制<sup>[20]</sup>的方法,去掉所有冗余框,保留置信度最大的框。

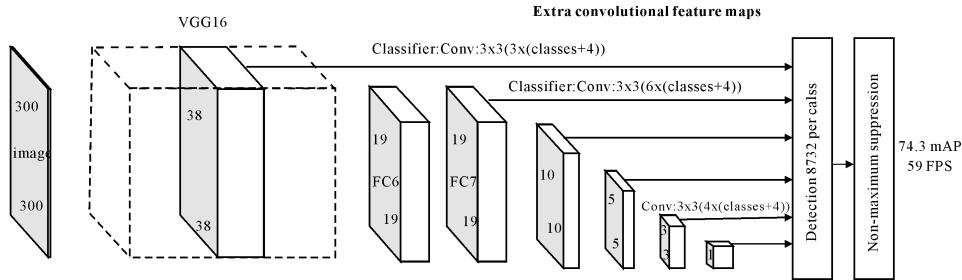


图 3 改进前的 SSD 网络结构

Fig.3 Improving the former SSD network structure

$$L(x, c, l, g) = \frac{1}{N} (L_{\text{conf}}(x, c) + \alpha L_{\text{loc}}(x, l, g)) \quad (1)$$

其中:

$$L_{\text{loc}}(x, l, g) = \sum_{i \in \text{Pos}} \sum_{m \in \text{Box}} x_{ij,k} S_{L1}(l_{i,m} - \hat{g}_{j,m}) \quad (2)$$

$$L_{\text{conf}}(x, c) = - \sum_{i \in \text{Pos}} x_{ij,p} \lg \hat{c}_{i,p} - \sum_{i \in \text{Neg}} \lg \hat{c}_{i,0} \quad (3)$$

$$\hat{c}_{i,p} = \frac{\exp c_{i,p}}{\sum_p \exp c_{i,p}} \quad (4)$$

式(1)~(4)中:N 为匹配到真实目标的先验框的数量;l 为先验框;g 为真实框;c 为 Softmax 函数对每类别的置信度;α 为调整位置损失和置信损失之间的比例; $\hat{c}_{i,p}$  为类别 p 第 i 个默认框的置信度; $\hat{g}_{j,m}$  为获取真实标签的近似回归先验框。

2) 改进 SSD 模型。SSD-ResNet 网络是基于经典的 SSD 网络改进而来的目标检测算法,ResNet 在 ImageNet 数据集上的错误率远低于 VGG 网络。本研究选用 ResNet50 代替原有的 VGG16 模型作为 SSD 的基础网络,来提取群体棉籽的特征。考虑到生产实际应用,需要综合考虑算法的效率以及精度,模型大小不宜太大,因此未选用更深层次的网络对 SSD 进行改进。经典 SSD 网络可分为基础网络和辅助网络,SSD-ResNet50 将基础网络使用 ResNet50 的前 4 个卷积阶段代替原 SSD 中的 VGG16 的前 5 个卷积阶段,并将 ResNet50 的第 4 个卷积阶段的步长都改为 1,以实现对输入图像特征的提取。ResNet 网络引入了一种残差学习框架,以应对传统网络的退化问题,由试验可知,这种残差学习框架有利于棉籽的检测,如图 4 所示为普通结构与残差结构的差异。

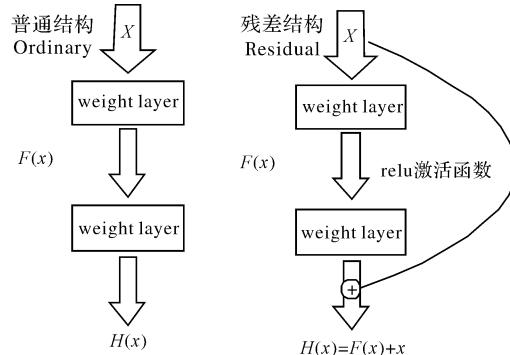


图 4 普通与残差对比图

Fig.4 Comparison between ordinary and residual

ResNet 使用的方法就是对多层的残差映射进行拟合,假设  $H(x)$  为几个网络层堆叠的期望映射,  $x$  即为当前堆叠块的入口,使用的 relu 激活函数缩短学习周期,使用 relu 激活函数越多,节省的参数就越多,在使用 relu 激活函数过程中,必须保持经过残差的堆叠块  $x$  和最后输出的期望映射  $H(x)$  的高度、宽度、维度一致。假设  $n$  个非线性层可近似地表达为某个复杂函数(残差函数),将堆叠起来的网络层拟合成为另一个映射,映射为  $F(x) = H(x) - x$ ,从而基础映射便成为  $H(x) = F(x) + x$ 。引入残差学习框架,可以大幅降低提取特征的重复度、减少网络模型的计算量、增强网络的泛化能力,对比 VGG 网络,ResNet 网络训练完成后,随机去除某个网络层,并不会导致网络大幅度的退化,而 VGG 网络删减任何一层都会导致模型的性能崩溃。

3) 基于 ResNet50 改进 SSD 模型。基础网络(base network)是用于图像分类的标准神经网络,基础网络的理论基础是生成模型,生成模型可以自适应地从输入图像中学习重要的特征,以解决传统

的全连接网络特征提取能力不足的问题。但是由于生成模型提取到的特征信息冗余太多,无法得到棉籽的准确信息,所以使用基础网络对输入数据进行特征提取,为后续网络层提供输入信息,以加快后续棉籽检测的训练速度,提高网络的表达能力。ResNet 通过引入残差学习来提高模型的检测性能,残差学习模块的构造是合并  $n$  个堆叠块,构造块定义为:

$$y = F(x, W_i) + x \quad (5)$$

式(5)中: $x$  为模块输入向量; $y$  为输出向量; $W_i$  为权重层参数; $F(x, W_i)$  为当前网络想要学习的残差结构。

在残差模块的第 1 层公式  $F = W_2 \sigma(W_1 x)$  中,  $\sigma$  为 relu 激活函数;第 2 层则通过快捷连接来执行  $F + x$  操作,当公式(5)中的输入向量  $x$  和函数  $F$  的维度一致时,需要对输入向量  $x$  执行线性投影  $W_i$  来实现维度匹配:

$$y = F(x, W_i) + W_i x \quad (6)$$

为了提升棉籽检测网络的检测速度以及减小其

参数量,选择 Resnet 50 残差结构进行网络替换,选取的特征提取层为 conv1\_x(卷积核大小为  $7 \times 7 \times 64$ ), conv2\_x(卷积核大小为  $1 \times 1 \times 64, 3 \times 3 \times 64, 1 \times 1 \times 256$ ), conv3\_x(卷积核大小为  $1 \times 1 \times 128, 3 \times 3 \times 128, 1 \times 1 \times 512$ ), conv4\_x(卷积核大小为  $1 \times 1 \times 256, 3 \times 3 \times 256, 1 \times 1 \times 1024$ ), 卷积过后得到第 1 个预测特征层大小为  $(38 \times 38 \times 1024)$ , 再经过 conv5\_x 分别进行卷积、Batch Normalization、relu 得到第 2 个预测特征层,后 5 个阶段均按照此方法进行,一共得到 6 个预测特征层。其中在 conv5\_x、conv6\_x、conv7\_x 前 3 个阶段中步距均为 2, Padding 均为 1, 后 2 个阶段的步距均为 1, Padding 均为 0, 最后得到的后 5 个预测特征层大小分别为  $(19 \times 19 \times 512), (10 \times 10 \times 512), (5 \times 5 \times 256), (3 \times 3 \times 256), (1 \times 1 \times 256)$ , 后 5 个阶段相比较原来的 VGG16 都添加了 Batch Normalization 层, 可加速网络的收敛并提升准确率<sup>[21]</sup>。图 5 为 SSD-ResNet50 的网络结构图。

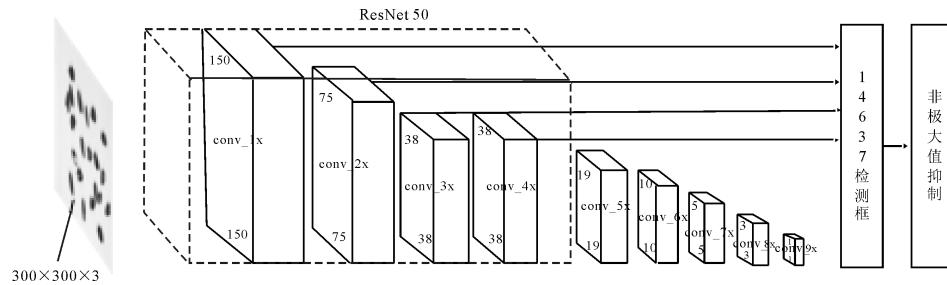


图 5 基于 ResNet50 改进的 SSD 网络结构

Fig.5 Improved SSD network structure based on ResNet50

4)训练参数设置。不同层的特征图在卷积神经网络中感受域大小也不同,此处采用不同位置的默认框对应不同的区域和目标尺寸。设有  $m$  个特征图用来预测,则每个特征图中默认框的尺寸为:

$$S_i = S_{\min} + \frac{S_{\max} - S_{\min}}{m-1}(i-1), i \in [1, m] \quad (7)$$

式(7)中: $S_{\min}$  为网络结构中最底层的默认框尺度,值为 0.2; $S_{\max}$  为最高层的默认框尺度,值为 0.95。

使用不同的宽高比取值  $a_r \in \{1, 2, 3, 1/2, 1/3\}$ , 则计算默认框的宽、高分别为:

$$w_i = S_i \sqrt{a_r} \quad (8)$$

$$h_i = S_i \sqrt{a_r} \quad (9)$$

设置默认框的中心点为  $(\frac{i+0.5}{|fk|}, \frac{j+0.5}{|fk|})$ , 其中,  $i, j \in [0, |fk|]$ ,  $|fk|$  是第  $k$  个特征图的尺寸。提取不同尺寸和宽高比的所有默认框的特征并进行结果

预测,并整合结果,解决不同尺度的目标检测问题。

SSD-ResNet 50 检测模型的生成需要使默认框匹配为真实标签框,但不是每个默认框都匹配给了真实框,1 个真实框可以匹配多个默认框,对每个真实框匹配 1 个交并比(intersection over union)值最大默认框,保证每个真实框至少能与 1 个默认框对应,之后再将剩余还没有配对的默认框与任意一个真实框尝试配对,只要两者之间的交并比大于设定的阈值(阈值为 0.5),就进行匹配。这样调整有利于弱化学习过程,也可以允许网络模型自适应地计算多个默认框的重叠情况。损失函数就是在进行模型训练时,始终存在 1 个目标函数,作用是将预测框和默认框的差距尽可能地和真实框与默认框的差距接近。算法持续对损失函数进行优化,直至损失值最低的目标函数,损失函数用来衡量网络模型的输出

值  $\hat{y}_i$  和真实值  $y_i$  的差异程度, 损失函数的目的是使损失值最小化, 其公式为:

$$L = \sum_{i=1}^N l(y_i, \hat{y}_i) \quad (10)$$

本研究建立 SSD-ResNet50 训练模型, 选择 Softmax 作为损失函数, 公式为:

$$L = -\log\left(\frac{e^{S_{yi}}}{\sum_j e^{S_j}}\right) = -S_{yi} + \log\left(\frac{e^{S_{yi}}}{\sum_j e^{S_j}}\right) \quad (11)$$

其中,  $S_j$  为类别  $j$  的得分,  $y_i$  为真实标签。不同类别上的目标分值离散程度越高, 损失值越低, 模型性能越好。

5) 迁移学习。在深度学习领域中, 利用迁移学习这种模型训练手段可以避免随机初始化网络权重和偏置等参数导致的模型收敛速度慢和网络的不稳定<sup>[22]</sup>等问题。利用大数据集下已经训练好的预训练模型, 通过共享网络前端的权重参数和微调网络后端的参数, 可以使新任务下的模型收敛更为迅速, 在小数据集下也只需要少量的训练时间和计算资源开销, 便可以实现对新任务的适应。本研究利用 COCO 数据集下预训练好的模型权重初始化 SSD 网络的权重, 通过训练对 SSD 网络参数进行微调, 进一步减少了模型的训练时间与资源。

## 2 结果与分析

### 2.1 模型训练

模型训练的硬件平台为: GPU 为 NIVIDIA GeForce RTX 2080Ti; CPU 为 AMD Ryzen Threadripper 2920X @3.5 GHz; 内存为 128 G。软件平台为: CUDA Toolkit 9.2 + CUDNN V9.2 + Python 3.5 + Tensorflow-GPU 1.8.0。由于深度学习需要大量的数据进行训练, 本研究在采集得到图片后, 将其标为破损与非破损, 训练、验证、测试集比例按 7:2:1 划分, 本研究训练和测试时均使用 GPU 加速。

模型训练的过程是模型不断调整自身参数, 使得预测值与真实值不断接近的过程, 训练过程中使用的目标损失函数为公式(11), 基于 ResNet 50 改进的 SSD 模型底层权重均使用迁移学习的方法进行初始化, 其余权重采用均值为 0、方差为 0.03 的高斯分布进行随机初始化。梯度更新是训练过程中不断调整权重和偏置的方法, 从而使模型的性能不断优化, 本研究使用批量随机梯度下降算法寻找梯度下降方向。初始学习率设置为 0.03, 学习率衰减方法为指数衰减法, 衰减率设置为 0.98。训练的批次

图像数量(batch size)设置为 16, 最大训练次数设置为  $1 \times 10^5$ , 单步训练时间约为 2 s, 每训练 10 min 保存 1 次模型的参数, 改进前后模型训练参数保持一致。以基于 ResNet 50 网络改进的 SSD 群体棉籽检测网络的训练过程为例, 改进前的 VGG16 版本训练过程与之类似。图 6 为使用了迁移学习损失函数变化曲线图, 从曲线可以看出模型的训练过程。基于 ResNet50 改进的 SSD 网络在使用迁移学习进行训练 10 743 次后维持在较低水平的小幅度震荡。

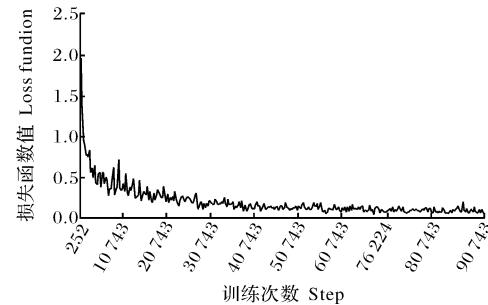


图 6 迁移学习损失函数曲线

Fig.6 Transfer learning loss function curve

### 2.2 评价指标及结果

目前评价深度学习目标检测模型性能优劣的主要手段是测试模型在验证集数据上的检测效果。常用的评价指标有准确率(precision)、召回率(recall)、漏检率(missed detection rate)、检测速度(speed)和模型大小(parameter size)。本研究利用训练好的 SSD-ResNet50 模型和改进前 SSD 模型分别对验证集 50 幅图像(含 1 000 粒棉籽)进行测试, 采用上述 5 个评价指标对比两者的优劣。漏检率定义为:

$$M = \frac{B}{N} \quad (12)$$

式(12)中,  $N$  为棉籽的总数量,  $B$  为误将棉籽当做背景的数量。

对 50 幅群体棉籽图像的测试结果进行统计, 包括准确率、召回率、漏检率、检测速度和模型大小 5 项评价指标。共 1 000 粒棉籽, SSD-ResNet50 错判了 39 粒棉籽, 其中将无损棉籽错判为破损 22 粒, 将破损棉籽判为无损棉籽 17 粒, 漏检 0 粒。SSD-VGG16 错判了 75 粒, 其中将无损棉籽错判为破损 36 粒, 将破损棉籽判为无损棉籽 25 粒, 漏检 14 粒。此外, 使用本研究所提出的方法与目前比较成熟 YOLOV3 进行了详细的对比, 3 种模型的测试结果如表 2 所示。其中, YOLOV3 网络参数设置值与上述 SSD 网络参数设置值保持一致。

表 2 模型测试结果

Table 2 Model test results

模型 Model	准确率/% Precision	召回率/% Recall	漏检率/% Missed detection rate	检测速度/FPS Speed	模型大小/Mb Parameter size
SSD-VGG16	92.5	96.4	1.4	59	93.00
基于 ResNet50 改进的 SSD Improved SSD based on ResNet50	96.1	97.3	0	86	11.10
YOLOV3	87.2	90.4	3.1	72	24.22

由表 2 可知, 基于 ResNet50 改进的 SSD 模型在精度和速度方面均优于改进前的 SSD-VGG16 网络, 精度更优于现在成熟的 YOLOV3 网络。在使用 ResNet50 过程中, 将破损棉籽错判为无损的原因是部分破损棉籽的裂纹过细, 即棉籽损伤过小, 导致误判。将非破损棉籽错判为破损的原因是部分棉籽脱绒不干净, 棉籽上面残留少量棉绒, 导致误判为破损的棉籽。YOLOV3 的准确率较低的原因在于 YOLO、FasterRCNN 等算法只会在最后一层做锚框(anchor), 因此它没有多尺度的特征, 对尺度变化不敏感。因此, 对于棉籽这类尺寸较小的物体使用 SSD 目标检测效果比较好。

综合考虑检测精度和速度, 基于 ResNet50 改进的 SSD 模型对群体棉籽的识别效果最佳。基于 ResNet50 改进的 SSD 模型的检测速度可以达到 86

FPS, 快于改进前的 59 FPS, 且模型大小仅为 11.10 Mb, 满足部署到嵌入式系统的模型参数要求。此外, 漏检率是衡量模型可靠性的重要指标, 改进后 SSD 模型的漏检率为 0, 说明模型的可靠性高于 SSD-VGG16 网络以及 YOLOV3 的网络。图 7 A、B 分别为同一幅群体棉籽图像模型和基于 ResNet50 和 VGG16 改进的 SSD 模型的检测结果, 图 7C 为 YOLOV3 模型检测结果。本研究并未使用当前最新网络对棉籽进行检测, 主要原因是后续将群体棉籽检测模型应用到生产实际, 需要考虑网络的成熟度以及抗干扰能力, 所以选用已经成熟的网络作为棉籽检测的框架。将模型应用到实际生产, 检测精度需大于 95%, 速度不小于 20 FPS, 综合考虑速度与精度, 本研究选用 SSD-ResNet50 对群体棉种进行破损识别。

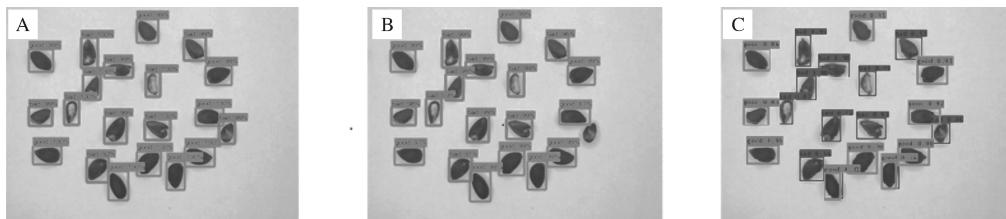


图 7 不同模型的检测图

Fig.7 Test charts of different models

### 3 讨 论

本研究以新路早-50#脱绒群体棉籽为研究对象, 使用基于 ResNet50 网络改进的 SSD 模型对群体棉籽中的破损棉籽和无损棉籽进行识别检测, 研究结果显示其对群棉籽中无损和破损的识别性能最佳, 准确率和召回率分别为 96.1%、97.3%。模型大小为 11.10 Mb, 检测速度达到了 86 FPS。本研究通过迁移学习, 将 COCO 大数据集下预训练好的模型权重迁移到群体棉籽中破损和无损检测的任务下, 加速了群体棉籽检测模型的收敛速度, 节约了模型的训练成本; 同时本研究解决了群体棉籽图像难以分割的问题, 直接利用卷积神经网络获取棉籽的位

置和类别信息, 不必再使用传统图像识别方法, 将群体棉籽中的个体单独分割出来进行检测。

### 参考文献 References

- [1] 张亮, 宋均燕, 夏鑫, 等. 棉籽仁含糖量测试方法探索[J]. 上海纺织科技, 2018, 46(5): 14-16, 23. ZHANG L, SONG J Y, XIA X, et al. Testing method of sugar content in cotton seeds[J]. Shanghai textile science & technology, 2018, 46(5): 14-16, 23 (in Chinese with English abstract).
- [2] 王亚丽, 彭彦昆, 赵鑫龙, 等. 玉米种子活力逐粒无损检测与分级装置研究[J]. 农业机械学报, 2020, 51(2): 350-356. WANG Y L, PENG Y K, ZHAO X L, et al. Design and experiment of non-destructive testing and grading device for corn seed vigor [J]. Transactions of the Chinese society for agricultural ma-

- achinery, 2020, 51(2): 350-356 (in Chinese with English abstract).
- [3] TIGABU M, DANESHVAR A, JINGJING R, et al. Multivariate discriminant analysis of single seed near infrared spectra for sorting dead-filled and viable seeds of three pine species: does one model fit all species? [J]. *Forests*, 2019, 10(6): 1-13.
- [4] 董翠翠, 李景彬, 坎杂, 等. 基于机器视觉的脱绒棉种染色特征的提取算法[J]. *农机化研究*, 2015(11): 51-54. DONG C C, LI J B, KAN Z, et al. The extract algorithms of staining characteristics of cottonseeds based on machine vision[J]. *Journal of agricultural mechanization research*, 2015(11): 51-54 (in Chinese with English abstract).
- [5] 闫彬, 杨福增, 郭文川. 基于机器视觉技术检测裂纹玉米种子[J]. *农机化研究*, 2020, 42(5): 181-185, 235. YAN B, YANG F Z, GUO W C. Detection of maize seeds with cracks based on machine vision technology[J]. *Journal of agricultural mechanization research*, 2020, 42(5): 181-185, 235 (in Chinese with English abstract).
- [6] 余淑华, 刘艳丽, 王世璞, 等. 基于 FPGA 的脱绒棉种色选机实现[J]. *农业机械学报*, 2015, 46(8): 20-26. YU S H, LIU Y L, WANG S P, et al. Implementation of delinted cottonseeds color sorter based on FPGA[J]. *Transactions of the Chinese society for agricultural machinery*, 2015, 46(8): 20-26 (in Chinese with English abstract).
- [7] 余淑华, 刘艳丽, 王世璞, 等. 脱绒棉种色选机双 CCD 检测系统设计与仿真分析[J]. *农业机械学报*, 2015, 46(9): 55-61. YU S H, LIU Y L, WANG S P, et al. Design and simulation analysis of dual CCD detection system for delinted cottonseeds color sorter[J]. *Transactions of the Chinese society for agricultural machinery*, 2015, 46(9): 55-61 (in Chinese with English abstract).
- [8] GOLSHAN M, KARIMI D, MAHDACI S, et al. Automatic detection of brachytherapy seeds in 3D ultrasound images using a convolutional neural network[J/OL]. *Physics in medicine & biology*, 2020, 65(3): 035016 [2021-04-25]. <https://iopscience.iop.org/article/10.1088/1361-6560/ab64b5>. DOI: 10.1088/1361-6560/ab64b5.
- [9] BRENNEN T J, GALLI L, CUTAJAR D L, et al. Brachy view: development of an algorithm for real-time automatic LDR brachytherapy seed detection[J/OL]. *Physics in medicine & biology*, 2020, 65(21): 215015 [2021-04-25]. <https://iopscience.iop.org/article/10.1088/1361-6560/abac9e>. DOI: 10.1088/1361-6560/ab64b5.
- [10] DEAN G H, PANG F, HAUGHN G W, et al. A simple, non-toxic method for separating seeds based on density, and its application in isolating *Arabidopsis thaliana* seed oil mutants[J]. *Applications in plant sciences*, 2020, 8(4): 11332 [2021-04-25]. <https://doi.org/10.1002/aps3.11332>.
- [11] 李庆旭, 王巧华, 顾伟, 等. 基于深度学习的种鸭蛋孵化早期受精信息无损检测[J]. *农业机械学报*, 2020(1): 188-194. LI Q X, WANG Q H, GU W, et al. Non-destructive testing of early fertilization information in duck egg laying based on deep learning [J]. *Transactions of the Chinese society for agricultural machinery*, 2020(1): 188-194 (in Chinese with English abstract).
- [12] 杨万里, 段凌凤, 杨万能. 基于深度学习的水稻表型特征提取和穗质量预测研究[J]. *华中农业大学学报*, 2021, 40(1): 227-235. YANG W L, DUAN L F, YANG W N. Deep learning-based extraction of rice phenotypic characteristics and prediction of rice panicle weight[J]. *Journal of Huazhong Agricultural University*, 2021, 40(1): 227-235 (in Chinese with English abstract).
- [13] 刘芳, 刘玉坤, 林森, 等. 基于改进型 YOLO 的复杂环境下番茄果实快速识别方法[J]. *农业机械学报*, 2020, 51(6): 229-237. LIU F, LIU Y K, LIN S, et al. Fast recognition method for tomatoes under complex environments based on improved YOLO [J]. *Transactions of the Chinese society for agricultural machinery*, 2020, 51(6): 229-237 (in Chinese with English abstract).
- [14] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot multibox detector[C/OL]//LEIBE B, MATAS J, SEBE N, et al. European conference on computer vision. Cham: Springer, 2016: 21-37 [2021-04-25]. [https://doi.org/10.1007/978-3-319-46448-0\\_2](https://doi.org/10.1007/978-3-319-46448-0_2).
- [15] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]//Anon. Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Vegas, NV, USA, 2016: 770-778.
- [16] GIRSHICK R, DONAHUE J, DARRELL T, et al. Region-based convolutional networks for accurate object detection and segmentation[J]. *IEEE transactions on pattern analysis and machine intelligence*, 2015, 38(1): 142-158.
- [17] GIRSHICK R. Fast R-CNN[C]//Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Vegas, NV, USA, 2015: 1440-1448.
- [18] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[J]. *Advances in neural information processing systems*, 2015, 39(6): 91-99.
- [19] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection[C]//Anon. Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Vegas, NV, USA, 2016: 779-788.
- [20] NEUBECK A, GOOL L V. Efficient non-maximum suppression [C/OL]//International Conference on Pattern Recognition. IEEE Computer Society, 2006, 3: 850-855 [2021-04-25]. <https://doi.org/10.1109/ICPR.2006.479>.
- [21] IOFFE S, SZEGEDY C. Batch normalization: accelerating deep network training by reducing internal covariate shift[J/OL]. *JMLR.org*, 2015 [2021-04-25]. <https://dl.acm.org/doi/10.5555/3045118.3045168>.
- [22] PAN S J, YANG Q A. A survey on transfer learning[J]. *IEEE transactions on knowledge and data engineering*, 2010, 22(10): 1345-1359.

## Improved SSD based detection of damaged cottonseed

GU Wei, WANG Qiaohua, LI Qingxu, SHI Hang, ZHANG Hongzhou

College of Engineering, Huazhong Agricultural University / Ministry of Agriculture and Rural Affairs Key Laboratory of Agricultural Equipment in Mid-Lower Yangtze River, Wuhan 430070, China

**Abstract** Cotton production in China is huge. The southern region of Xinjiang is one important cotton production base of China. At the same time, cotton production is a pillar industry in this region as well. However, the quality and sorting problem of lint-free cotton has severely limited the development of the cotton industry in this region. In order to realize the detection of damaged population cottonseed, the population cottonseed of Xinluzao-50# lint-free cottonseed were randomly arranged. The CCD camera was used to collect the image of the population cottonseed. The classic single-step multi-frame detection (single shot multibox detector, SSD) algorithm was improved. Based on the improved SSD, the ResNet50 network was used to replace the VGG network in the classic SSD algorithm. ResNet50 was used as the basic network of the SSD to quickly extract the image characteristics of the population cottonseed, and to finally realize the accurate identification of the damaged cottonseed in the population lint-free cottonseed. The results showed that the detection accuracy, recall rate, and missed detection rate of the model established by this method for the damaged cottonseed and the non-destructive cottonseed in the population cottonseed was 96.1%, 97.3%, and 0%, respectively. It is higher than that (92.5%, 96.4%, 1.4%) of the classic SSD network model. This study transfers the pre-trained model weights under the COCO large dataset to the task of damage and non-destructive detection in the population cottonseed, which accelerates the convergence speed of the population cottonseed detection model and saves the training cost of the model. It solves the problem of the difficulty of segmentation of the population cottonseed image. It directly uses the convolutional neural network to obtain the position and category information of the cottonseed. It is not necessary to use traditional image recognition methods to separate the individuals in population cottonseed for detection. It will provide a novel idea for detecting population cottonseed damage to accelerate the intelligent sorting of cottonseed and a technical support for subsequently studying and developing related automation equipment.

**Keywords** lint-free cottonseed; detection of damaged cottonseed; damage-free detection; ResNet50; SSD; accurate identification; intelligent sorting; selection of cottonseed

(责任编辑:陆文昌)