

基于机器视觉的樱桃缺陷检测与识别

Cherry defect detection and recognition based on machine vision

裴悦琨^{1,2}连明月^{1,2}姜艳超^{1,2}PEI Yue-kun^{1,2} LIAN Ming-yue^{1,2} JIANG Yan-chao^{1,2}叶家敏^{1,2}韩心新^{1,2}谷 宇^{1,2}YE Jia-min^{1,2} HAN Xin-xin^{1,2} GU Yu^{1,2}

(1. 大连大学辽宁省北斗高精度位置服务技术工程实验室, 辽宁 大连 116622;

2. 大连大学大连市环境感知与智能控制重点实验室, 辽宁 大连 116622)

(1. Beidou High Precision Positioning Service Technology Engineering Laboratory of Liaoning Province, Dalian University, Dalian, Liaoning 116622, China; 2. Environment Sensing and Intelligent Control Key Laboratory of Dalian, Dalian University, Dalian, Liaoning 116622, China)

摘要:以机器视觉技术为基础,利用卷积神经网络对樱桃缺陷进行检测与识别,并进行验证。结果表明,正常果樱桃识别准确率为 99.25%,缺陷果樱桃识别准确率为 97.99%,识别速度为 25 个/s;通过与其他方法进行对比,试验方法能够准确检测并识别多种缺陷类型。

关键词:缺陷检测;樱桃分级;机器视觉;卷积神经网络

Abstract: Based on the machine vision technology, convolutional neural network (CNN) was used to detect and recognize, and verified the cherry defects. The results showed that the recognition accuracy of intact cherry was 99.25%, with the average recognition accuracy of defective cherry of 97.99%, and the recognition speed was 25 per second. Compared with other research methods, this method could accurately detect and identify various types of defects.

Keywords: defect detection; cherry grading; machine vision; convolutional neural network (CNN)

樱桃种植主要分布于美国、加拿大、澳洲、欧洲等地,中国主要有辽宁、山东等^[1-2],因其营养价值丰富而深受人们喜爱。樱桃缺陷主要有裂口、腐烂、双胞胎、刺激生长、未成熟、鸟啄和鼻尖裂口等。目前的分级主要依靠人眼进行识别,会造成可食用缺陷种类的浪费,而基于机器视觉技术的水果自动分级方法可快速、准确、无损地检测

樱桃表面缺陷和判定分类级别。Balestani 等^[3]利用图像处理技术对樱桃进行缺陷检测,其裂口、双胞胎检测率分别为 95%,89%;王昭等^[4]设计了一套基于视觉处理软件的樱桃缺陷检测装置,检测率为 86.8%,但未对缺陷类型进行精确分类。

卷积神经网络(Convolution Neural Network, CNN)是将特征提取与分类融为一体,可直接对原始图像进行模式识别,已成功应用于手写数字识别及车牌字符识别^[5-6],并且都取得了较好的效果。刘云等^[7]改进了一种基于 RGB 彩色分量算术运算的背景分割算法,使用分块策略对苹果缺陷进行检测,检测速度可达 5 个/s,准确率为 97.3%;吴志洋等^[8]采用双网络并行的模型训练方法,结合特征图优化卷积核参数的模型压缩算法,实现了单色布匹瑕疵的快速检测,检测速度为 135 m/min,准确率达 96.99%;郝鑫等^[9]设计了基于块的 CNN 缺陷检测算法,发现晶圆缺陷可分为 10 个缺陷种类,准确率达 97%;许伟栋等^[10]融合了支持向量机方法对马铃薯缺陷进行检测,其运行速度快且准确率达 99.2%;卞国龙等^[11]采用数据增强技术对轮胎 X 射线图像进行了缺陷检测,所需时间短且准确率达 91.3%。

为了避免樱桃可食用缺陷种类的浪费,收获更高的经济效益、更标准的樱桃分级水平、更快的速度以及更高的准确性,试验拟使用图像增强、增广技术对图片进行部分预处理,利用卷积神经网络对樱桃缺陷进行检测识别,为樱桃产后处理提供理论依据。

1 整体模型与算法

1.1 模型框架

典型的卷积神经网络主要由数据输入层、卷积计算

基金项目:国家自然科学基金项目(编号:61601076);国防重点实验室开放基金项目(编号:614240101060217);辽宁省博士启动基金项目(编号:20170520159)

作者简介:裴悦琨(1985—),男,大连大学讲师,研究生导师,博士。E-mail: peiyuekun@dlu.edu.cn

收稿日期:2019-08-06

层、池化层、激励层和全连接层组成。试验设计的基本模型结构如图 1 所示。

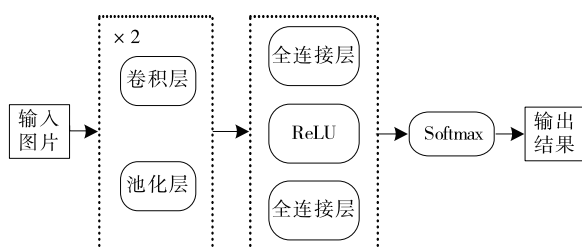


图 1 CNN 基本模型结构

Figure 1 Basic model structure of CNN

1.2 输入层

训练前,需对原始图像进行旋转、缩放、裁剪、归一化等预处理操作,即标准化处理。在图片输入层,对指定大小的生成图片进行样本与标签分类制作,获得神经网络输入的以特征命名的文件,为了方便网络的训练,对输入数据进行批处理。经输入层的图片预处理操作可避免人为干扰,保证输入图片的准确性,提高训练的正确率,得到精确的训练结果。

目前,常用的图片类型主要有 bmp、png、jpg 及 gif 格式。训练时需统一图片格式,彩色图、灰度图不可同时进行训练。试验主要采用 bmp 格式的 RGB 彩色图片,即三通道输入。为加快处理速度与时间,将图片尺寸统一规定为 64×64 。

1.3 卷积层

卷积层是 CNN 网络的核心结构,通过局部感知和参数共享两个原理实现对高维输入数据的降维处理,并且能够自动提取原始数据的优秀特征。卷积层将输入层输出的数据与一组可学习的滤波器进行卷积操作,每个滤波器在输出的图像中产生一个特征映射。试验中采用尺寸为 3×3 的卷积核(filter)。按式(1)计算卷积层特征图大小。

$$\begin{cases} height_{out} = \frac{height_{in} - height_{filter} + 2 \times padding}{stride} + 1; \\ width_{out} = \frac{width_{in} - width_{filter} + 2 \times padding}{stride} + 1, \end{cases} \quad (1)$$

式中:

$height_{in}$ 、 $width_{in}$ ——输入图像尺寸;

$height_{out}$ 、 $width_{out}$ ——输出图片尺寸;

$height_{filter}$ 、 $width_{filter}$ ——卷积核尺寸;

$padding$ ——取 0 或 1;

$stride$ ——卷积滑动步长。

卷积得到的结果加上一个偏置,通过激活函数,作为该层某个神经元的值。试验选用修正线性单元进行激活,不存在饱和问题,其收敛速度大于 S 型生长曲线和双

曲正切,且得到的结果是稀疏的。按式(2)计算卷积层。

$$a_c^i = f[\sum (w_{a_c^{i-1}} + b^i)], \quad (2)$$

式中:

W ——卷积核;

f ——激活函数;

b^i ——偏置。

1.4 池化层

池化层又称子采样层,主要作用是减少参数数量,防止过拟合,降低特征向量维度,并且能够保持局部线性变换的不变性。常见的池化操作主要包括:① 最大池化函数,即对邻域内特征点只取最大值;② 平均池化函数,即对邻域内特征点只求平均值。分别按式(3)、(4)计算最大池化函数、平均池化函数。

$$pooling_{max}(\partial_k) = \max_{i \in \partial_k} a_i, \quad (3)$$

$$pooling_{ave}(\partial_k) = \frac{1}{|\partial_k|} \sum_{i \in \partial_k} a_i, \quad (4)$$

式中:

$pooling_{max}$ ——最大池化函数;

$pooling_{ave}$ ——平均池化函数;

∂_k ——卷积层输出的特征向量划分的多个区域, $k = 1, 2, \dots, K$ 。

1.5 全连接层和输出层

全连接层的输入特征是经过反复提炼过的,因此直接使用原始数据作为输入能够取得更好的效果。试验模型中有两层全连接层,使用 Relu 激活函数,采用 softmax 函数做分类器,输出相应的类别,试验中有两种类别:第一种为正常果和缺陷果;第二种为完好樱桃、裂口樱桃、腐烂樱桃、双胞胎樱桃、刺激生长樱桃及鼻尖裂口樱桃。由于试验中主要是基于缺陷分类,故模型最终的输出根据第二种来定。

训练 CNN,最终形成的训练模型如表 1 所示。

2 算法验证及数据分析

2.1 试验环境

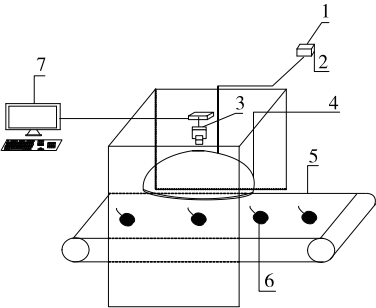
基于机器视觉的樱桃缺陷检测系统由硬件和软件两部分组成,硬件系统用来采集樱桃图像,软件部分对图像进行处理。

图片采集系统主要由照明系统、CMOS 摄像机、吸光背景布、POE 千兆网卡、计算机等组成,如图 2 所示。照明系统包含圆顶光源和控制器,圆顶光源将内嵌的高强度白色 LED 光经圆顶内壁漫反射到被测物体表面,很大程度上避免了樱桃表面的反射和底部阴影,且有利于提高被测物体的轮廓、斑点等细节的清晰度,进而提高检测精度。光源控制器具有同步触发模式,将其与相机触发输出接口相连可以通过相机快门同步触发频闪信号。在整个采集系统的外围,定制了一个立方体的遮光盒,其内

表 1 CNN 模型设计

Table 1 Design of CNN model in this paper

网络架构	具体层	s 参数	步长	输出
Input	数据层			$64 \times 64 \times 3$
Conv1	卷积层 1	$3 \times 3 \times 64 \times 3 + 64$	1	$62 \times 62 \times 64$
Relu1	激活层			
Pool1	3×3 MAX 池化层 1		2	$30 \times 30 \times 64$
Conv2	卷积层 2	$3 \times 3 \times 64 \times 16 + 16$	1	$28 \times 28 \times 16$
Relu2	激活层			
Pool2	3×3 MAX 池化层 2		2	$24 \times 24 \times 16$
FC1	全连接层 1	$24 \times 24 \times 16 \times 128 + 128$		128
Relu3	激活层			
FC2	全连接层 2	$128 \times 2 + 128$		6
Relu4	激活层			
Softmax	损失函数层			



1. 同步光源 2. 控制器 3. CMOS 相机 4. 圆顶光源 5. 传送带 6. 樱桃样本 7. 电脑

图 2 硬件框架图

Figure 2 Hardware framework diagram

壁由吸光背景布覆盖。该外壳能有效屏蔽外界环境光的影响,减少内部反射,使得采集的图片处于同一光照条件下。CMOS 相机、镜头及网卡均为主流的工业级产品。

试验使用 TensorFlow 框架,Python3.7,软件平台采用 Windows10 系统、Pycharm,硬件环境为 4 GB 内存, intel(R)Core(TM) i5-6500 CPU @ 3.20GHzHP LV2011 的计算机。

2.2 试验设计

每个樱桃在采集图像的过程中,根据不同角度采集 6 张图片,以便能覆盖樱桃的全部表面信息。为了扩大样本数据集,先使用旋转、缩放、按比例裁剪以及镜像等 Matlab 操作方法扩大样本数据,在此基础上,建立樱桃样本库,样本数量 15 万张,用样本库进行模型训练。图片采集平台统一规格 2 046×1 080,格式为 bmp,为了较大地缩短训练时间,加快训练速度,所有照片在预处理时都被统一成 64×64 尺寸。具体的流程框架以及样本示例如图 3、4 所示。

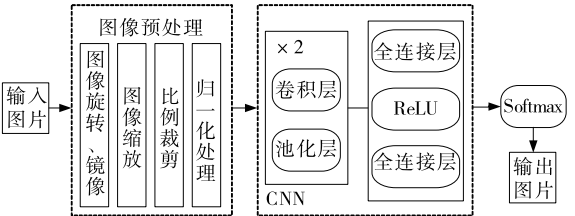
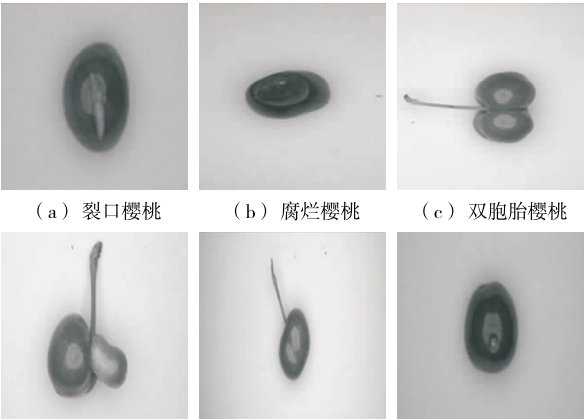


图 3 流程框图

Figure 3 Flow chart



(a) 裂口樱桃 (b) 腐烂樱桃 (c) 双胞胎樱桃 (d) 刺激生长樱桃 (e) 完好樱桃 (f) 鼻尖裂口樱桃

图 4 样本示例

Table 4 Sample examples

试验模型采用 AdamOptimizer 法来最小化损失函数,并对其进行优化。为了避免训练过程中产生过拟合现象,采用随机失活方法,虽然时间增加了近 3 倍,但防过拟合效果很好^[12]。经过多次模型调试,最终将模型训练的批次大小设为 50,即每次的输入样本数为 50;队列容量设为 300,即一个队列的最大输入样本数为 300;训练总次数设为 3 000,即进行 3 000 次迭代训练,每隔 100 步

输出一次训练结果;基础学习率设为 0.001。

试验将 6 种预处理好的分类样本各 25 000 张进行训练学习。首先,采用人工标记的方法对训练样本进行标记,并用数字表示樱桃样本(0 代表裂口樱桃,1 代表腐烂樱桃,2 代表双胞胎樱桃,3 代表刺激生长樱桃,4 代表完好樱桃,5 代表鼻尖裂口樱桃)。这种利用一组已知分类样本来调整分类器参数以满足所需性能的过程即监督学习。经过多次训练,最终得到一个最优模型。对于每一种分类情况,由于其正确率的计算相同,以完好樱桃为例,将完好樱桃用正常果表示,将有缺陷樱桃用缺陷果表示,如果将正常果识别为正常果表示为正预测的正类数,用 TP 表示;将缺陷果识别为缺陷果表示为正预测的负类数,用 TN 表示;将正常果识别为缺陷果表示为正预测的负类数,用 FN 表示;将缺陷果识别为正常果表示为负预测的正类数,用 FP 表示。按式(5)计算正确率(acc)。

$$acc = \frac{TP + TN}{TP + TN + FN + FP} \quad (5)$$

由图 5 可知,整体训练准确率随训练步数的增加逐渐提高;当训练次数为 2 300 步时,整体损失逐渐趋于 0,樱桃识别准确率达 100%。结果表明,试验设计的 6 层 CNN 结构能在短时间内有效提取樱桃图像的特征,使训练精度达到较高水平。

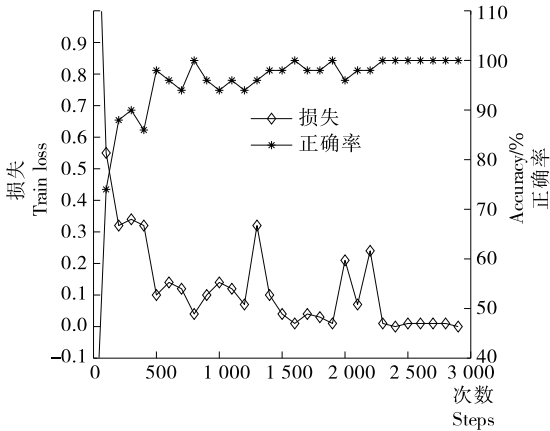


图 5 损失值及正确率随迭代次数的变化趋势图
 Figure 5 The trend of loss value and accuracy with iteration times

2.3 试验结果分析

为了检测使用 64×64 的樱桃图像进行训练是否会对试验结果产生一定影响,在试验测试时选取 720 张樱桃原图进行测试,每个分类各 120 张。由表 2 可知,完好、双胞胎、鼻尖裂口及腐烂樱桃测试结果均正确;裂口樱桃有 3 个被识别错误,刺激生长樱桃有 3 个被识别错误;完好、双胞胎、鼻尖裂口和腐烂樱桃识别正确率为 100%,裂口、刺激生长樱桃识别正确率为 97.5%,处理速度可达 25 个/s,满足了实时性要求。通过查看训练样本可知,刺

表 2 识别正确率比较

Table 2 The correct rate of classification according to the number of recognition

类别	按识别个数计			试验方法 正确率/%
	正确数	错误数	正确率/%	
完好樱桃	120	0	100.0	99.52
双胞胎樱桃	120	0	100.0	97.98
鼻尖裂口樱桃	120	0	100.0	99.99
刺激生长樱桃	117	3	97.5	96.72
裂口樱桃	117	3	97.5	95.60
腐烂樱桃	120	0	100.0	99.64

激生长樱桃被识别为双胞胎、裂口和完好樱桃是因为图片在人为标记过程中出现了分类错误。说明使用预处理后的图像进行训练未对结果产生影响。

2.4 与其他方法的对比

根据樱桃的实际售卖情况可知,裂口、双胞胎、刺激生长及鼻尖裂口樱桃是可以进行售卖的,且裂口、鼻尖裂口樱桃甜度更高,很多人更喜欢食用这两种樱桃,故试验将有缺陷的樱桃划分为裂口、鼻尖裂口、刺激生长、腐烂、双胞胎 5 种类型,同时对完好樱桃进行了精确识别。由表 3 可知,试验方法优于其他方法,有较高的精度及较快的识别速度。

表 3 缺陷种类识别结果对比

Table 3 Comparison of defect type recognition results

方法	缺陷类型	正确率/%	识别速度/ s^{-1}
文献[4]的方法	裂口樱桃	95.00	/
	双胞胎樱桃	89.00	/
文献[5]的方法	正常果	90.90	/
	缺陷果	82.60	/
试验方法	裂口樱桃	97.50	25
	双胞胎樱桃	97.80	25
	完好樱桃	99.52	25
	正常果	99.00	25

3 结论

试验将卷积神经网络应用于樱桃的缺陷检测上,实现了缺陷种类的精确分类,相比其他方法,分类效果得到了明显提升,该方法可以准确高效地识别出正常果和缺陷果,也能精确地细分 5 种缺陷果。根据识别个数分类可知,正常果识别准确率为 100%,缺陷果识别准确率为 99%;根据相似程度判断可知,正常果识别准确率为 99.25%,缺陷果识别准确率为 97.99%,识别速度达 25 个/s。试验表明,该算法实时性好、识别率高,且通过

(下转第 226 页)