

王佳盛, 陈燕, 曾泽钦, 等. 基于全卷积神经网络的荔枝表皮缺陷提取[J]. 华南农业大学学报, 2018, 39(6): 104-110.

基于全卷积神经网络的荔枝表皮缺陷提取

王佳盛, 陈燕, 曾泽钦, 李嘉威, 刘威威, 邹湘军
(华南农业大学 工程学院/南方农业机械与装备关键技术教育部重点实验室, 广东 广州 510642)

摘要:【目的】增强荔枝表皮缺陷提取效果, 满足其品质检测分级准确性要求。【方法】采用 Tensorflow 框架构建基于 AlexNet 的全卷积神经网络 AlexNet-FCN, 以 ReLU 为激活函数, Max-pooling 为下采样方法, Softmax 回归分类器的损失函数作为优化目标, 建立荔枝表皮缺陷提取的全卷积神经网络模型, 并用批量随机梯度下降法对模型进行优化。【结果】模型收敛后在验证集上裂果交并比 (IoU_d) 为 0.83, 褐变交并比 (IoU_b) 为 0.60, 褐变与裂果的总体交并比 (IoU_a) 为 0.68; 与利用线性 SVM、朴素贝叶斯分类器缺陷提取效果相比, 该模型的特征提取能力显著提高。【结论】全卷积神经网络在水果表面缺陷提取中具有良好的应用前景。

关键词: 荔枝; 图像处理; 缺陷提取; 深度学习; 全卷积神经网络; 品质检测
中图分类号: S24; TP391.41 文献标志码: A 文章编号: 1001-411X(2018)06-0104-07

Extraction of litchi fruit pericarp defect based on
a fully convolutional neural network

WANG Jiasheng, CHEN Yan, ZENG Zeqin, LI Jiawei, LIU Weiwei, ZOU Xiangjun
(College of Engineering, South China Agricultural University/Key Laboratory of Key Technology on Agricultural Machine and Equipment, Ministry of Education, Guangzhou 510642, China)

Abstract: 【Objective】 To enhance the effects of litchi fruit pericarp defect extraction and satisfy the accuracy requirements of quality detection and classification. 【Method】 A fully convolutional neural network was built up based on AlexNet (AlexNet-FCN) using Tensorflow framework, with ReLU as the activation function, Max-pooling as the down-sampling method and loss function of Softmax regression classifier as the optimization target. Mini-batch stochastic gradient descent (Mini-batch SGD) was used to optimize the model. 【Result】 When the model was converged, the intersection-over-union of dehiscent area (IoU_d) of litchi fruit cracking was 0.83 for the validation set, the intersection-over-union of brown area (IoU_b) was 0.60, and the intersection-over-union of both dehiscent and brown area (IoU_a) was 0.68. Compared with linear-support vector machine (SVM) and Naïve Bayes classifier, AlexNet-FCN had a stronger defect extraction ability. 【Conclusion】 Fully convolutional networks (FCN) have a good prospect for application of fruit pericarp defect extraction.

Key words: litchi; image processing; defect extraction; deep learning; fully convolutional neural network; quality detection

荔枝的品质检测分级主要采用人工观察或机械分级方法, 存在不够客观准确、效率低下或分级指标单一、容易损伤荔枝等缺点, 而采用基于机器视觉的图像处理技术对荔枝进行自动分级可消除这些弊端^[1-3]。近年来, 国内外许多研究者对基于机器视觉的图像处理技术应用在水果品质的检测进行了大量研究, 并取得较大的进展^[4-6]。采用图像处理技术检测水果品质时, 缺陷提取是其关键步骤, 缺陷区域的提取效果影响缺陷的分析, 从而影响水果品质检测分级的准确性。

果蔬表面的缺陷区域与正常区域通常存在着颜色差异, 有研究提出了基于颜色特征的缺陷提取算法, 赵娟等^[7]利用大津法对苹果图像做阈值分割, 从而提取缺陷区域; 李江波等^[8]提出了基于照度-反射模型的脐橙表面缺陷提取方法。但这些算法需要设定阈值, 自适应性不高。有些水果缺陷区域与正常区域有明显的边缘, 因此有学者提出了基于边缘信息的缺陷提取算法, Miller 等^[9]用基于梯度的高斯滤波器来检测桃子的缺陷边缘; 杜永忠等^[6]用基于 Sobel 与分数阶微分的边缘检测算法提取了圣女果的表面缺陷, 这些算法在表皮光滑的水果上取得了较好的效果。基于无监督学习的水果缺陷检测也有相关研究, 田杰等^[10]提出了基于主成分分析 (Principal component analysis, PCA) 和高斯混合模型 (Gaussian mixture model, GMM) 的小麦病害区域分割; 朱伟华等^[10]提出了在 HSL 颜色模型下的模糊聚类算法对西红柿缺陷进行分割。此外, 还有基于贝叶斯分类器^[11]、SVM^[12]等传统监督学习的缺陷提取方法, 根据像素值将像素点分为“正常”与“缺陷”2 类, 从而实现缺陷分割。由于荔枝表皮布满鳞斑状突起, 上述的算法不能很好地应用于荔枝的表皮缺陷提取, 基于机器视觉的荔枝品质检测的相关研究较少。

卷积神经网络 (Convolutional neural network, CNN) 直接由数据本身驱动特征学习, 对特征具有很强的表达能力。CNN 已成功应用于手写数字识别、人脸识别等研究领域^[13-16], 在农作物的识别分类上也取得了一定的进展^[17-19]。但 CNN 主要应用于图像的识别, 而全卷积神经网络 (Fully convolutional neural network, FCN)^[20]能进行图像的像素级分类, 实现图像分割, 但目前关于其在水果缺陷提取应用的研究较少。本研究采用 FCN 对荔枝表皮缺陷进行提取, 构建基于 AlexNet^[21]的全卷积神经网络 AlexNet-FCN, 优化网络参数, 建立荔枝缺陷提取模型, 通过大量样本进行模型训练, 最终实现荔枝表

<http://xuebao.scau.edu.cn>

皮缺陷的快速准确提取, 为保证荔枝品质检测分级的准确性提供科学依据, 以便后续结合文献^[22-23]实现荔枝品质的检测与分级。

1 材料与方法

1.1 硬件设备

本研究处理平台为台式计算机, 处理器为 Intel(R) Core(TM) i5-4590, 主频为 3.3 GHz, 8 G 内存, GTX1060, 4 G 显卡, 运行系统为 Windows 7 64 位, IDE 为 Pycharm2017, 配置 Anaconda3+TensorFlow-GPU1.4.0+CUDA8.0+cuDNN6.0。图像采集设备为维视 MV-EM130C 工业相机。

1.2 图像采集

荔枝于 2017 年 6 月至 2017 年 8 月采集自广东省广州市增城区和从化区, 品种为“桂味”和“黑叶”。采摘后 3 h 内用维视 MV-EM130C 工业相机进行图像采集, 采集区域布置纯色背景以便荔枝图像分割。通过 2 块互为 60° 的平面镜反射, 每个荔枝可以获得 3 个角度的图像。图像保存为 JPEG 格式, 分辨率为 1 280×960, 采集的荔枝图像如图 1 所示。

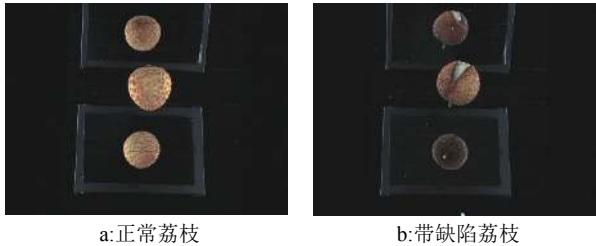


图 1 工业相机采集的荔枝图像
Fig. 1 Litchi figures captured by industrial camera

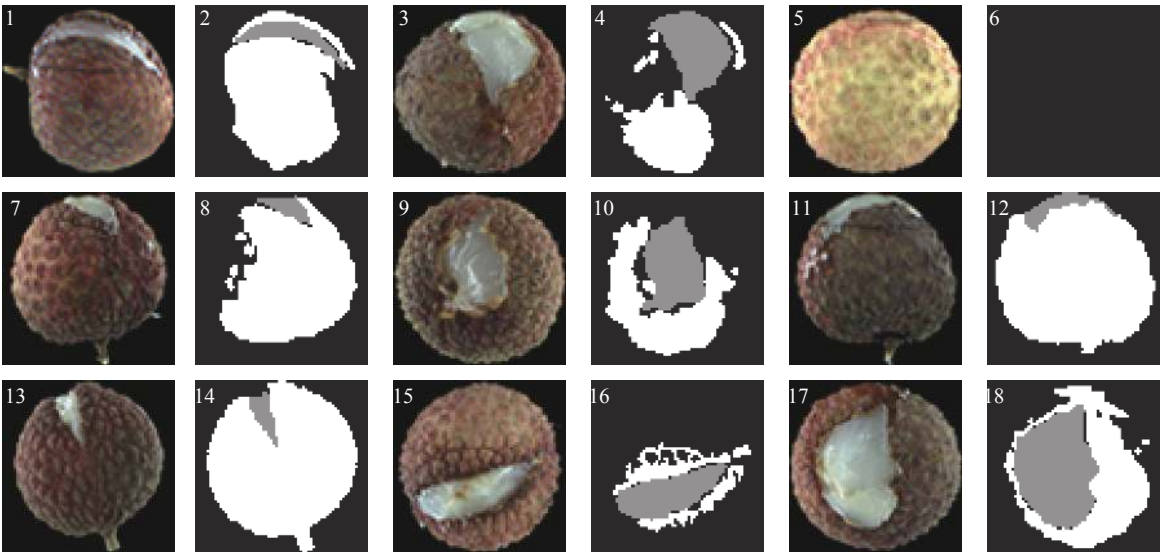
1.3 样本数据集

通过图像分割提取一张图像中的 3 个荔枝, 并分别保存为试验用的样本。对于每个荔枝图像, 手工制作其样本标签, 荔枝分割结果与其对应的标签如图 2 所示。最终试验数据集由 1 068 张图像以及对应的标签图像组成, 在其中随机抽取 70% 作为训练集 (748 张), 20% 作为验证集 (214 张), 剩下 10% 作为测试集 (106 张)。

模型训练完成后, 使用测试集对模型进行测试评估。以人工提取的结果为准, 分析训练模型结果。同时与 SVM、朴素贝叶斯分类方法的缺陷提取结果进行对比。

1.4 全卷积神经网络的构建

本研究采用 Tensorflow 框架构建了基于 AlexNet 的全卷积神经网络 AlexNet-FCN。网络的



1、3、5、7、9、11、13、15 和 17 表示部分试验样本, 2、4、6、8、10、12、14、16、18 分别表示 1、3、5、7、9、11、13、15、17 对应的标签

图 2 试验样本及其标签
Fig. 2 Samples and annotations

卷积层、池化层参照 AlexNet 建立, 用卷积层代替原 AlexNet 中的全连接层, 最后做反卷积进行上采样, 将图像恢复到原图大小, 以获取缺陷的提取结果。AlexNet-FCN 主要网络结构包括卷积层、池化层和反卷积层。

卷积核的尺寸与数量是影响 AlexNet-FCN 模型特征表达能力的重要因素。卷积层的计算如式 (1) 所示。

$$x_j^l = f \left(\sum_{i \in M_j} x_i^{l-1} w_{ij}^l + b_j^l \right), \tag{1}$$

式中, x_j^l 表示 l 层的第 j 个特征图; x_i^{l-1} 表示 $l-1$ 层的第 i 个特征图; $f(\cdot)$ 表示激励函数; M_j 表示输入的特征图; w_{ij}^l 表示第 l 层第 j 个滤波器中的第 i 个卷积核; b_j^l 表示特征的偏置项。

池化层是对特征图的下采样, 一般有平均池化和最大池化 2 种方式, 本研究使用最大池化进行下采样。池化层的计算见式 (2)。

$$x_{m,n}^{l,j} = \text{Max} \left(D_{w \times h}^{m \times s, j \times s} \right), \tag{2}$$

式中, $x_{m,n}^{l,j}$ 表示下采样后第 l 层的第 j 个特征图中, 第 m 行第 n 列的值; $\text{Max}(\cdot)$ 表示取范围内的最大值; $D_{w \times h}^{m \times s, j \times s}$ 表示下采样前第 l 层的第 j 个特征图中, 以 $(m \times s, j \times s)$ 为起始, $(m \times s + w, j \times s + h)$ 为止的范围; s 表示步长; $w \times h$ 表示核的大小。

反卷积^[24]的作用是上采样, 通过多层反卷积将 CNN 结果还原至原图大小, 实现特征可视化。反卷积的大小由反卷积核 (k) 大小和滑动步长 (s) 决定, 计算如式 (3) 所示。

$$S_o = (S_i - 1) \times s + k, \tag{3}$$

式中, S_i 表示输入图像的尺寸; S_o 表示输出图像的尺寸; k 表示卷积核尺寸; s 表示步长。

反卷积过程的具体计算操作如下: 对输入图像的每个像素点进行 1 次全卷积, 即可得到 $S_i \times S_i$ 个 $k \times k$ 特征图; 将特征图进行融合, 根据输入图像像素点将对应的特征图进行排布, 每隔 s 个像素点进行融合, 重叠部分进行相加。该过程如图 3 所示。

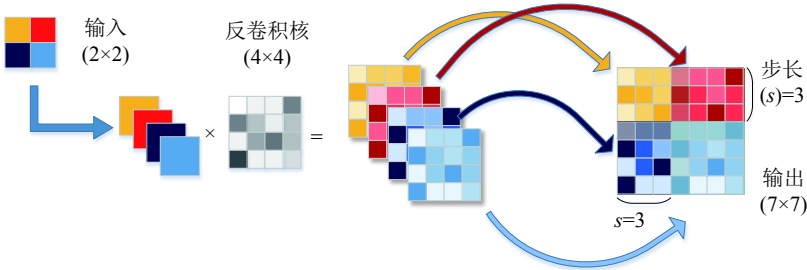


图 3 反卷积计算过程
Fig. 3 Computational procedure of deconvolution

1.5 基于全卷积神经网络的荔枝表皮缺陷提取

1.5.1 输入图像归一化 为提高收敛速度, 输入网络的图像需进行像素值的均值归一化, 采用去均值标准化方法, 具体计算公式见式 (4)。

$$X_n = \frac{X - \mu}{255}, \tag{4}$$

式中, X 表示图像任意一点的像素值; μ 表示图像各通道的像素均值; X_n 表示均值归一化结果。

1.5.2 确定模型 若将 AlexNet 模型直接引入, 将不能进行端到端的特征提取工作。因此将 AlexNet 的全连接层用卷积层代替。尽管 FCN 没有尺寸统一的要求, 但考虑到网络构建的便捷性以及试验样本与 AlexNet 原始网络的差异, 本研究将采用统一的输入尺寸。将样本图像通过插值缩放变化为 227×227 的矩阵, 而标签图像通过插值变化为 $1 \times 227 \times 227$ 。标签图像必须为只有标记 0、1 和 2, 其中 1 和 2 分别代表裂果区域和褐变区域, 0 代表剩下的区域。

AlexNet 全连接部分的结构为 $6 \times 6 \times 256C-4\ 096-4\ 096-1\ 000$, 该结构应用于 1 000 类的分类问题。该结构没有上采样层, 从而无法获得缺陷提取结果。采用 3 次卷积操作代替 AlexNet 中的 3 个全连接层, 3 个卷积核大小分别为 6×6 、 1×1 和 1×1 , 不进行边缘扩充, 最终得到网络结构为 $6 \times 6 \times 256-1 \times 4\ 096C-1 \times 4\ 096C-1 \times 2C$ 的 3 个卷积层。最后, 对 $1 \times 2C$ 做 3 次反卷积, 将结果恢复到 $3 \times 227 \times 227$, 最终实现端到端的像素级分类, 完成缺陷特征提取。

激活函数采用 ReLU。池化层采用最大池化采

样方法, 可以减小因邻域大小受限造成的估计值方差。损失函数采用交叉熵 (对应于 Softmax 逻辑回归分类器), 采用批量随机梯度下降法进行训练。

由于神经网络的计算过程中, 下一层的输入是上一层非线性映射的输出, 因此会造成各层的输入分布不同, 从而导致收敛速度变慢, 甚至会影响精度。这个现象称为 Internal covariate shift。为了解决这一问题, 本研究引入批次正则化 (Batch normalization, BN) 法, 通过 BN, 可以加快网络收敛速度。将输出按照同一批次的特征数值进行去均值归一化。具体公式如下所示。

$$\mu_B = \frac{1}{m} \sum_{i=1}^m x_i, \tag{5}$$

$$\sigma_B^2 = \frac{1}{m} \sum_{i=1}^m (x_i - \mu_B)^2, \tag{6}$$

$$\hat{x}_i = \frac{x_i - \mu_B}{\sqrt{\sigma_B^2 + \varepsilon}}, \tag{7}$$

$$y_i = \gamma \hat{x}_i + \beta \equiv \text{BN}_{\gamma, \beta}(x_i), \tag{8}$$

式中, x_i 表示输入值; m 表示批量化的数目; γ 、 β 表示学习参数, 通过训练网络确定; μ_B 表示批次输入值的均值; σ_B^2 表示批次输入值的方差; $\hat{x}_i$ 表示规范化结果; ε 表示微小常量, 保证数值稳定性; $\text{BN}_{\gamma, \beta}(x_i)$ 表示 BN 的输出。

最终确立 AlexNet-FCN 网络模型结构为 $227 \times 227 \times 3-96C-96S-256C-256S-384C-384C-256C-256S-4\ 096C-4\ 096C-2C-256D-96D-227 \times 227 \times 3$, 具体结构如图 4 所示。

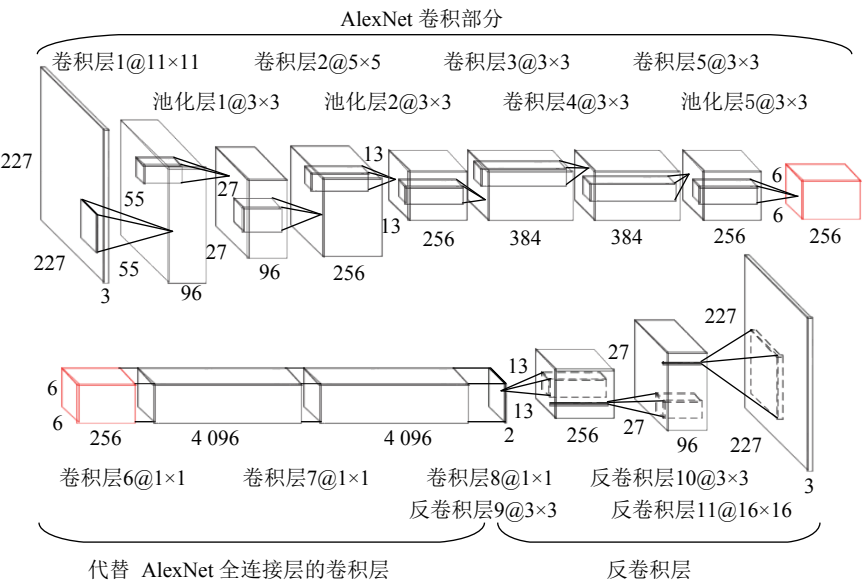


图 4 AlexNet-FCN 结构图

Fig. 4 Structure diagram of AlexNet-FCN

2 结果与分析

2.1 特征可视化分析

按照图 4 所示的网络结构图,使用训练好的模型对荔枝果实样本进行表皮缺陷提取,图 5 为输入荔枝图像前 3 层卷积的部分特征图输出结果。从

图 5 中可以观察到从第 1 层卷积到第 3 层卷积,模型学习的特征从整体到局部,逐渐细化。由图 5 可知,神经网络卷积层能够很好地提取荔枝的特征,说明本研究中运用的 AlexNet-FCN 网络能够很好的达到提取荔枝表皮缺陷特征的效果。

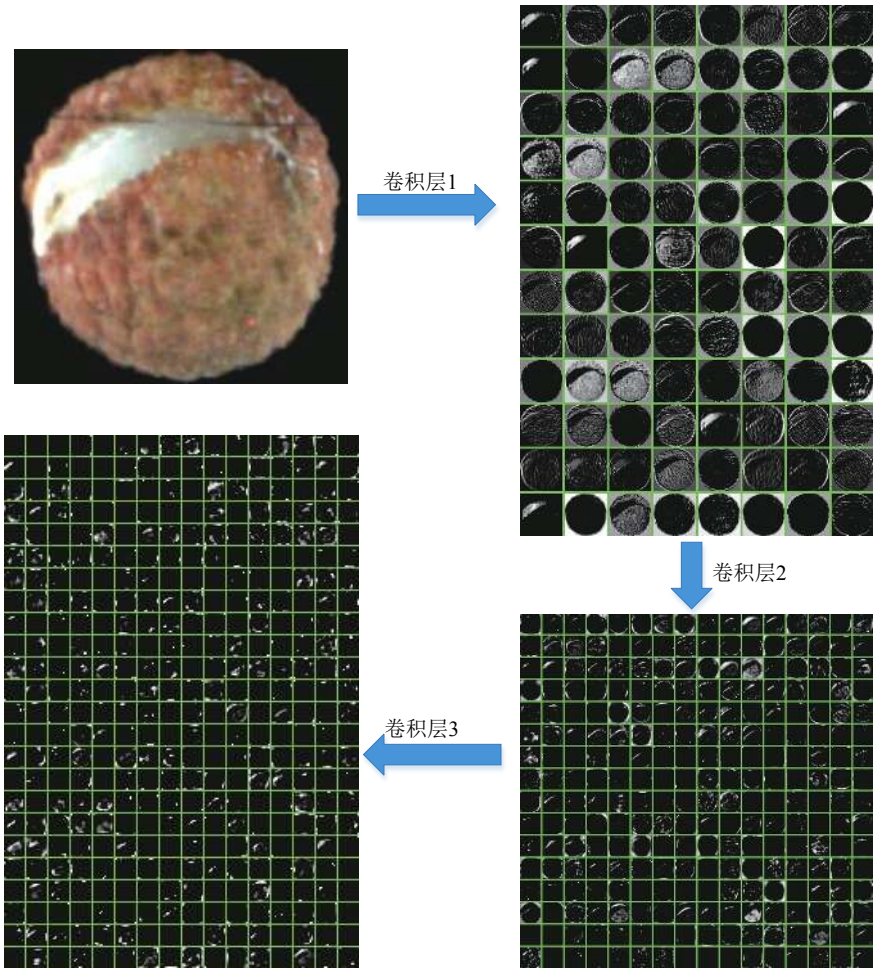


图 5 模型前 3 层卷积处理结果

Fig. 5 Results of the first three convolutional layers of AlexNet-FCN

2.2 训练过程分析

使用训练样本对上述 AlexNet-FCN 模型进行训练。采用均值为 0, 方差为 0.001 的高斯分布进行初始化。设定最大训练步数为 6 000 步,批处理的批尺寸为 24,初始学习率为 0.000 1,学习率衰减因子为 0.9,其损失函数曲线如图 6 所示。由图 6 可知,模型在训练的前 1 000 步损失函数下降迅速,模型的收敛速度较快;当训练步数到达 5 500 步后训练集的损失函数趋于稳定,模型收敛,模型运行到 6 000 步后停止计算。

参照 Long 等^[20]的方法,采用交并比 (Intersection-over-union, IoU) 来验证图像分割效果。交并比是指预测结果与标记结果的交集与并集之比,

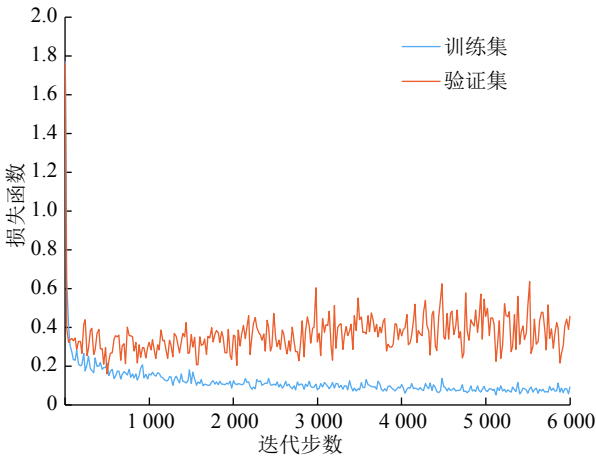


图 6 损失函数曲线

Fig. 6 Curves of loss function

反映了 2 个区域的重叠程度。IoU 求解方法如式 (9) 所示。

$$\text{IoU}_c = \frac{\sum_1^n (S_p \cap S_g)}{\sum_1^n (S_p \cup S_g)}, \quad (9)$$

式中, IoU_c 表示类别 c 的 IoU 得分, n 为验证集样本个数, S_p 表示类别 c 的预测区域面积, S_g 表示类别人工标记标签的区域面积。为方便描述, 本研究采用 IoU_d 表示裂果交并比, 用 IoU_b 表示褐变交并比, 用 IoU_a 表示褐变和裂果的总体交并比。

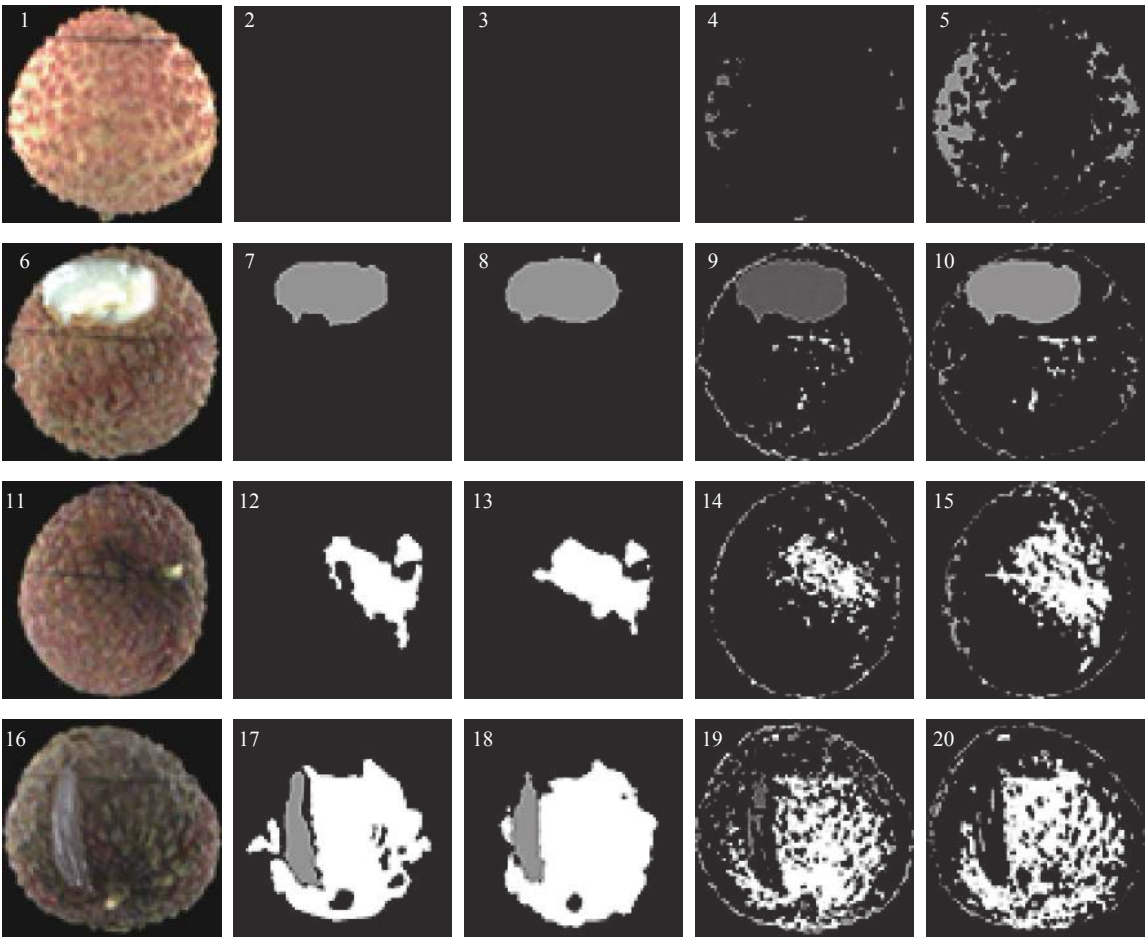
求验证集所有图像的各个评估项目平均交并比得分, 结果如下: $\text{IoU}_d=0.83$, $\text{IoU}_b=0.60$, $\text{IoU}_a=0.68$ 。

2.3 对比分析

因荔枝表皮的缺陷提取相关文献较少, 为验证算法性能, 借鉴其他水果的缺陷提取算法, 运用基

于朴素贝叶斯分类器和线性 SVM 分类器的监督学习缺陷提取算法与本研究的 AlexNet-FCN 模型进行对比分析。在验证集中选取无缺陷样本、裂果缺陷样本、褐变缺陷样本以及裂果和褐变缺陷样本这 4 种典型样本进行对比分析, 图 7 是集中算法的分割效果对比, 表 1 是几种算法总体交并比的结果。由图 7 和表 1 可知, 当缺陷区域与正常区域的像素值差异较大时, 本文模型以及上述 2 种算法均能提取缺陷区域, 但本文模型的 IoU_a 得分明显高于 2 种对比算法, 误识别与漏识别较少, 提取效果更优。而当缺陷区域与正常区域的像素值差异较小时, 本文模型也能较好地分割缺陷区域, 优势明显, 而 2 种对比算法的分割能力略显不足。

AlexNet-FCN 通过卷积学习可得到更加丰富的特征。此外, 卷积和池化操作使得像素分类不仅依靠本身像素值, 还考虑了该点感受野区域内的像素值。因此, 其分类结果优于常规方法。



1~5 表示无缺陷样本; 6~10 表示裂果样本; 11~15 表示褐变样本; 16~20 表示裂果与褐变融合样本; 1、6、11 和 16 表示验证集样本; 2、7、12 和 17 表示样本所对应的标签; 3、8、13 和 18 表示 AlexNet-FCN 模型的提取效果; 4、9、14 和 19 表示线性 SVM 的提取效果; 5、10、15 和 20 表示朴素贝叶斯分类器的提取效果

图 7 几种算法与 AlexNet-FCN 模型的缺陷提取效果对比

Fig. 7 Comparisons of the defect extraction effects among AlexNet-FCN and several other methods

表 1 几种算法与 AlexNet-FCN 模型的总体交并比
Table 1 IoU_a of AlexNet-FCN and several other methods

样本类型	AlexNet-FCN	线性 SVM	朴素贝叶斯分类器
无缺陷	1.00	0.00	0.00
裂果	0.91	0.68	0.81
褐变	0.70	0.38	0.46
裂果和褐变	0.78	0.44	0.55

3 结论

针对荔枝果皮布满斑状突起, 基于颜色以及纹理的表皮缺陷特征提取效果不佳的状况, 本研究提出基于全卷积神经网络的荔枝表皮缺陷提取方法, 构建了 AlexNet-FCN 模型。通过验证试验表明该模型能够自动学习荔枝特征, 可满足荔枝表皮缺陷提取要求。特别是当缺陷区域与正常区域的像素值差异较小时, 模型表现明显优于常规算法。模型收敛后在验证集上裂果交并比 (IoU_d) 为 0.83, 褐变交并比 IoU_b 为 0.60, 褐变与裂果的总体交并比 IoU_a 为 0.68。本研究对荔枝裂果与褐变缺陷做了相关试验与结果分析, 下一步将改进网络结构, 扩充样本, 构建多缺陷特征提取模型, 提高网络的多类缺陷提取能力, 以满足荔枝品质检测分级的实际要求。

参考文献:

[1] 杨松夏, 陆华忠, 吕恩利, 等. 球形水果机械式大小分级技术研究现状[J]. 农机化研究, 2012, 34(8): 235-239.

[2] 赵茂程, 侯文军. 我国基于机器视觉的水果自动分级技术及研究进展[J]. 包装与食品机械, 2007(5): 5-8.

[3] 饶秀勤. 基于机器视觉的水果品质实时检测与分级生产线的关键技术研究[D]. 杭州: 浙江大学, 2007.

[4] CEN H Y, LU R F, ZHU Q B, et al. Nondestructive detection of chilling injury in cucumber fruit using hyperspectral imaging with feature selection and supervised classification[J]. Postharvest Biol Tec, 2016, 111: 352-361.

[5] 熊俊涛, 邹湘军, 刘念, 等. 基于机器视觉的荔枝果实采摘时品质检测技术[J]. 农业机械学报, 2014, 45(7): 54-60.

[6] 杜永忠, 平雪良, 何佳唯. 圣女果表面缺陷检测与分级系统研究[J]. 农业机械学报, 2013, 44(S1): 194-199.

[7] 赵娟, 彭彦昆, SAGAR D, 等. 基于机器视觉的苹果外观缺陷在线检测[J]. 农业机械学报, 2013, 44(S1): 260-263.

[8] 李江波, 饶秀勤, 应义斌. 基于照度-反射模型的脐橙表面缺陷检测[J]. 农业工程学报, 2011, 27(7): 338-342.

[9] MILLER B. Peach defect detection with machine

vision[J]. Trans Asae, 1991, 34(6): 2588-2597.

[10] 田杰, 韩冬, 胡秋霞, 等. 基于 PCA 和高斯混合模型的小麦病害彩色图像分割[J]. 农业机械学报, 2014, 45(7): 267-271.

[11] LEEMANS V, MAGEIN H, DESTAIN M F. Defect segmentation on ‘Jonagold’ apples using colour vision and a Bayesian classification method[J]. Comput Electron Agr, 1999, 23(1): 43-53.

[12] MIZUSHIMA A, LU R. An image segmentation method for apple sorting and grading using support vector machine and Otsu’s method[J]. Comput Electron Agr, 2013, 94: 29-37.

[13] YANG W, JIN L, TAO D, et al. Dropsample: A new training method to enhance deep convolutional neural networks for large-scale unconstrained handwritten Chinese character recognition[J]. Pattern Recogn, 2016, 58(4): 190-203.

[14] ALBU R D. Human face recognition using convolutional neural networks[J]. JEEE, 2009(2): 110-113.

[15] RAMAIAH N P, IJJINA E P, MOHAN C K. Illumination invariant face recognition using convolutional neural networks[C]// IEEE International Conference on Signal Processing, US: IEEE Xplore, 2015: 1-4.

[16] SINGH R, OM H. Newborn face recognition using deep convolutional neural network[J]. Multimed Tools Appl, 2017, 76(18): 1-11.

[17] 高震宇, 王安, 刘勇, 等. 基于卷积神经网络的鲜茶叶智能分选系统研究[J]. 农业机械学报, 2017, 48(7): 53-58.

[18] 马晓丹, 关海鸥, 祁广云, 等. 基于改进级联神经网络的大豆叶部病害诊断模型[J]. 农业机械学报, 2017, 48(1): 163-168.

[19] 谭文学, 赵春江, 吴华瑞, 等. 基于弹性动量深度学习神经网络的果体病理图像识别[J]. 农业机械学报, 2015, 46(1): 20-25.

[20] LONG J, SHELHAMER E, DARRELL T. Fully convolutional networks for semantic segmentation[J]. TPAMI, 2014, 39(4): 640-651.

[21] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks[C]// Proceedings of the 15th International Conference on Neural Information Processing Systems. Cambridge: Mit Press, 2012: 1097-1105.

[22] 陈石榕. 国际鲜荔枝标准[J]. 标准化信息, 1997(4): 11-12.

[23] 周先治, 陈阳, 唐建阳, 等. 荔枝褐变损失估计研究[J]. 热带作物学报, 2012, 33(8): 1403-1408.

[24] ZEILER M D, KRISHNAN D, TAYLOR G W, et al. Deconvolutional networks[C]//Computer Vision and Pattern Recognition. US: IEEE Computer Science, 2010: 2528-2535.

【责任编辑 霍 欢, 周志红】