

DOI: 10.13652/j.issn.1003-5788.2021.01.015

# 基于高光谱成像技术的水果表面 农药残留无损检测

## Nondestructive detection of pesticide residues on fruit surface by hyperspectral imaging technology

张 萌<sup>1</sup> 贾世杰<sup>2</sup>ZHANG Meng<sup>1</sup> JIA Shi-jie<sup>2</sup>

(1. 吉林铁道职业技术学院电气工程分院, 吉林 吉林 132200;

2. 大连交通大学电子信息工程学院, 辽宁 大连 116028)

(1. *Electrical Engineering Branch, Jilin Railway Vocational and Technical College, Jilin, Jilin 132200, China;*2. *School of Electronic Information Engineering, Dalian Jiaotong University, Dalian, Liaoning 116028, China)*

**摘要:**在高光谱成像技术的基础上,提出了一种应用于水果表面农药残留的无损检测方法。对采集数据进行预处理和特征提取,通过细菌群体趋药性算法找到最优的最小二乘支持向量机参数,建立农残检测模型,并与最小二乘支持向量机模型进行比较,验证该模型的优越性和准确性。结果表明,基于连续投影法特征波长结合文中检测模型具有最高的检测精度,其准确率达 97.92%。

**关键词:**农药残留;无损检测;高光谱成像技术;细菌群体趋药性算法;最小二乘支持向量机

**Abstract:** Based on the hyperspectral imaging technology, a non-destructive detection method for pesticide residues on fruit surfaces is proposed. By preprocessing the collected data and extract features, finding the optimal least squares support vector machine parameters through the bacterial population chemotaxis algorithm, a pesticide residue detection model was established, which was compared with the least squares support vector machine model to verify the superiority and accuracy the model. The results showed that the detection model had the highest detection accuracy based on the characteristic wavelength of the continuous projection method combined with the detection model, and its accuracy rate was 97.92%, which had certain application value.

**Keywords:** pesticide residues; nondestructive testing; hyperspectral imaging technology; bacterial colony chemotaxis algorithm; least squares support vector machine

基金项目:辽宁省教育厅科学研究项目(编号:JDL2019006)

作者简介:张萌,女,吉林铁道职业技术学院讲师。

通信作者:贾世杰(1969—),男,大连交通大学教授,博士。

E-mail: stamadeng@foxmail.com

收稿日期:2020-08-14

农药为世界农业的发展和粮食可持续供应提供了极大帮助<sup>[1]</sup>。中国作为一个农业大国,每年农药使用量超过 50 万 t,居世界第一,是世界平均水平的 2.5 倍。进入 21 世纪以来,食品安全逐渐成为社会关注的焦点<sup>[2]</sup>。由于农药的广泛使用会直接或间接进入人体,危害人体健康,因此农药残留检测已成为重中之重。

目前,常见的食品农药残留检测方法主要为气相色谱法、液相色谱法、气质联用法、液质联用法等,而基于高光谱成像技术的水果表面农药残留无损检测技术研究较少。杨昌标等<sup>[3]</sup>采用液相色谱—串联质谱分析法快速检测出西红柿中的 18 种残留农药,该方法具有良好的线性关系,且在批量样品的检测中既快速又稳定。薄璐等<sup>[4]</sup>在视觉成像技术的基础上,提出了一种番茄表面农残无损检测方法,该方法具有良好的无损检测性能,提高了番茄表面农残的检测能力。姚云恒等<sup>[5]</sup>提出了一种使用气相色谱—质谱分析技术快速测定苹果梨中 19 种残留农药的方法,该检测方法简便、准确,适用于苹果梨中 19 种残留农药的筛选和测定。张晶等<sup>[6]</sup>建立了一种气相色谱—质谱分析方法,用于检测蔬菜中的 8 种氨基甲酸酯农残,该检测方法灵敏度高、操作简便、定性和定量分析准确可靠,在植物氨基甲酸酯的检测中准确率较高且稳定。徐洁等<sup>[7]</sup>提出了一种哈密瓜表面农残判别分析方法,并验证了该检测方法的准确性。李增芳等<sup>[8]</sup>提出了一种用高光谱成像技术对赣南脐橙样品农药残留进行无损检测,并验证了该检测方法的准确性。赵曼彤等<sup>[9]</sup>提出了一种利用高光谱成像技术对香梨表面低农残进行检测,通过多元线性回归法建立农残检测模型,并验证了该检测方法的准确性。以上检测技术不仅操作复杂、耗时

长等,甚至还会破坏样品,因此需要寻找更加快速、精确、高效的农残无损检测技术。

试验拟基于高光谱成像技术,提出细菌群体趋药性(BCC)算法和最小二乘支持向量机(LS-SVM)算法相结合建立水果表面农残检测模型,通过对采集数据进行预处理和特征提取,建立农药残留浓度检测模型,并与改进前进行试验对比,验证该方法的优越性和准确性,旨在为水果表面农残的无损检测提供依据。

## 1 系统概述

高光谱成像技术是计算机技术和光学技术的结合。由于其丰富的图像信息和光谱数据,近年来已被应用于医学和农业领域。原始图像所包含的光谱信息不仅仅是特定像素,还有特定波长的图像信息。高光谱成像平台主要由高光谱相机(CCD 相机、光谱仪、镜头)、光源(紫外线灯、卤素灯)、电动平台和控制计算机等组成(见图 1),采用推扫成像原理,高光谱相机使用分束器将宽波混合反射光散射为各种频率的单波长光。

## 2 食品表面农残无损检测

### 2.1 光谱数据预处理

使用标准正态变换(SNV)结合去趋势算法完成数据预处理,用 SNV 校正由粒子散射导致的误差<sup>[10]</sup>,其原理是每个光谱中不同波长的吸收是不同的,因此有必要校正每个原始光谱,并按式(1)进行计算。

$$X_{iSNV} = \frac{x_i - \bar{x}}{\sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}}}, \quad (1)$$

式中:

$\bar{x}$  ——第  $i$  条光谱  $n$  个光谱点平均值;

$X_{iSNV}$  ——第  $i$  条光谱经 SNV 处理后的数据。

去趋势主要解决 SNV 校正光谱数据中基线漂移问题<sup>[11]</sup>,即使用多项式拟合光谱的吸收率和波长以获得趋

势线,用 SNV 校正光谱数据减去趋势线。

### 2.2 特征波长的选择

高光谱图像含有大量光谱数据,由数百个波段构成,并且每个波段都高度相关<sup>[12]</sup>。对整个频带进行分析和建模不仅会影响试验的准确性,而且计算量大大增加,降低了建模的速度<sup>[13]</sup>。故采用主成分分析(PCA)和连续投影算法(SPA)提取光谱数据的特征波长进行比较分析。

2.2.1 PCA 特征提取 PCA 具有降维和降噪的效果。降维后得到一组新的自变量,这些自变量将原始变量的所有信息最大化。处理后的图像如式(2)所示<sup>[14]</sup>。

$$Pc_m = \sum_{i=1}^n \alpha_i I_i, \quad (2)$$

式中:

$Pc_m$  ——第  $m$  个主成分图像;

$I_i$  ——第  $i$  波段对应的图像;

$\alpha_i$  ——第  $i$  波段图像权重系数;

$n$  ——图像的数量。

第一个主成分包含最原始的信息,贡献最大。通常前 10 个主成分图像的累积贡献为原始图像的 99%。因此,只要选择适当的主成分图像,就可以获得高光谱数据的特征波长。

2.2.2 SPA 特征提取 SPA 可以从光谱信息中充分找到冗余信息最低的变量组,减少信息重叠。同时,减少建模中使用的变量数量,从而提高建模的速度和效率<sup>[15]</sup>。SPA 算法的主要原理是:设置  $\mathbf{X}_{n \times m}$  为光谱矩阵,其中  $n$  为样本数, $m$  为光谱波长数, $N$  为需选择的变量数。

(1) 首先选取  $j$  列光谱矩阵  $x_j$ , 初始迭代向量为  $x_{k(0)}$ 。

(2) 将剩余的光谱数据放入集合  $S$  中,  $S = \{j, 1 \leq j \leq J, j \notin [k(0), \dots, k(n-1)]\}$ 。

(3) 计算  $x_j$  在剩余列向量上的投影,如式(3)所示<sup>[16]</sup>。

$$p_{x_j} = x_j - [x_j^T x_{k(n-1)}] x_{k(n-1)} [x_{k(n-1)}^T x_{k(n-1)}]^{-1}, \quad j \in S. \quad (3)$$

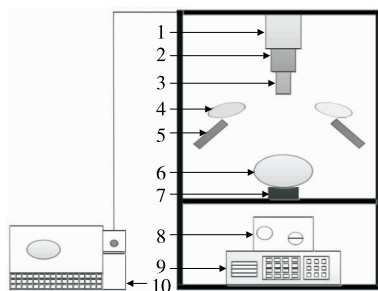
(4) 记  $k(n) = \arg[\max(\|p_{x_j}\|)]$ ,  $j \in S$ , 同时令  $x_j = p_{x_j}$ ,  $j \in S$ 。

(5) 令  $n = n + 1$ , 若  $n < N$ , 跳到 2.2.2(2) 重新计算,最终选定的特征波长为  $\{x_{k(n)} = 0, \dots, N-1\}$ 。

对于每个  $k(0)$  和  $N$ , 在一个周期后进行多元线性分析,获得的最小预测标准偏差为  $k(0)$  和  $N$  对应的最优值。

### 2.3 数学模型

2.3.1 LS-SVM 算法 LS-SVM 解决了 SVM 算法在处理大量数据时训练速度较慢的问题,大大节省了收敛时间<sup>[17]</sup>。LS-SVM 是处理小样本集的最合适算法之一,该算法可以从有限数量的样本中提取信息并预测未知



1. CCD 相机 2. 光谱仪 3. 镜头 4. 紫外灯 5. 卤素灯 6. 试验样本 7. 电动平移台 8. 光源控制箱 9. 电动平移台控制箱 10. 计算机

图 1 高光谱成像采集系统

Figure 1 Hyperspectral imaging acquisition system

数据<sup>[18]</sup>。

LS-SVM 使用非线性映射函数  $\varphi(x)$  来建立方程,并使用拉格朗日算子求解最优化问题。根据 Mercer 条件,映射函数  $\varphi(x)$  和核函数  $K(x_i, x_j)$  的对数函数如式(4)所示<sup>[19]</sup>。

$$\varphi(x_k)^T \varphi(x_1) = K(x_k, x_1), \quad (4)$$

式中:

$x_k$  ——第  $k$  个输入样本。

文中采用 RBF 核函数:

$$K(x_k, x_1) = \exp\left(-\frac{\|x_k - x_1\|^2}{\sigma^2}\right). \quad (5)$$

LS-SVM 算法的函数方程如式(6)所示<sup>[20]</sup>。

$$y(x) = \sum_{k=1}^N \alpha_k K(x, x_k) + b, \quad (6)$$

式中:

$\alpha_k$  ——Lagrange 乘子;

$b$  ——偏差参数。

**2.3.2 BCC-LS-SVM 算法** LS-SVM 算法通常使用 RBF 内核函数。采用两步网格搜索和留一法确定 RBF 内核参数,对 RBF 函数的正则化参数  $\gamma$  和核参数  $\sigma^2$  进行优化<sup>[21]</sup>。 $\gamma$  和  $\sigma^2$  反映了算法的学习能力和泛化能力。

BCC 算法是一种群体智能算法。细菌不仅依靠自身行为产生的信息,而且优化周围细菌运动产生的信息。BCC 算法具有全局搜索功能、收敛速度快、精度高等优点。该方法提高了支持向量机参数选择的效率,避免了人工参数设置的不足,大大缩短了优化时间。

文中提出了基于 BCC 改进的 LS-SVM 算法,通过 BCC 找到最优的  $\gamma$  和  $\sigma^2$ 。BCC-LS-SVM 算法的基本步骤:

(1) 初始化。初始细菌种群、位置、优化参数的限制、初始迭代次数和最大迭代次数等。

(2) 计算参数  $T_0$ 、 $b_1$ 、 $t_c$ , 如式(7)所示<sup>[22]</sup>。

$$\begin{cases} T_0 = \epsilon^{0.30} \times 10^{-1.73} \\ b_1 = T_0 \times (T_0^{-1.54} \times 10^{0.60}) \\ t_c = \left[\frac{b_1}{T_0}\right]^{0.31} \times 10^{1.16} \end{cases}, \quad (7)$$

式中:

$T_0$  ——与计算精度  $\epsilon$  有关的时间参数;

$t_c$  ——与细菌在运动方向上的旋转角度有关的细菌移动时间;

$b_1$  ——与维度无关的参数。

(3) 根据细菌的位置对 LS-SVM 模型进行训练,对初始目标函数值进行评估,记录最大目标函数及其对应的正则化参数  $\gamma$  和核参数  $\sigma^2$ <sup>[23]</sup>。

(4) 通过 BCC 算法找到最佳的  $\gamma$  和  $\sigma^2$ ,为每种细菌计算新位置,并对细菌种群进行更新。

(5) 确定最佳细菌并求得最佳函数值。

(6) 如果前后函数值的差的绝对值小于给定精度或达到最大迭代次数,则执行下一步,否则跳到步骤 2.3.2 (3)。

(7) 使用最佳细菌位置( $\gamma$  和  $\sigma^2$ )作为 LS-SVM 模型中的参数。

BCC-LS-SVM 算法流程图如图 2 所示。

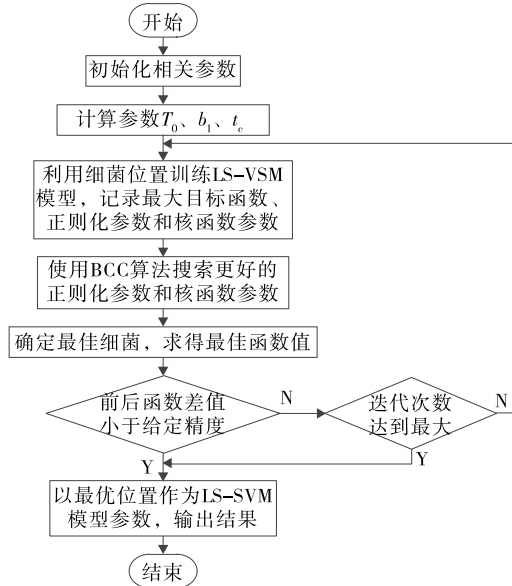


图 2 BCC-LS-SVM 算法流程

Figure 2 BCC-LS-SVM algorithm flow

### 3 结果与分析

#### 3.1 试验设备

CCD 相机:ORCA-R2 型,日本滨松公司;  
光谱仪:V10E-QE 型,芬兰 SPECIM 公司;  
超亮紫外线灯:ML3500 型,美国 SP 公司;  
光纤卤素灯:DCR III 型,150 W,德国肖特公司。

#### 3.2 试验分析

以吡螨灵农药为试验对象,对 80 份水果样品进行洗涤,置于通风处干燥<sup>[24]</sup>。随机分为 4 组,每组 20 份。从每组中随机选择 8 个样本作为训练数据,共选择 32 个样本,其余 48 个样本作为测试集。提取全光谱数据的特征波长后,PCA 提取的特征波长分别为 500,580,680,850,930,980 nm;SPA 提取的特征波长分别为 461,549,600,680,740,830,900,960 nm。建立 LS-SVM 和 BCC-LS-SVM 检测模型,输出数据在样本类别中。水果表面上吡螨灵残留浓度( $V_{\text{吡螨灵}} : V_{\text{水}}$ )可分为 4 类<sup>[25]</sup>:1 : 400,1 : 800,1 : 1 500 以及无吡螨灵残留。

**3.2.1 LS-SVM 模型试验分析** 使用 PCA、SPA 提取的特征波长作为输入数据,建立 LS-SVM 模型判别水果表面农药残留浓度。基于 PCA 特征波长的 LS-SVM 判别

水果表面农药残留浓度如图 3 所示。由图 3 可知,基于 PCA 特征波长的 LS-SVM 模型成功预测了 43 个样本,准确率为 89.58%。

图 4 为基于 SPA 特征波长的 LS-SVM 判别水果表面上的残留农药浓度,其成功预测了 44 个样本,每种类型吡啶螨均存在误判,准确率为 91.67%。

比较图 3、4 可知,基于 PCA 和 SPA 的 LS-SVM 模型检测水果表面农药残留浓度准确度分别为 89.58%, 91.67%,采用 SPA 的模型优于 PCA 的。这是因为采用 SPA 算法挑选出的有效波长能充分代表原始光谱的有效信息,预测效果较好。因此,将 LS-SVM 模型应用于水果表面残留农药检测的连续投影更为有效。

3.2.2 BCC-LS-SVM 模型试验分析 使用 PCA 和 SPA 提取的特征波长作为输入数据,建立 BCC-LS-SVM 模型,用于水果表面农药残留浓度的判别,基于 PCA 特征波长结合改进的最小二乘支持向量机模型(BCC-LS-SVM)检测水果表面农药残留如图 5 所示。由图 5 可知,基于 PCA 的特征波长的 BCC-LS-SVM 成功预测了 45 个样本,准确率为 93.75%。

图 6 为基于 SPA 的特征波长信息的 BCC-LS-SVM 模型检测水果表面上的残留农药浓度。由图 6 可知,该模型成功预测了 47 个样本,准确率为 97.92%。

比较图 5、6 可知,基于 PCA 和 SPA 这两种特征波长提取方法建立的改进的最小二乘支持向量机模型(BCC-LS-SVM)可以检测水果表面上残留的农药浓度,准确率分别为 93.75%,97.92%,SPA 选择特征波长的模型精度高于 PCA 的。因此,BCC-LS-SVM 模型可用于检测水果

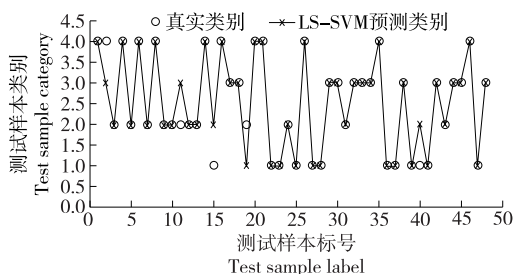


图 3 基于 PCA 的 LS-SVM 模型判别图

Figure 3 PCA based LS-SVM model discriminant graph

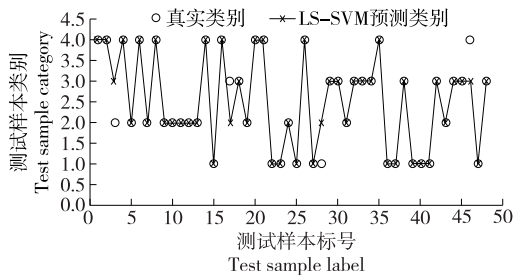


图 4 基于 SPA 的 LS-SVM 模型判别图

Figure 4 SPA based LS-SVM model discriminant graph

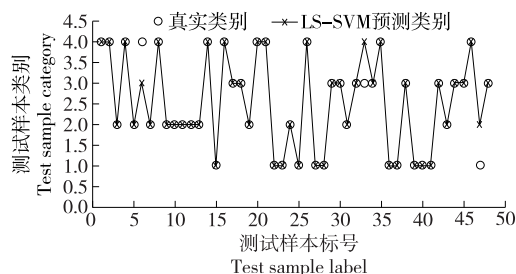


图 5 基于 PCA 的 BCC-LS-SVM 模型判别图

Figure 5 PCA based BCC-LS-SVM model discriminant graph

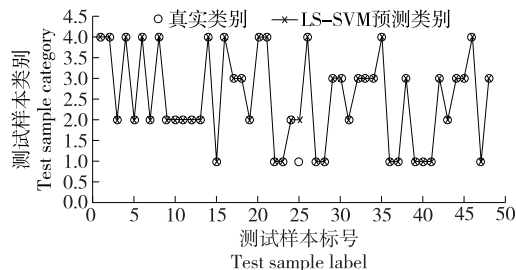


图 6 基于 SPA 的 BCC-LS-SVM 模型判别图

Figure 6 SPA based BCC-LS-SVM model discriminant graph

表面残留的农药浓度,其中连续投影更为有效。

3.2.3 不同模型对比分析 特征波长提取方法分别选择 PCA 和 SPA,并分别使用 LS-SVM 和 BCC-LS-SVM 进行建模,预测模型的结果见表 1。由表 1 可知,基于连续投影法(SPA)的特征波长的检测模型整体精度高于基于主成分分析(PCA)的特征波长的检测模型,说明连续投影法的特征提取效果极好。在 SPA 特征波长下,BCC-LS-SVM 模型的准确率为 97.92%,优于 LS-SVM 模型,是因为 BCC 算法找到最优的正则化参数和核参数,使 BCC-LS-SVM 模型更加精准。

### 4 结论

试验表明,在高光谱成像中,连续投影结合 BCC-LS-SVM 模型可以有效地检测水果表面上残留的吡啶灵浓度,且优于 LS-SVM 模型,通过 BCC 算法找到最优的 LS-SVM 模型参数,将检测准确率从 93.75%提高至 97.92%。该检测方法仍处于起步阶段,后续将不断改进检测模型

表 1 不同建模结果对比

Table 1 Comparison of different modeling results

特征选择	模型	准确率/%
PCA	LS-SVM	89.58
	BCC-LS-SVM	93.75
SPA	LS-SVM	91.67
	BCC-LS-SVM	97.92



的性能,完善基于高光谱成像技术的水果表面残留农药的无损检测方法。

### 参考文献

- [1] 赵利平, 吴德刚. 基于小波与模糊相融合的苹果分级算法[J]. 食品与机械, 2020, 36(4): 142-145.
- [2] 项辉宇, 薛真, 冷崇杰, 等. 基于 Halcon 的苹果品质视觉检测试验研究[J]. 食品与机械, 2016, 32(10): 123-126.
- [3] 杨昌彪, 付春燕, 罗廷武, 等. QuEChERS-LC-MS/MS 快速检测番茄 18 种农药残留[J]. 江苏农业科学, 2020, 48(9): 218-222.
- [4] 薄璐, 王立霞. 基于视觉图像识别的番茄表面农药残留量无损检测方法[J]. 食品与机械, 2019, 35(3): 63-71.
- [5] 姚蕴恒, 白龙律, 武伦鹏, 等. SPE-GC/MS 测定苹果梨中 19 种有机氯农药残留[J]. 食品与机械, 2019, 35(12): 87-95.
- [6] 张晶, 韩建龙, 沈海涛, 等. 气相色谱—串联质谱法检测蔬菜中 8 种氨基甲酸酯类农药残留[J]. 食品安全质量检测学报, 2020, 11(15): 5 045-5 051.
- [7] 徐洁, 杨杰, 孙静涛, 等. 基于高光谱技术的哈密瓜表面农药残留判别分析[J]. 江苏农业科学, 2016, 44(12): 338-340.
- [8] 李增芳, 楚秉泉, 章海亮, 等. 高光谱成像技术无损检测赣南脐橙表面农药残留研究[J]. 光谱学与光谱分析, 2016, 36(12): 4 034-4 038.
- [9] 赵曼彤, 李柏承, 周瑶, 等. 香梨表面低浓度农药残留高光谱检测研究[J]. 光学技术, 2016, 42(5): 408-412, 418.
- [10] HU Jian-jiang, FICHTNER Maximilian, BARICCO Marcello. Preparation of Li-Mg-N-H hydrogen storage materials for an auxiliary power unit[J]. International Journal of Hydrogen Energy, 2017, 42(27): 17 144-17 148.
- [11] LIU Guang-ming, YANG Ming-gan, LANG Guang, et al. Analysis of the heat generation of lithium-ion battery during charging and discharging considering different influencing factors[J]. Journal of Thermal Analysis & Calorimetry, 2014, 116(2): 1 001-1 010.
- [12] ULRIKE K, FRIDOLIN R, ERANDA H, et al. Review: Dynamic models of Li-Ion batteries for diagnosis and operation: A review and perspective[J]. Journal of the Electrochemical Society, 2018, 165(16): 3 656-3 673.
- [13] HOU Kai-yuan, SHAO Guang-hui, WANG Hai-ming, et al. Research on practical power system stability analysis algorithm based on modified SVM[J]. Protection and Control of Modern Power Systems, 2018, 3(2): 2-7.
- [14] GAUTAN J V, PRAJAPATI H B, DABHI V K, et al. Empirical study of job scheduling algorithms in hadoop MapReduce[J]. Cybernetics and Information Technologies, 2017, 21(1): 146-163.
- [15] CAETANOC E F, LIMAA B, PAULINO J O S, et al. A conductor arrangement that overcomes the effective length issue in transmission line grounding[J]. Electric Power Systems Research, 2018, 46(5): 159-162.
- [16] SONG Zi-you, LI Jian-qi, HOU Jun, et al. The optimization of a hybrid energy storage system at subzero temperatures: Energy management strategy design and battery heating requirement analysis[J]. Applied Energy, 2015, 159: 576-588.
- [17] KERESZTES J C, GOODARZI M, SAEYS W. Real-time pixel based early apple bruise detection using short wave infrared hyperspectral imaging in combination with calibration and glare correction techniques [J]. Food Control, 2016, 66(1): 215-226.
- [18] JIA Zhi-wei, WANG Li-jun, ZHANG Jin-chuan, et al. High efficiency, low power-consumption DFB quantum cascade lasers without lateral regrowth [J]. Nanoscale Research Letters, 2017, 12(1): 88-95.
- [19] AZUMAYA C M, DAYS E L, VINSON P N, et al. Screening for AMPA receptor auxiliary subunit specific modulators[J]. PLoS One, 2017, 12(3): 1 523-1 538.
- [20] FAN Cheng-li. Hybrid artificial bee colony algorithm with variable neighborhood search and memory mechanism[J]. Journal of Systems Engineering and Electronics, 2018, 29(2): 405-414.
- [21] 王冉冉, 刘鑫, 尹孟, 等. 面向苹果硬度检测仪的声振信号激励与采集系统设计[J]. 浙江大学学报(农业与生命科学版), 2020, 46(1): 111-118.
- [22] 杨双艳, 杨紫刚, 张四伟, 等. 基于近红外光谱和 PSO-SVM 算法的烟叶自动分级方法[J]. 贵州农业科学, 2018, 46(12): 141-144.
- [23] 于蒙, 李雄, 杨海潮, 等. 基于图像识别的苹果的等级分级研究[J]. 自动化与仪表, 2019, 34(7): 39-43.
- [24] 王阳阳, 黄勋, 陈浩, 等. 基于同态滤波和改进 K-means 的苹果分级算法研究[J]. 食品与机械, 2019, 35(12): 47-51, 112.
- [25] 王立扬, 张瑜, 沈群, 等. 基于改进型 LeNet-5 的苹果自动分级方法[J]. 中国农机化学报, 2020, 41(7): 105-110.

### 信息窗

## 更正启事

本刊 2020 年第 12 期刊发的《复合营养粉对 2 型糖尿病大鼠的辅助降血糖试验》一文,表 5 中胰岛素抵抗指数数据勘误:复合营养粉低剂量组由  $8.998 \pm 2.765$  改为  $9.204 \pm 2.386$ ;复合营养粉中剂量组由  $10.130 \pm 2.939$  改为  $13.149 \pm 11.232$ ;复合营养粉高剂量组由  $11.278 \pm 4.649$  改为  $9.159 \pm 3.025$ ;阳性对照组由  $8.292 \pm 1.605$  改为  $8.998 \pm 2.765$ 。本数据订正不影响文章结果及讨论等内容。

《食品与机械》编辑部

2021 年 1 月