

可见-近红外光谱的鸭梨黑心缺陷在线检测 AdaBoost 集成模型研究

郝勇¹, 王起明¹, 张书敏²

1. 华东交通大学机电与车辆工程学院, 江西 南昌 330013

2. 南昌海关技术中心, 江西 南昌 330038

摘要 黑心病是鸭梨贮藏期间发生的生理病害, 其病变初期表现在内部果核处出现褐色斑块, 而在果实外观上与正常果几乎没有任何差异, 严重影响鸭梨的贮藏时间和品质, 亟需一种快速无损的检测方法为鸭梨质量保驾护航。采用可见-近红外光谱法对鸭梨黑心缺陷进行在线检测和识别, 结合平滑(Smoothing)、标准正态变量变换(SNV)、多元散射校正(MSC)、SG一阶导数(SG 1st-Der)以及小波变换(WT)预处理方法和主成分分析(PCA)、*k*近邻(kNN)、朴素贝叶斯(NBC)、支持向量机(SVM)以及基于 Adaboost 的集成学习等方法对鸭梨黑心病进行判别研究。Adaboost 集成了 kNN、NBC 和 SVM 三个独立学习器。将 120 个健康鸭梨和 165 个黑心鸭梨共计 285 个样品划分为训练集和测试集进行模型的构建和评价, 采用训练集的查准率/查全率的调和平均值(F-measure)和正确识别率(Accuracy)对分类模型进行优化和评价。研究结果表明: 不同属性(正常和黑心)鸭梨样品光谱的前三主成分分布图相互交错, 很难直观地对黑心鸭梨进行区分。样品光谱经小波变换(小波基为“Haar”)预处理后的 kNN 模型训练集的 F-measure 和 Accuracy 分别为 78.98% 和 82.62%; 经过 SG 一阶导数预处理后的 NBC 模型训练集的 F-measure 和 Accuracy 分别为 80.90% 和 82.11%; 经过小波变换预处理后的 SVM 模型训练集的 F-measure 和 Accuracy 分别为 90.24% 和 91.58%; 经小波变换预处理的 AdaBoost 模型训练集的 F-measure 和 Accuracy 分别为 91.46% 和 92.63%。通过测试集对模型进行验证可知: 光谱经小波变换预处理后建立的 Adaboost 分类模型最优, 分类的 F-measure 达到 90.91%, 较 WT-kNN, SG 1st-Der-NBC 和 WT-SVM 模型分别提高了 11.39%, 15.23% 和 2.30%; Accuracy 达到 92.63%, 分别提高了 10.52%, 11.58% 和 2.10%; 模型对测试集样品预测时的计算时间约为 0.12s, 满足在线分选要求。可见-近红外光谱结合 AdaBoost 分类方法, 可以为鸭梨黑心病的在线检测提供一种快速简便的分析方法。

关键词 鸭梨; 黑心病; 可见-近红外光谱; 集成学习; 在线检测

中图分类号: O657.33 **文献标识码:** A **DOI:** 10.3964/j.issn.1000-0593(2021)09-2764-06

引言

鸭梨在储藏过程中由于低温、低氧和高二氧化碳浓度会出现内部褐变的现象^[1]。这与鸭梨组织内的多种酶类、酚类以及膜脂过氧化等因素有关^[2], 主要是多酚氧化酶与酚类物质在有氧的条件下反应生成复杂的醌类化合物^[3], 醌类物质会与氨基酸、蛋白质等发生反应产生褐色的高分子络合物。

可见-近红外光谱应用于鸭梨黑心病检测的主要原因是果实褐变过程中, 果实内部的 C—H, N—H, O—H 等含氢基团振动的倍频和合频在不同化学环境中对可见-近红外光的吸收波长都有明显差别。Gabriëls^[4]等将可见-近红外光谱

与芒果内部颜色分析获得的褐变指数值联系起来, 与神经网络结合可以对芒果内部褐变进行鉴别, 准确度超过 80%。Sun^[5]等使用近红外光谱结合偏最小二乘判别分析模型检测苹果内部褐变, 有 21% 的健康苹果被错误分类。Mogollon^[6]等针对苹果在冷藏期间容易出现内部褐变现象, 提出可见-近红外光谱结合偏最小二乘判别分析模型来预测存储开始时损坏的水果, 达到约 87% 的正确识别率。Khatiwada^[7]采用可见-短波近红外光谱法结合偏最小二乘对完整苹果内部果肉褐变进行静态评估, 并进行线性判别分析和支持向量机分类, 正确识别率为 95% 以上。由文献报道分析可知, 在采用可见-近红外光谱分析方法对果品缺陷进行分析时, 静态分析效果较好, 而动态在线分选由于采用开放式的

收稿日期: 2020-08-01, 修订日期: 2020-12-15

基金项目: 国家自然科学基金项目(31960497)资助

作者简介: 郝勇, 1978年生, 华东交通大学机电与车辆工程学院副教授 e-mail: haonm@163.com

采集装置和较高分级效率的要求, 导致分析模型的精度较低。

针对可见-近红外光谱法在鸭梨黑心缺陷在线分选精度较低的不足, 提出了不同光谱预处理方法结合 k 近邻法 (k -nearest neighbor, kNN)、朴素贝叶斯法 (naive Bayes classifier, NBC)、支持向量机法 (support vector machines, SVM) 以及基于 Adaboost 的集成学习等方法对鸭梨黑心病进行在线判别, 以期得到一种高精度的果品病害在线判别模型的构建方法, 提升我国新鲜果品质量分级水平。

1 实验部分

1.1 样品与仪器

样品: 选用河北鸭梨作为实验样品, 采用冷藏运送, 到达实验室后将鸭梨样品置于 20 °C 恒温条件下保存, 实验前擦除鸭梨表面的污渍与水分。鸭梨样品共计 285 个, 采用 KS (Kennard-Stone) 方法^[8]对样品进行 2 : 1 划分为训练集样本与测试集样本。样品集信息如表 1 所示, 训练集共 190 个鸭梨样品, 包含正常鸭梨 80 个, 黑心鸭梨 110 个; 测试集共 95 个鸭梨样品, 包含正常鸭梨 40 个, 黑心鸭梨 55 个。

表 1 样品集信息

Table 1 Sample set information

样本集	正常梨	黑心梨
训练集	80	110
测试集	40	55

鸭梨可见-近红外光谱在线分选装置示意如图 1 所示, 装置包括输送模块、光谱采集模块和控制模块。输送模块是由变频器和异步电动机控制, 传输速度约为每秒过 5 个鸭梨 (约 0.2 s · 个⁻¹)。光谱采集模块是由卤钨灯和光纤探头组成, 光纤探头安装在托盘的下方, 与输送线的距离约为 120 mm; 光源布置方式如图 2 所示, 由 20 盏 100 W 卤钨灯组成, 每侧 10 盏等距排列。控制模块由 PLC 控制电磁阀和光电接近传感器来触发光谱仪, 完成对鸭梨样品的可见-近红外光谱的采集。

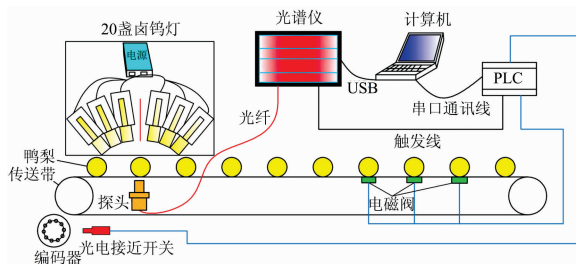


图 1 鸭梨的可见-近红外光谱在线分选装置示意图

Fig. 1 Schematic diagram of the vis-near infrared spectroscopy online sorting device for 'Yali' pear

光谱仪选用美国的 QE65Pro 型高精度光谱仪 Ocean optics INC, 波长范围为 372 ~ 1 154 nm。开机前预热 30

min, 以 6.5 mm 厚度的聚四氟乙烯白板作为标准参比, 校正光源能量谱; 鸭梨在输送线上的托盘里的放置方式要求果柄 (C) 和果蒂 (D) 的连线方向保持与传送带运行方向垂直。

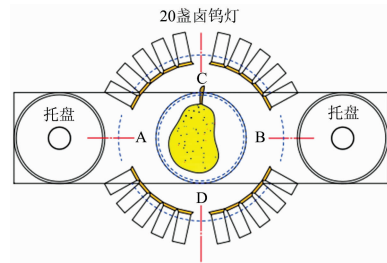


图 2 卤钨灯的排列俯视图

Fig. 2 Arrangement top view of halogen lamp

1.2 黑心梨的破损判别

采集可见-近红外光谱后, 采用传统的人工切开识别法, 对黑心梨进行破损判别。切开鸭梨时, 注意切开的部位是垂直于果柄与果蒂连线方向的中部, 即如图 2 所示的 AB 连线方向。切开后观察鸭梨果核部位有无黑心症状, 对于出现褐色麻点、整个果核褐变及果肉褐变的情况是否被认定为鸭梨患黑心病, 通过 3 位长期从事鸭梨种植和销售的从业者对鸭梨内部的黑心标准进行评价并综合得出结论。

1.3 判别模型构建方法

kNN 算法^[9]是通过计算训练集中的每个样本与测试集样本的距离, 通过对距离的排序, 取距离最近的 k 个点, 这 k 个样本中具有最多的那个类别就是测试集样本的类别。kNN 算法中 k 值的设定影响着模型分类的精确度, k 值选择过大或过小, 都会降低分类精度, 同时也会造成噪声增加, 因此 k 值在选择时一般遵从低于训练样本数的平方根的原则。

NBC 通过属性条件独立性假设, 假设所有属性相互独立^[10]。基于贝叶斯判定准则, 选择每个样本 x 中使后验概率 $P(c|x)$ 最大的类别标记, 最优分类器为

$$h_{nb} = \arg \max_c P(c) \prod_{i=1}^d P(x_i | c) \quad (1)$$

式(1)中, c 属于类别标记, d 为属性值, x_i 为 x 在第 i 个属性上的取值。

SVM 主要是通过找到最大间隔的划分超平面, 使得不同类别之间的间隔最大化, 在处理小样本、非线性及高维数据等问题中具有一定的优势^[11]。通过间隔的概念, 确定模型的约束参数, SVM 的优化目标为

$$\begin{cases} \min \frac{1}{2} \|w\|^2 + c \sum_{i=1}^n \xi_i & \xi_i \geq 0 \\ \text{s. t. } \begin{cases} y_i (wx_i + b) \geq 1 - \xi_i \\ c \geq 0 \end{cases} & i = 1, 2, \dots, n \end{cases} \quad (2)$$

式(2)中, n 为训练样本数, x_i 为训练样本的支持向量, y_i 表示对应样本的类别, 取值为 +1 或 -1, w 为超平面的法向量, b 为偏置向量, c 为惩罚因子, ξ_i 为松弛变量。

基于 AdaBoost 的集成学习可以通过训练多个独立的弱学习器来得到泛化性能优良的强学习器^[12]。集成学习不仅拥有更好的预测性能, 而且解决了单个学习器容易欠拟合及

过拟合的问题。算法的中心思想是通过改变样本的权值分布以得到更好的训练模型。

1.4 模型的评价

F-measure 和 Accuracy 作为分类模型的评价指标, F-measure 和 Accuracy 的结果越接近 1, 表明分类模型越好。其中 F-measure 是由 P (查准率或者准确率, Precision) 和 R (查全率或者召回率, Recall) 加权调和平均得出^[13]。混淆矩阵如表 2 所示。

表 2 分类结果混淆矩阵

Table 2 confusion matrix for classification result

真实情况	预测结果	
	正常梨	黑心梨
正常梨	TP	FN
黑心梨	FP	TN

Precision 与 Recall 的计算公式如式(3)

$$P = \frac{TP}{TP + FP}, R = \frac{TP}{TP + FN} \quad (3)$$

对于鸭梨进出口贸易, 希望既要正常梨尽可能多的挑选出来又要使得挑选出来的鸭梨中正常梨的比例尽可能的高, 即查准率和查全率都需要得到提高。F-measure 公式如式(4)

$$F\text{-measure} = \frac{2PR}{P + R} \quad (4)$$

Accuracy 指的是正确预测的样本数占总预测样本数的比值, 公式如式(5)

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

2 结果与讨论

2.1 鸭梨可见-近红外光谱分析

鸭梨的可见-近红外光谱如图 3 所示, 从光谱的全波段来看, 正常梨与黑心梨的光谱在 695 和 797 nm 左右范围内, 都存在明显的吸收峰, 正常梨的能量谱总体上是高于黑心梨的光谱能量, 但也存在部分正常梨的光谱能量低于黑心梨, 而且光谱存在重叠, 没有一个准确的阈值分割线, 因此不能直接从光谱图中区分鸭梨是否黑心。为了消除可见-近红外光谱存在的光散射、基线漂移等问题, 实验分别采用平滑(Smoothing)、标准正态变量变换(standard normal variate, SNV)、多元散射校正(multiplicative scatter correction, MSC)、SG 一阶导数(savitzky golay first-derivative, SG 1st-Der)和小波变换(wavelet transform, WT)等预处理方法增强光谱的特异性, 以便更好地建模分析。

2.2 正常和黑心病鸭梨光谱的 PCA 分析

采用主成分分析法(principal components analysis, PCA)对两种鸭梨样品光谱的空间分布情况进行分析。建模集含有 120 个正常梨, 165 个黑心梨, 对其进行主成分分析, 前三个主成分累计贡献率占 99.70%, 表示这 3 个主成分能够解释原始波长变量的 99.70%^[14]。图 4 所示(PC1, PC2,

PC3)为正常鸭梨与黑心鸭梨样品的前三个主成分分布图。图中黑心梨与正常梨的光谱点相互交叉在一起, 无法区分, 表明采用可见-近红外光谱结合 PCA 方法对鸭梨是否黑心进行定性鉴别具有一定难度, 需要进一步探讨区分黑心梨与正常梨的判别方法。

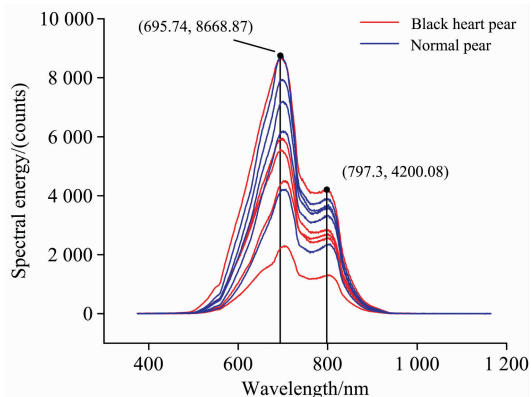


图 3 正常梨与黑心梨的能量谱曲线

Fig. 3 Energy spectra curve of normal pear and black heart pear

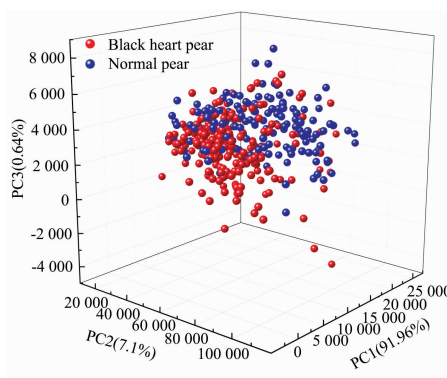


图 4 正常梨与黑心梨的前三主成分分布图

Fig. 4 Distribution of the first three principal components of normal pears and black heart pears

2.3 鸭梨黑心病在线判别模型构建

2.3.1 独立判别模型的构建

采用 kNN 方法进行建模时, 其核心思想是利用训练集数据训练一个 kNN 分类器, 其中 k (预测变量中最近邻的数量)为 5。计算训练集中的每个样本与测试集的欧氏距离, 取距离最近的 5 个点, 选择这 5 个样本中出现最多的类别标记作为预测结果。从表 3 中可知, 经过小波变换预处理的 kNN 模型训练集的 F-measure 和 Accuracy 最高, 分别为 78.98% 和 82.62%。

NBC 分类方法建模的主要原理是使用具有一定平均值和标准偏差的高斯分布对正常鸭梨与黑心鸭梨两个类别内的预测变量分布进行建模。利用训练集的数据, 估算出基于贝叶斯分布的概率分布参数, 测试集数据根据 NBC 模型估计出的分布概率参数, 计算出测试集样本属于正常梨或黑心梨

的先验概率。从表 4 中可知, 鸭梨光谱经过 SG 一阶导数预处理后的 NBC 模型可达到最好的识别效果, 训练集的 F-measure 为 80.90%, Accuracy 为 82.11%。

表 3 不同预处理方法的鸭梨定性识别 kNN 模型判别结果
Table 3 kNN model results of qualitative identification of 'Yali' pears with different pretreatment methods

预处理方法	F-measure/%	Accuracy/%
Raw	77.38	80.00
Smooth	73.49	76.84
SNV	77.38	80.00
MSC	77.38	80.00
SG 1 st -Der	72.11	78.42
WT	78.98	82.62

表 4 不同预处理方法的鸭梨定性识别 NBC 模型判别结果
Table 4 NBC model results of qualitative identification of 'Yali' pears with different pretreatment methods

预处理方法	F-measure/%	Accuracy/%
Raw	53.93	56.84
Smooth	53.93	56.84
SNV	53.93	56.84
MSC	54.02	57.89
SG 1 st -Der	80.90	82.11
WT	68.57	71.05

SVM 分类是利用训练二进制支持向量机分类器建立二分类 SVM 模型。使用训练集和对应的类别标签训练的 SVM 分类器进行二分类, 使用径向基核训练 SVM 分类器, 找到内核函数的比例值。从表 5 中可知, 鸭梨光谱经过小波变换预处理后的 SVM 模型可达到最好的识别效果, 训练集的 F-measure 为 90.24%, Accuracy 为 91.58%。

表 5 不同预处理方法的鸭梨定性识别 SVM 模型判别结果
Table 5 SVM model results of qualitative identification of 'Yali' pears with different pretreatment methods

预处理方法	F-measure/%	Accuracy/%
Raw	85.19	87.37
Smooth	84.81	87.37
SNV	85.19	87.37
MSC	85.19	87.37
SG 1 st -Der	56.52	68.42
WT	90.24	91.58

2.3.2 AdaBoost 集成模型的构建

AdaBoost 分类方法集成 kNN、NBC 和 SVM 三种分类方法, 设独立学习器数目为 3。AdaBoost 模型的原理如图 5 所示, 通过训练集(黑心梨样品集的 2/3 与正常梨样品集的 2/3 组成)来训练 kNN 学习器, 根据 kNN 学习器对鸭梨黑心病的判别表现来调整样本权重, kNN 学习器错误分类的鸭

梨样本的权重得到提高, 正确分类的鸭梨样本的权值将被降低^[15]; 同理, 基于调整后的鸭梨样本分布来训练 NBC 学习器, 鸭梨样本权值根据分类结果再一次得到重新分布; 最后, 将调整后的鸭梨样本作为 SVM 模型的训练集。完成独立学习器的模型建立后, 通过加权投票法 (weighted voting) 得到最终的强学习器。

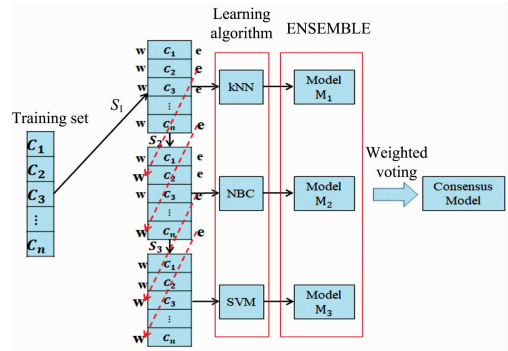


图 5 AdaBoost 算法原理

Fig. 5 AdaBoost algorithm principle

表 6 所示为不同预处理方法下的 AdaBoost 模型的训练集样品的查准率/查全率的调和平均和正确识别率结果。从表中可知, 不同的处理方法其分类结果不尽相同, 鸭梨光谱经过小波变换预处理后的 AdaBoost 模型可达到最好的识别效果, 训练集的 F-measure 为 91.46%, Accuracy 为 92.63%。

表 6 不同预处理方法的鸭梨定性识别 AdaBoost 模型判别结果

Table 6 AdaBoost qualitative identification results of 'Yali' pears with different pretreatment methods

预处理方法	F-measure/%	Accuracy/%
Raw	84.21	87.37
Smooth	84.42	87.37
SNV	85.88	87.37
MSC	83.23	85.79
SG 1 st -Der	77.19	79.47
WT	91.46	92.63

2.3.3 最优分类模型分析

通过模型查准率/查全率的调和平均和正确识别率, 综合比较 kNN 模型、NBC 模型、SVM 模型和 AdaBoost 模型。表 7 可以得出, 鸭梨黑心病鉴别最优模型是 WT-AdaBoost 模型, 其测试集分类结果的 F-measure 为 90.91%, Accuracy 为 92.63%, 模型对测试集样品预测时的计算时间约为 0.12 s, 满足在线分选要求。WT-AdaBoost 的训练集(a)与测试集(b)的预测类别(红色三角形)与实际类别(黑色三角形)比较如图 6 所示, 其中 1 代表正常梨, -1 代表黑心梨, 红色三角形与黑色三角形重合代表预测正确, 否则预测错误(红色三角形+垂线)。

表 7 kNN, NBC, SVM 和 AdaBoost 模型测试集预测结果

分类模型	F-measure / %	Accuracy / %	预测时间估算 * /s
WT-kNN	79.52	82.11	0.04
SG 1 st -Der-NBC	75.68	81.05	0.03
WT-SVM	88.61	90.53	0.04
WT-AdaBoost	90.91	92.63	0.12

注: * 所用电脑配置: CPU: Intel(R) Core(TM) i5RAM: 8GB

Note: * Computer configuration; CPU: Intel (R) Core (TM) i5RAM: 8GB

3 结论

可见-近红外光谱结合 PCA, kNN, NBC, SVM 和基于 Adaboost 的集成学习法建模对鸭梨黑心病进行判别研究, 原始光谱和 5 种预处理光谱结合 kNN, NBC, SVM 和 AdaBoost 方法用于鸭梨黑心病判别模型的建立和优化。实验结果表明: PCA 方法无法在主成分空间将黑心鸭梨与正常鸭梨区分开来; 小波变换预处理方法结合由 kNN, NBC 和 SVM 集成的 AdaBoost 分类方法建立的分类模型最优, 训练集与测试集的 F-measure 分别为 91.46% 和 90.91%, Accuracy 分别为 92.63% 和 92.63%, 且模型对测试集样品预测时间约为 0.12 s, 满足在线分选要求。可见-近红外光谱结合 WT-AdaBoost 分类方法, 可以实现对鸭梨黑心病的在线检测。

References

- [1] Sun H J, Luo M L, Zhou X, et al. Food Chemistry, 2020, 306: 626.
- [2] Zhou H S, Tian M Y, Huang W, et al. Gene Expression Patterns, 2020, 36: 113.
- [3] LI Yue-yuan, FAN Xiao-lan, ZHANG Yin-yin, et al(李月圆, 樊晓岚, 张引引, 等). Food Technology(食品科技), 2018, 43(4): 23.
- [4] Gabriëls S H E J, Mishra P, Mensink M G J, et al. Postharvest Biology and Technology, 2020, 166: 11206.
- [5] Sun J, Künnemeyer R, McGlone A, et al. Computers and Electronics in Agriculture, 2018, 155: 32.
- [6] Mogollon M R, Jara A F, Contreras C, et al. Postharvest Biology and Technology, 2020, 161: 60.
- [7] Khatiwada B P, Subedi P P, Hayes C, et al. Postharvest Biology and Technology, 2016, 120: 103.
- [8] Sun Xudong, Liu Yande, Li Yifan, et al. Postharvest Biology and Technology, 2016, 116: 80.
- [9] Pan Z B, Wang Y D, Ku W P. Expert Systems with Applications, 2017, 67: 115.
- [10] Zhang H, Liu C T, Mao J, et al. Toxicology in Vitro, 2020, 65: 812.
- [11] PENG Yan-kun, ZHAO Fang, LI Long, et al(彭彦昆, 赵芳, 李龙, 等). Transactions of the Chinese Society of Agricultural Engineering(农业工程学报), 2018, 34(5): 159.
- [12] Shahraki Amin, Abbasi Mahmoud, Haugen Øystein. Engineering Applications of Artificial Intelligence, 2020, 94: 103.
- [13] Berger A, Guda S. Pattern Recognition, 2020, 102: 107.
- [14] HAO Yong, SHANG Qing-yuan, RAO Min, et al(郝勇, 商庆园, 饶敏, 等). Spectroscopy and Spectral Analysis(光谱学与光谱分析), 2019, 39(3): 705.
- [15] Li J L, Sun L J, Li R N. Optik, 2020, 206: 164.

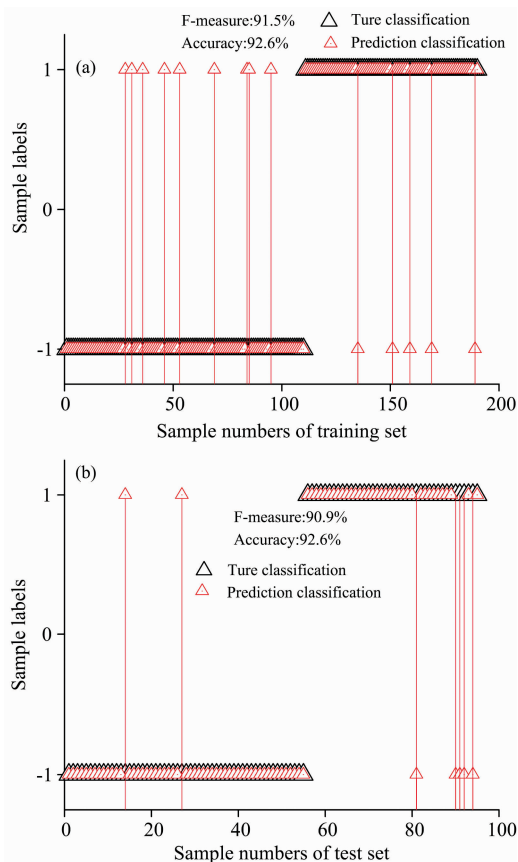


图 6 鸭梨样品实际类别与 WT-AdaBoost 模型预测类别比较图

Fig. 6 Comparison of actual categories and predicted categories in WT-AdaBoost model for 'Yali' pear samples

Study on Online Detection Method of “Yali” Pear Black Heart Disease Based on Vis-Near Infrared Spectroscopy and AdaBoost Integrated Model

HAO Yong¹, WANG Qi-ming¹, ZHANG Shu-min²

1. School of Mechatronics and Vehicle Engineering, East China Jiaotong University, Nanchang 330013, China

2. Technology Center of Nanchang Customs District, Nanchang 330038, China

Abstract Black heart disease is a physiological disease that occurs during the storage of “Yali” pears. The initial stage of the disease manifests itself in brown plaques on the inner core, but there is no difference in the appearance of the fruit from normal fruits, which seriously affects the storage time and quality of “Yali” pears. A fast and non-destructive testing method is urgently needed to escort the quality of “Yali” pears. The vis-near infrared spectroscopy method was used to explore the feasibility of online detection of “Yali” pear black heart disease, combined with principal component analysis (PCA), k-nearest neighbor (kNN), naive Bayes classifier (NBC), support vector machines (SVM), and integrated learning based on Adaboost modeling were used to identify “Yali” pear black heart disease. Standard normal variable (SNV), multiplicative scatter correction (MSC), Savitzky Golay first-derivative derivative (SG 1st-Der) and wavelet transform (WT) were used to preprocess the spectra. Adaboost integrates three base learners: kNN, NBC and SVM. A total of 285 samples, including 120 normal pears and 165 black hearted pears, divided into the training set and test set for model construction and evaluation. The harmonic average of the precision/recall rate (F-measure) and accuracy were used to optimize and evaluate the classification model. The results of the study show that the first three principal components of the spectrum of the samples of different attributes (normal and black heart) “Yali” pears were interlaced with each other, and it was difficult to distinguish the black heart pears visually. The F-measure and accuracy of the training set of the kNN model, in which the spectra of the samples were preprocessed by wavelet transform (the wavelet basis is “Haar”), were 78.98% and 82.62%, respectively. The F-measure and accuracy of the training set of NBC model after the Savitzky Golay first-derivative pretreatment were 80.90% and 82.11%, respectively. The F-measure and accuracy of the training set of SVM model after the wavelet transform pretreatment were 90.24% and 91.58%, respectively. The F-measure and accuracy of the training set of AdaBoost model after the wavelet transform pretreatment were 91.46% and 92.63% respectively. By verifying the model through the test set, it can be known that; the Adaboost classification model after the wavelet transform pretreatment was the best, and the F-measure reached 90.91%, which was 11.39%, 15.23% and 2.30% higher than that of WT-kNN model, SG 1st-Der-NBC model and WT-SVM model, respectively. Accuracy reached 92.63%, improved by 10.52%, 11.58% and 2.10% respectively. The calculation time of the model for the prediction of test set samples was about 0.12 s, which meets the requirements of online sorting. The combination of vis-near infrared spectroscopy and the AdaBoost classification method can provide a quick and easy analysis method for online detection of “Yali” pear blackheart disease.

Keywords “Yali” pear; Black heart disease; Vis-near infrared spectroscopy; Integrated learning; Online detection

(Received Aug. 1, 2020; accepted Dec. 15, 2020)