

半透射高光谱多指标同时检测马铃薯内外部缺陷

徐梦玲, 李小昱*, 库静, 曲宝羊
(华中农业大学工学院, 湖北 武汉 430070)

摘要: **目的** 应用半透射高光谱成像技术结合支持向量机(support vector machine, SVM)模型实现马铃薯内外部缺陷多指标同时检测。**方法** 采集 310 个马铃薯样本半透射高光谱图像, 并分别采用标准正态变量变换(standard normalized variate, SNV)、归一化(normalize)和平滑处理(smoothing)对光谱信息进行预处理。进一步采用竞争性自适应重加权算法结合无信息变量消除法(competitive adaptive reweighted sampling algorithm, uninformative variable elimination, CARS-UVE)进行特征波长选择, 提高模型识别率。**结果** 原始光谱信息经归一化预处理和竞争性自适应重加权算法结合无信息变量消除法(CARS-UVE)降维后所建的支持向量机(SVM)模型识别结果最优, 该方法对合格、绿皮和黑心马铃薯样本预测结果分别为 90.7%、88.9%、95.7%, 混合识别率为 91.3%。**结论** 采用半透射高光谱成像技术结合 CARS-UVE 方法所建 SVM 模型能够实现马铃薯内外部缺陷多指标同时检测。

关键词: 高光谱成像; 支持向量机; 内外部缺陷; 马铃薯

Simultaneous detection of multiple index for internal and external defects of potato based on semi-transmission hyperspectral

XU Meng-Ling, LI Xiao-Yu*, KU Jing, QU Bao-Yang

(College of Engineering, Huazhong Agricultural University, Wuhan 430070, China)

ABSTRACT: Objective To realize multi-index simultaneous detection of internal and external defects of potato by semi-transmissive hyperspectral imaging technology combined with support vector machine (SVM) model. **Methods** A total of 310 semitransparent hyperspectral images of potato samples were collected, spectral information of that were pretreated by using standard normal variate transformation (SNV), normalize and smoothing, respectively. Further, characteristics of wavelength were selected using competitive adaptive reweighted sampling algorithm combined with uninformative variable elimination (CARS-UVE) to improve pattern recognition rate. **Results** The model of SVM which was built by original spectral information that got normalize pretreatment and CARS-UVE dimension reduction was the optimal identification results. Using this method to forecast eligibility, green skin and black heart potato sample rate was 90.7%, 88.9% and 95.7%, respectively, and hybrid recognition rate could reach to 91.3%. **Conclusion** The established SVM model of potato by semi-transmissive hyperspectral imaging technologies combined with CARS-UVE method can realize potato internal and external defects simultaneous detection in multiple index.

基金项目: 国家自然科学基金项目(61275156)、湖北省自然科学基金重点资助项目(2011CDA033)

Fund: Supported by the National Natural Science Foundation of China (61275156) and the Hubei Natural Science Foundation Funded Project (2011CDA033)

*通讯作者: 李小昱, 教授, 博士生导师, 主要研究方向为农产品无损检测检测技术。E-mail: lixiaoyu@mail.hzau.edu.cn

*Corresponding author: LI Xiao-Yu, Professor, College of Engineering, Huazhong Agricultural University, Wuhan 430070, China. Email: lixiaoyu@mail.hzau.edu.cn

KEY WORDS: hyperspectral imaging; support vector machine; internal and external defects; potato

1 引言

我国是世界上最大的马铃薯种植国,但是马铃薯的深加工相对发达国家较为落后,产品的附加值较低,因此,完善马铃薯品质的检测方法和分类技术是提高马铃薯价值的关键^[1]。

马铃薯的品质可分为能用肉眼观察的外部品质和不易直观识别的内部品质^[2,3]。针对马铃薯外部缺陷, Dacal 等^[4]通过高光谱成像结合特征波段的方法建立支持向量机模型对马铃薯疮痂病进行检测,模型识别率达到 97.1%;周竹等^[5]采用反射高光谱成像技术结合波段比算法和均匀二次差分算法对 6 类外部缺陷马铃薯进行检测,总体识别率为 95.65%;高海龙等^[6]针对马铃薯损伤部位放置位置会影响检测精度的问题,提出了从背对相机、正对相机和侧对相机 3 个不同方向随机放置的方法采集马铃薯图像,建立了透射和反射高光谱成像的马铃薯机械损伤检测模型,最终确定采用透射光谱可实现任意放置方向马铃薯损伤检测。针对不易检测的马铃薯内部品质,黄涛等^[7]基于半透射高光谱成像技术使用竞争性自适应重加权算法结合连续投影算法(CARS-SPA)和人工鱼群算法结合支持向量机法(AFSA-SVM)对马铃薯空心病进行准确检测,采用支持向量机建立的模型准确率为 100%。高海龙等^[8]采用无信息变量消除算法结合连续投影算法(UVE-SPA)对马铃薯透射高光谱数据中光谱变量进行选择,最终建立偏最小二乘判别分析(PLS-DA)模型对马铃薯内部黑心病进行检测,识别率达到 100%。

综上所述,国内外已有基于高光谱成像技术的马铃薯内外部缺陷指标检测的相关研究^[9-11],然而,针对马铃薯内外部多个缺陷同时分类检测的相关研究还需深入。因此,开展基于高光谱成像技术的马铃薯内外部缺陷多分类检测方法研究具有重要的科学意义和应用价值。

2 材料与方方法

2.1 试验材料

试验用马铃薯样本购于武汉洪山区关山农副产品批发市场。在样本高光谱图像采集完成后将马铃薯

沿最大横截面切开以判定样本是否黑心,最终选定合格样本 145 个,绿皮样本 106 个,黑心样本 59 个。

按约 2:1 的比例,采用随机样本生成法将 310 个马铃薯划分为校正集与测试集,校正集中共有样本 206 个(合格马铃薯 91 个、绿皮马铃薯 30 个、黑心马铃薯 85 个);测试集中共有样本 104 个(合格马铃薯 54 个、绿皮马铃薯 27 个、黑心马铃薯 23 个)。

图 1 分别为马铃薯典型黑心样本的剖面和绿皮样本。

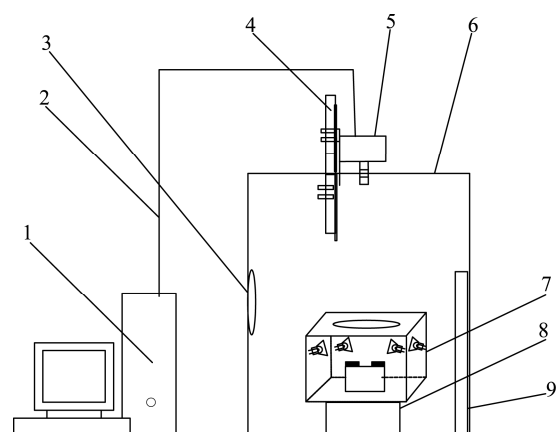


图 1 马铃薯样本

Fig. 1 Test samples of potatoes

2.2 仪器设备及参数选择

半透射高光成像系统如图 2 所示。该系统主要由高光谱成像仪(SPECIM, V10E, Finland),半透射光源箱(专利号: ZL201320694193.6)、高精度电控移动平台,排风扇和专用计算机等组成。可采集到的光谱波长范围为 390~1040 nm,光谱的分辨率为 2.8 nm。半透射光源箱由 4 个 50 W 卤钨灯、采集箱体和载物台等组成。根据研究结果,设置摄像头曝光时间为 100 ms,高精度电控移动平台运行速度为 2 mm/s。



1. 专用计算机 2. 数据线 3. 排风扇 4. 高度可调支架 5. 高光谱成像仪 6. 采集暗箱 7. 半透射光源箱 8. 高精度电控移动平台 9. 箱门

图 2 半透射高光谱成像系统

Fig. 2 Transmission hyperspectral imaging system

2.3 支持向量机判别分析(SVM)

支持向量机通过寻找高维特征空间中的最优分离超平面进行线性分类,具有以有限的样本信息获得精准稳定的模型能力。在解决少量样本、高维模式以及非线性的模式识别过程中具有明显的优势,因此,应用十分广泛^[13,14]。本研究以此作为建模方法。

3 结果与讨论

3.1 马铃薯样本原始光谱分析

应用高光谱图像处理软件 ENVI4.7 随机选取半透射高光谱图像 80×80 像素的感兴趣区域,计算区域内的平均光谱作为该样本的原始光谱,图 3 分别为马铃薯黑心、合格、绿皮样本的半透射光谱曲线。由图 3 可知,黑心样本与合格样本的光谱存在明显差异,黑心样本由于内部黑心,对光的吸收能力较强,在全波段范围内黑心样本的光谱透射率均低于合格样本的透射率。合格马铃薯和绿皮马铃薯样本的光谱曲线相似,在光谱区域总体呈现正态分布趋势。在 700 nm 和 790 nm 附近两种样本光谱都呈现明显峰值,但在 560 nm 至 840nm 附近的绿皮样本光谱峰值要高于合格样本光谱峰值,这可能是由于马铃薯在产生绿皮过程中其内部组织中龙葵素的有毒生物碱含量会急剧增长所造成的。

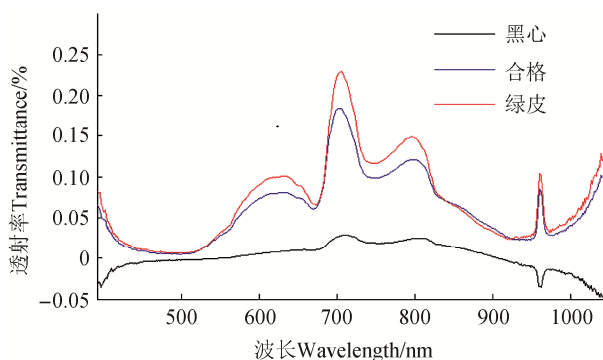


图3 平均光谱曲线

Fig. 3 Average spectral curve

3.2 不同预处理方法马铃薯缺陷多分类模型

由于系统噪声、图像背景等因素会对原始光谱数据造成一定的影响,因此,需要对光谱数据进行预处理,以提高模型性能。

分别采用变量标准化(normalize)、平滑处理

(smoothing)及标准正态变量校正(standard normal variate, SNV)对光谱数据进行预处理,不同的预处理方法所建立的 SVM 模型结果如表 1 所示。

表1 不同光谱预处理法的多分类 SVM 模型检测结果
Table 1 Results of the multi-category SVM model based on different preprocessing methods

预处理方法	类型	校正集		测试集	
		误判数	识别率/%	误判数	识别率/%
变量标准化	合格	6	93.4%	6	88.9%
	绿皮	11	86.2%	4	85.2%
	黑心	1	97.1%	2	91.3%
	混合	18	91.3%	12	88.5%
平滑处理	合格	2	97.8%	3	94.4%
	绿皮	4	95.0%	1	96.3%
	黑心	0	100%	1	95.6%
	混合	6	97.1%	5	95.2%
标准正态变量校正	合格	13	85.7%	9	83.3%
	绿皮	12	77.5%	4	85.2%
	黑心	0	100%	1	95.7%
	混合	25	87.9%	14	86.5%

由表1可知,对比几种预处理方法的校正集与测试集识别准确率,原始光谱经过变量标准化预处理后建立的 SVM 模型识别率最优,该模型对测试集合格、绿皮和黑心马铃薯样本的识别率分别为 88.9%、85.2%、91.3%,混合识别率达到 88.5%。原始光谱经过平滑滤波预处理后出现过拟合现象,不符合数据处理的要求。因此,选取变量标准化作为最优预处理方法。

3.3 基于特征波长选择的多分类模型优化

为了提高模型性能,需要对全波段光谱进行变量选择,剔除冗长光谱信息。

竞争性自适应重加权算法(CARS)是建立在模仿达尔文进化理论中“适者生存”的原则基础上提出的变量选择方法,经过多次的自适应重加权采样技术得出偏最小二乘模型中需要的波长变量,可以有效地选择和样品特性有关的波长组合^[15,16]。

图 4 是马铃薯内外部缺陷检测模型光谱变量的 CARS 选择示意图,图 4(a)表示在特征波长变量选择过程中,波长数量随采集次数的变化趋势。所选变量

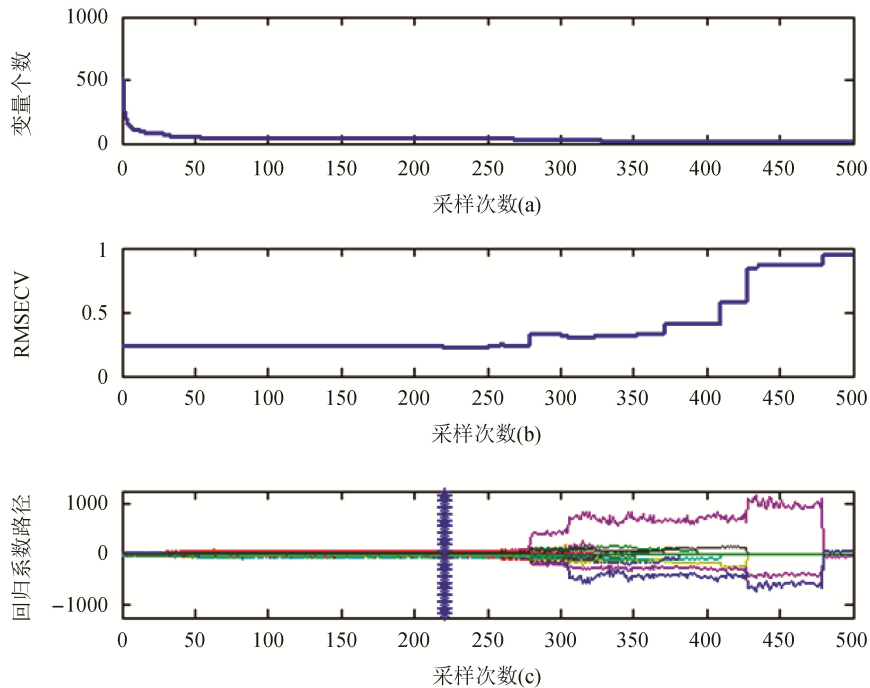


图 4 CARS 选择过程

Fig. 4 CARS selection process

数目随采集次数的增加而由快到慢逐渐递减, 展示出粗选阶段和精选阶段, 最终保留一定数量的有效变量; 图 4(b)表示马铃薯内外缺陷样本校正集经 10 折交叉检验所得到的 RMSECV 变化趋势图, 由图 4(b)可以看出, 在前 270 次采样过程中, 所建模型的 RMSECV 逐渐减小, 表明变量选择过程中去除马铃薯内外部缺陷无关的变量, 第 271 次以后采样建模 RMSECV 开始逐渐增大, 表明变量选择过程中可能去除了与马铃薯内外部缺陷检测有关的变量; 图 4(c)的“*”表示马铃薯内外外部缺陷样本校正集最小 RMSECV 所对应的采样次数, 使用 CARS 特征波长选择方法最终选择的波长变量个数为 33 个。

无信息变量消除法(UVE)是一种基于回归系数稳定性分析的新变量选择方法。它能有效剔除无信息的波长变量, 防止预测模型过拟合, 改善模型的预测能力。 UVE 的详细算法与原理见文献^[17]。

采用 CARS-UVE 特征波长选择方法得到的 29 个特征变量建立马铃薯合格、绿皮和黑心多分类的 SVM 检测模型, 模型识别效果如表 2 所示。

由表 2 可知: 采用 CARS-UVE 特征波长选择方法所建 SVM 检测模型与 520 个特征变量的全波段模型比较, 样本光谱的波段数量有较大减少。该模型测

试集合格、绿皮和黑心马铃薯样本的识别率分别为 90.7%、88.9%、95.7%, 混合识别率达到 91.3%。其中混合模型测试集误判个数由原来的 12 个减少到 9 个, CARS-UVE 特征波长选择方法不仅简化了模型, 而且提高了模型性能。

表 2 CARS-UVE 波长选择方法的 SVM 模型结果
Table 2 Results of SVM model based on CARS-UVE wavelength selection method

类型	校正集		测试集	
	误判数	识别率	误判数	识别率
合格	5	94.5%	5	90.7%
绿皮	8	90.0%	3	88.9%
黑心	0	100%	1	95.7%
混合	13	93.7%	9	91.3%

4 结 论

以合格、绿皮和黑心马铃薯为研究对象, 研究了基于半透射高光谱成像技术的马铃薯内外部缺陷的多分类模型检测方法。

比较了标准正态变量变换、归一化和平滑处理 3 种不同的光谱预处理方法, 确定了归一化为最优光

谱预处理方法,对预处理后的全波段光谱变量建立马铃薯内外部缺陷的SVM分类检测模型,该模型对测试集合格、绿皮和黑心马铃薯样本的识别率分别为88.9%、85.2%、91.3%,混合识别率达到88.5%。

为进一步剔除冗余光谱信息,提高模型的精度,采用CARS-UVE特征波长选择方法对全波段520个光谱变量进行筛选,确定了马铃薯合格、绿皮和黑心多分类SVM检测的29个光谱变量,基于所选变量建立半透射高光谱成像技术的马铃薯SVM多分类检测模型,该模型对测试集合格、绿皮和黑心马铃薯样本的识别率分别为90.7%、88.9%、95.7%,混合识别率达到91.3%。采用半透射高光谱成像技术结合CARS-UVE方法所建SVM模型能够实现马铃薯内外部缺陷多指标同时检测。

参考文献

- [1] 周竹, 李小昱, 高海龙, 等. 马铃薯干物质含量高光谱检测中变量选择方法比较[J]. 农业机械学报, 2012, 43(2): 128-133, 185.
Zhou Z, Li XY, Gao HL, *et al.* Comparison of different variable selection methods on potato dry matter detection by hyperspectral imaging technology [J]. T Chin Soc Agric Mach, 2012, 43(2): 128-133, 185.
- [2] 李江波, 饶秀勤, 应义斌. 农产品外部品质无损检测中高光谱成像技术的应用研究进展[J]. 光谱学与光谱分析, 2011, 31(8): 2021-2026.
Li JB, Rao XQ, Ying YB, *et al.* Advance on application of hyperspectral imaging to nondestructive detection of agricultural products external quality [J]. Spectrosc Spect Anal, 2011, 31(8): 2021-2026.
- [3] 单佳佳, 彭彦昆, 王伟, 等. 基于高光谱成像技术的苹果内外品质同时检测[J]. 农业机械学报, 2011, 42(3): 140-144.
Shan JJ, Peng YK, Wang W, *et al.* Simultaneous detection of external and internal quality parameters of apples using hyperspectral technology [J]. T Chin Soc Agric Mach, 2011, 42(3): 140-144.
- [4] Dacal-Nieto A, Formella A, Carrión P, *et al.* Non-destructive detection of hollow heart in potatoes using hyperspectral imaging [C]. Computer Analysis of Images and Patterns. Springer Berlin Heidelberg, 2011: 180-187.
- [5] 周竹, 李小昱, 陶海龙, 等. 基于高光谱成像技术的马铃薯外部缺陷检测[J]. 农业工程学报, 2012, 28(21): 221-228.
Zhou Z, Li XY, Tao HL, *et al.* Detection of potato external defects based on hyperspectral imaging technology [J]. T CSAE, 2012, 28(21): 221-228.
- [6] 高海龙, 李小昱, 徐森森, 等. 透射和反射高光谱成像的马铃薯损伤检测比较研究[J]. 光谱学与光谱分析, 2013, 33(12): 3366-3371.
Gao HL, Li XY, Xu SM, *et al.* Comparative study of transmission and reflection hyperspectral imaging technology for potato damage detection [J]. Spectrosc Spect Anal, 2013, 33(12): 3366-3371.
- [7] 黄涛, 李小昱, 徐梦玲, 等. 半透射高光谱成像技术与支持向量机的马铃薯空心病无损检测研究[J]. 光谱学与光谱分析, 2015, 35(1): 198-202.
Huang T, Li XY, Xu ML, *et al.* Non-destruction detection research for hollow heart of potato based on semi-transmission hyperspectral imaging and SVM [J]. Spectrosc Spec Anal, 2015, 35(1): 198-202.
- [8] 高海龙, 李小昱, 徐森森, 等. 马铃薯黑心病和单薯质量的透射高光谱检测方法[J]. 农业工程学报, 2013, 29(15): 279-285.
Gao HL, Li XY, Xu SM, *et al.* Transmission hyperspectral detection method for weight and black heart of potato [J]. T CSAE, 2013, 29(15): 279-285.
- [9] 李小昱, 陶海龙, 高海龙, 等. 马铃薯缺陷透射和反射机器视觉检测方法分析[J]. 农业机械学报, 2014, 45(5): 191-196.
Li XY, Tao HL, Gao HL, *et al.* Comparison of transmission and reflection imaging technologies to detect potato defects based on machine vision technology [J]. T Chin Soc Agric Mach, 2014, 45(5): 191-196.
- [10] 周竹, 李小昱, 高海龙, 等. 漫反射和透射光谱检测马铃薯黑心病的比较[J]. 农业工程学报, 2012, 28(11): 237-242.
Zhou Z, Li XY, Gao HL, *et al.* Comparison of diffuse reflection and transmission mode of visible/near infrared spectroscopy for detecting black heart of potato [J]. T CSAE, 2012, 28(11): 237-242.
- [11] 黄涛, 李小昱, 金瑞, 等. 半透射高光谱结合流形学习算法同时识别马铃薯内外部缺陷多项指标[J]. 光谱学与光谱分析, 2015, 35(4): 992-996.
Huang T, Li XY, Jin R, *et al.* Multi-target recognition of internal and external defects of potato by semi-transmission hyperspectral imaging and manifold learning algorithm [J]. Spectrosc Spec Anal, 2015, 35(4): 992-996.
- [12] 黄涛. 一种马铃薯漫透射高光谱图像采集装置[P]. 中国: ZL201320694193.6, 2013.
Huang T, A potato diffuse transmission hyperspectral image acquisition device [P]. China: ZL201320694193.6, 2013.

- [13] 高昆仑, 刘建明, 徐茹枝, 等. 基于支持向量机和粒子群算法信息网络安全态势复合预测模型[J]. 电网技术, 2011, 35(4): 176–182.
Gao KL, Liu JM, Xu RZ, *et al.* A hybrid security situation prediction model for information network based on support vector machine and particle swarm optimization [J]. Power Sys Technol, 2011, 35(4): 176–182.
- [14] 丁世飞, 齐丙娟, 谭红艳. 支持向量机理论与算法研究综述[J]. 电子科技大学学报, 2011, 40(1): 2–10.
Ding SF, Qi BJ, Tan HY, *et al.* An overview on theory and algorithm of support vector machines [J]. J Univ Elec Sci Technol China, 2011, 40(1): 2–10.
- [15] 孙通, 许文丽, 林金龙, 等. 可见/近红外漫透射光谱结合 CARS 变量优选预测脐橙可溶性固形物[J]. 光谱学与光谱分析, 2012, 32(12): 3229–3233.
Sun T, Xu WL, Lin JL, *et al.* Determination of soluble solids content in navel oranges by Vis/NIR diffuse transmission spectra combined with cars method [J]. Spectrosc Spec Anal, 2012, 32(12): 3229–3233.
- [16] 李江波, 郭志明, 黄文倩, 等. 应用 CARS 和 SPA 算法对草莓 SSC 含量 NIR 光谱预测模型中变量及样本筛选[J]. 光谱学与光谱分析, 2015, 35(2): 372–378.
Li JB, Guo ZM, Huang WQ, *et al.* Near -infrared spectra combining with CARS and SPA algorithms to screen the variables and samples for samples for quantitatively determining the soluble solids content in strawberry [J]. Spectrosc Spec Anal, 2015, 35(2): 372–378.
- [17] Centner V, Massart DL, deNoord OE, *et al.* Elimination of uninformative variables for multivariate calibration [J]. Anal Chem, 1996, 68(21): 3851–3858.

(责任编辑: 李振飞)

作者简介



徐梦玲, 硕士研究生, 主要研究方向为智能化检测技术。

E-mail: tina19910917@126.com



李小昱, 教授, 博士生导师, 主要研究方向为农产品检测技术。

E-mail: lixiaoyu@mail.hzau.edu.cn