

基于近红外高光谱的梨叶片炭疽病与黑斑病识别

刘莉¹ 陶红燕¹ 方静¹ 郑文娟¹ 王良龙² 金秀²
(1.安徽农业大学园艺学院,合肥 230036; 2.安徽农业大学信息与计算机学院,合肥 230036)

摘要: 针对梨炭疽病和黑斑病发病症状很相似,难以区分,导致实际生产中不便对症施药的问题,以砀山酥梨叶片为研究对象,探究利用高光谱技术来识别梨叶片炭疽病与黑斑病的可行性。首先,运用高光谱成像系统采集砀山酥梨正常叶片、炭疽病叶片和黑斑病叶片的高光谱图像,提取图像的平均光谱反射率。采用多元散射校正法(Multiplicative scatter correction, MSC)、Savitzky-Golay 卷积平滑法和标准正态变换法(Standard normal variate, SNV)分别对原始光谱数据进行预处理。然后,采用主成分分析算法(Principal component analysis, PCA)、连续投影算法(Successive projections algorithm, SPA)、无信息变量消除法(Uninformative variable elimination, UVE)、竞争性自适应重加权算法(Competitive adaptive reweighted sampling, CARS)、随机蛙跳算法(Shuffled frog leaping algorithm, SFLA)提取特征波长,分别获取了 27、12、15、26、20 条特征波长,并将其作为后期建模的输入变量。经对比发现,在各基于特征波长建立的支持向量机(SVM)分类识别模型以及 BP 神经网络分类识别模型中,SPA-SVM 识别模型效果最佳,测试集准确率为 93.25%,建模集准确率为 94.80%。试验结果证明,利用高光谱技术能够有效识别砀山酥梨叶片的黑斑病与炭疽病。

关键词: 砀山酥梨; 炭疽病; 黑斑病; 高光谱; 特征波长; 识别模型

中图分类号: S436.612.1; S661.2 文献标识码: A 文章编号: 1000-1298(2022)02-0221-10 OSID: 

Identifying Anthracnose and Black Spot of Pear Leaves on Near-infrared Hyperspectroscopy

LIU Li¹ TAO Hongyan¹ FANG Jing¹ ZHENG Wenjuan¹ WANG Lianglong² JIN Xiu²
(1. College of Horticulture, Anhui Agricultural University, Hefei 230036, China
2. College of Information and Computer Science, Anhui Agricultural University, Hefei 230036, China)

Abstract: Pear anthracnose and pear black spot are serious diseases that occur during the growth of pears. The symptoms of these two diseases are very similar and it is difficult to distinguish, which leads to the inconvenience of prescribing the right medicine to these two kinds of leaves in actual production. In response to the status quo, taking ‘Dangshan’ pear leaves as the study object, the feasibility of using hyperspectral technology to identify anthracnose and black spot on pear leaves was explored. First of all, the hyperspectral imaging system was used to collect the hyperspectral images of the normal leaves, anthracnose leaves and black spot leaves of ‘Dangshan’ pear, and extract the average spectral reflectance of the images. The multiplicative scatter correction method (MSC), Savitzky-Golay convolution smoothing method and standard normal variate method (SNV) were used respectively to preprocess the original spectral data. Then the principal component analysis (PCA), successive projections algorithm (SPA), uninformative variable elimination (UVE), competitive adaptive reweighted sampling algorithm (CARS), and shuffled frog leaping algorithm (SFLA) were used to extract characteristic wavelengths, respectively, and totally 27, 12, 15, 26 and 20 characteristic wavelengths were obtained, and using them as input variables for later modeling. After comparison, it was found that in the support vector machine (SVM) classification and recognition model based on characteristic wavelength and the BP

收稿日期: 2021-02-25 修回日期: 2021-04-20
基金项目: 财政部和农业农村部:国家现代农业产业技术体系建设专项(CARS-29-14)、国家重点研发计划项目(2018YFD0201401)和安徽省教育厅项目(KJ2019A0212)
作者简介: 刘莉(1980—),女,副教授,博士,主要从事果树资源遗传与育种研究,E-mail: 81735296@qq.com
通信作者: 金秀(1983—),男,副教授,博士,主要从事人工智能、机器学习和深度学习研究,E-mail: jinxiu123@ahau.edu.cn

neural network classification and recognition model based on characteristic wavelength, the SPA - SVM recognition model had the best effect during all models, the accuracy rate of the model's test set was 93.25%, and the accuracy rate of the model's modeling set was 94.80%. The test results proved that hyperspectral technology can effectively identify the black spot and anthracnose of 'Dangshan' pear leaves.

Key words: 'Dangshan' pear; anthracnose; black spot; hyperspectral; characteristic wavelength; identification model

0 引言

梨是我国的种植范围广泛,产量仅次于苹果、柑橘^[1]。在梨树的生长过程中,各种病害不仅威胁着梨树的生长,还会影响梨果实的品质,最典型的有黑斑病、炭疽病、锈病等,且病害的大规模爆发会造成严重的经济损失^[2]。因此,及时检测梨叶片病害,并采取针对性的病害防治措施,对于梨树病害防治、减少经济损失有着重要意义。

在梨树诸多病害中,炭疽病和黑斑病的发病症状相似,仅凭肉眼很难区分。尽管目前可见-近红外检测技术^[3-5]、机器视觉检测技术^[6-7]等已经广泛运用于农作物的无损检测,但是可见-近红外只能检测到农作物的内部信息,而无法获取其外部特征,机器视觉检测技术虽然可以获得农作物的外部特征,但是却无法检测其内部信息。相比之下,高光谱技术具有图谱合一的特点,既可以探测到农作物的外部图像信息,又可以获得其内部品质信息。

目前已有很多学者利用高光谱技术在农产品品质检测方面做了大量研究^[8-10]。国外学者侧重于将高光谱技术用于肉类以及水果的品质测定^[11-12],国内学者则多将此技术应用于蔬菜水果农产品的无损检测、病虫害识别与分类^[13-15]。ELMASRY等^[16]利用波长400~1 000 nm的高光谱成像仪采集“McIntosh”苹果的高光谱图像,建立了“McIntosh”苹果的早期损伤判别模型,并通过试验表明,利用高光谱技术可以有效识别损伤1 h以上的苹果与正常苹果;MEHL^[17]利用高光谱技术建立了Golden Delicious、Gala、Red Delicious 3个品种苹果的表面损伤程度的检测模型,检测结果的准确率分别为85%、95%和76%;吴龙国等^[18]利用高光谱图像采集技术采集了波长400~1 000 nm范围的高光谱图像,对灵武长枣表面的外部碰伤缺陷进行检测,结果较为理想;李勋兰等^[19]利用高光谱成像技术采集了4种柚子的上表面与下表面的高光谱图像,构建了柚子种类鉴别的识别模型,准确率分别为99.46%和98.44%。但目前有关利用高光谱技术对梨树病害种类进行识别研究的报道比较少。

本文以砀山酥梨叶片黑斑病与炭疽病为研究对

象,以砀山酥梨正常叶片、炭疽病叶片与黑斑病叶片的高光谱图像为试验样本,通过对高光谱图像光谱信息的预处理、特征波长提取,并结合支持向量机(SVM)与反向传播(BP)神经网络模式识别等建立砀山酥梨炭疽病和黑斑病的识别模型,以期为果树的病害识别提供参考依据与技术支持。

1 试验材料与识别方法

1.1 试验材料

本试验供试材料为安徽农业大学资源育种实验室盆栽砀山酥梨叶片,挑选采集叶片大小均匀、叶面较为宽大、位于新梢顶端且健康的嫩叶叶片,共90片。

提前在实验室配置好炭疽病与黑斑病菌丝。接种病菌前,用标签标注每片叶片,并采集所有叶片正反两面的高光谱图像,共获取180幅高光谱图像。然后对90片样品叶片进行菌丝块接种处理。

为了观察各叶片从健康状态到发病状态的连续动态过程,自接种病菌当天开始连续对叶片进行高光谱图像采集,每天记录发病叶片的标签号及其发病症状,并将肉眼所能看到的发病症状与高光谱图像上的症状作比较。

1.2 高光谱图像采集

试验所采用的高光谱图像采集系统如图1所示。采用美国OKSI公司的Hycan1211型高光谱相机,其图像分辨率为1 620像素×2 325像素,光谱

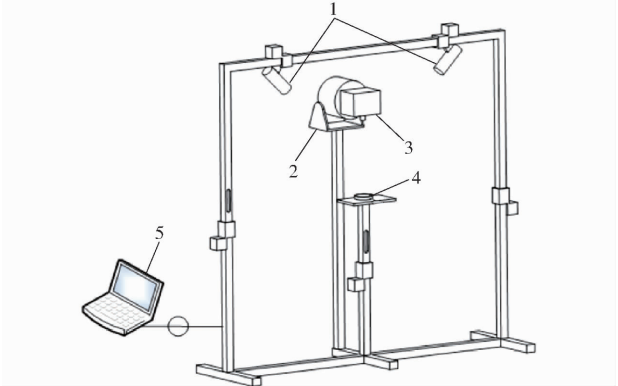


图1 高光谱图像采集系统

Fig. 1 Hyperspectral image acquisition system

1.50 W 卤素灯 2.精密云台 3.高光谱相机 4.样本 5.便携式计算机

分辨率为 1.79 nm, 波长范围为 400 ~ 1 000 nm, 光谱采集点有 339 个波段。

拍摄环境封闭, 仪器被置于用遮光布搭建的棚中, 拍摄时除了光源外无其他任何干扰。同时, 为了确保稳定的光源照射强度, 每次采集高光谱图像之前将高光谱仪预热 30 min, 然后在计算机上对配套软件进行参数设置以保证采集到的图像清晰且不失真: 扫描角度设置为 $-10^{\circ} \sim 10^{\circ}$, 扫描速度设置为 $1.844(^{\circ})/\text{s}$, 镜头焦距为 60 mm, 相机曝光时间为 10 ms。图 2 为利用高光谱仪器采集的高光谱图像样例。

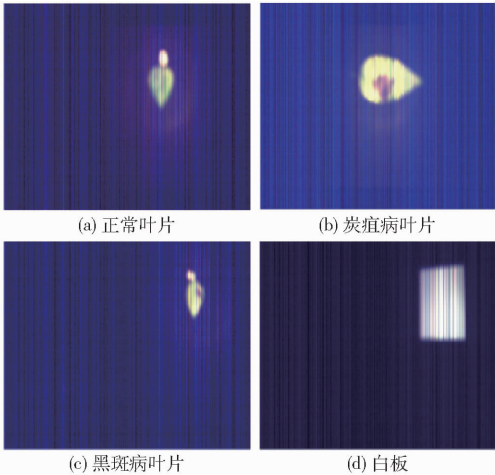


图2 高光谱图像
Fig.2 Hyperspectral images

为了减少试验过程中外界自然光、暗箱内照明以及采集系统本身暗电流对高光谱图像质量的影响, 需要进行黑白板图像校正。黑白板标定公式为

$$R = \frac{I - B}{W - B}$$

式中 R ——校正后图像 I ——原始图像
 W ——白板图像 B ——全黑图像

同时为了进一步从光谱图像上确定对应的病害区域, 用相机拍摄发病叶片每天的状态, 进行对照, 分辨率在 70ppi 以上。拍摄背景统一使用单一白色, 拍摄距离为 15 ~ 20 cm, 图像保存为 JPEG 格式, 采集的图像样例如图 3 所示。

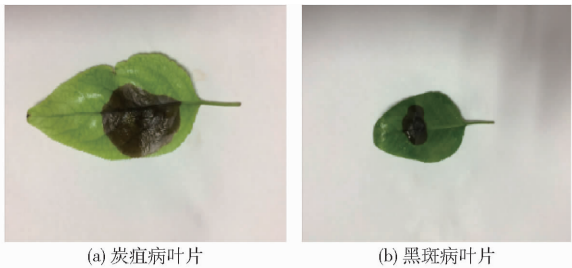


图3 病害叶片图像样例

Fig.3 Sample images of diseased leaves

1.3 高光谱图像信息提取及预处理

1.3.1 信息提取

基于 ENVI 4.7 软件对高光谱图像进行操作, 操作过程主要分为 2 部分: ①提取感兴趣区域, 获取病害区域与健康叶片的平均光谱反射率, 通过光谱反射率曲线从光谱角度找差别并建立模型, 从而进行炭疽病、黑斑病与健康叶片的高精度分类。②对所有的图像进行主成分分析, 一方面可以对高光谱数据进行数据降维, 获取每个感兴趣区域的数据矩阵, 另一方面可以通过计算主成分贡献率获取主成分图像与特征波长。

高光谱图像感兴趣区域提取的具体操作: 首先, 在 ENVI 4.7 软件中打开需处理的梨叶片高光谱图像 (图 4a); 然后, 采用 ROI TOOL 多边形感兴趣区域选择工具分割出感兴趣区域, 如图 4b 所示, 即叶片中形如圆形斑点所示的发病区域; 将感兴趣区域的数据保存为 .txt 文件和 .sta 文件, .txt 和 .sta 文件中都包括了 400 ~ 1 000 nm 高光谱图像 339 个波段, 曲线中间一条白色曲线即为 DN 值 (像素值) 的平均值; 用 Excel 将 .sta 文件与 .txt 文件打开, 获取到感兴趣区域的相关信息后, 通过 DN 平均值数据, 并根据黑白标定公式计算出感兴趣区域的平均光谱反射率; 最后, 得到有效光谱反射率图。

1.3.2 信息预处理

为了提高后期建模效率, 使得光谱变化轮廓更加清楚并提高光谱数据的信噪比, 需在建模前对提取出的高光谱图像感兴趣区域的平均反射率数据进行预处理。剔除不符合条件的曲线后, 获得可供建模的光谱反射曲线共计 440 条, 如图 5a 所示, 叶片平均光谱反射率如图 5b 所示。从整个光谱反射率曲线可以观察到, 健康叶片、炭疽病叶片与黑斑病叶片的光谱反射曲线整体走势一致, 差别在于反射率的大小。

获得原始曲线后, 采用多元散射校正 (MSC)、标准正态变换 (SNV) 和 Savitzky-Golay (S-G) 卷积平滑法分别对其进行预处理, 预处理后反射率见图 6。由图 6 可知, MSC 减少了样本之间发生基线偏移或平移现象, 最大限度地保留样本中与化学成分相关的光谱吸收信息; S-G 法对光谱数据进行平均计算并重新分配误差, 对每个数据点进行处理, 去除光谱数据中的高频噪声并保留有用的低频信息, 不存在波峰与波谷的区别, 这样的优点是可以让光谱曲线发生倾斜偏移与线性平移; 与 MSC 算法相比, SNV 算法的不同之处在于单独对每个样本的光谱进行校正, 并且不需要理想光谱, 但是每个波段吸光度均符合正态分布^[20]。

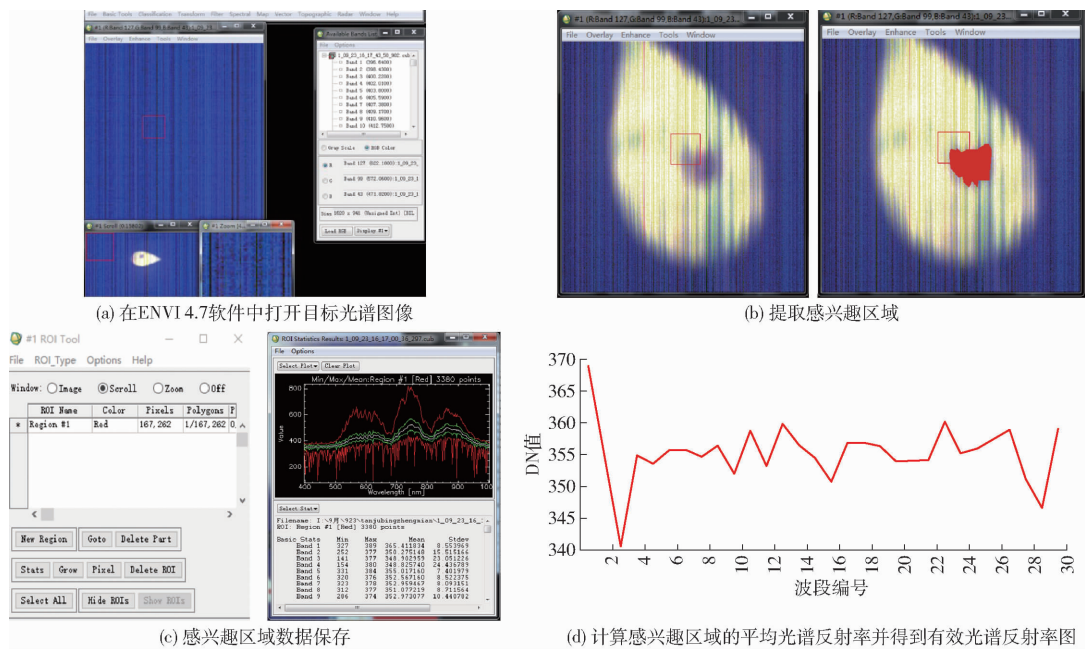


图 4 提取感兴趣区域流程图

Fig. 4 Flow chart of extracting region of interest

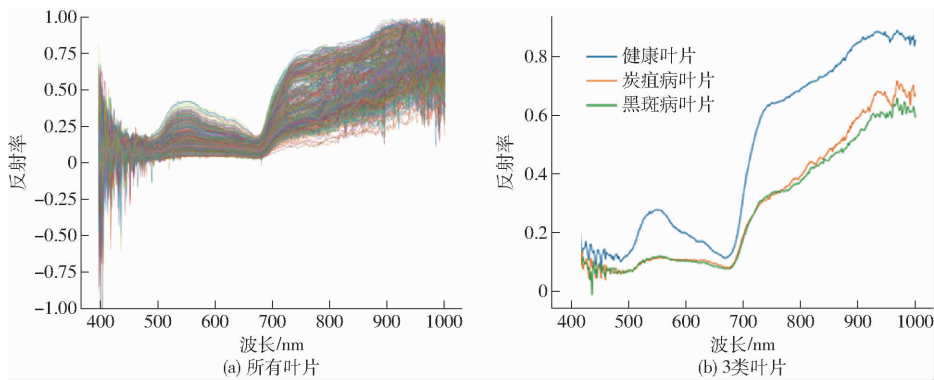


图 5 叶片光谱反射率

Fig. 5 Leaf spectral reflectance diagrams

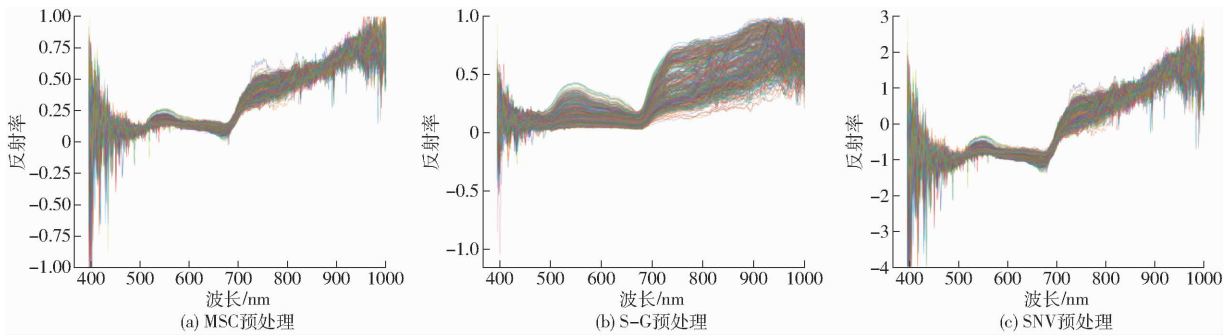


图 6 不同算法预处理后所有叶片原始平均光谱反射率

Fig. 6 Original average spectral reflectance of all leaves after preprocessing by different algorithms

3 种方法虽然具体运算过程不同,但都间接提高了原始光谱数据的信噪比及后期的建模效率。其中,基于 SNV 法的预处理效果最佳。

1.4 特征波长提取

通过高光谱成像技术得到的数据信息量丰富,但冗余信息也相对较多,继而对后期建模效率、模型精度等会产生一系列影响,因此需从原有光谱中提

取与样本化学成分相关的波段和波长来进行降维处理。本试验采取的特征提取方法有主成分分析法 (PCA)、连续投影算法 (SPA)、无信息变量消除法 (UVE)、竞争性自适应重加权算法 (CARS) 和随机蛙跳算法 (SFLA)。PCA 算法需要选择贡献率大的波长,且要求累计贡献率之和大于 85%;SPA 算法利用少数几列光谱概括大部分原始光谱的变量信

息^[21-22]；UVE 算法可以去除光谱中含有的较多噪声,优选出特征波长；CARS 算法不仅可以有效去除无信息变量,而且还可以最大程度地减少共线性变量对模型的影响^[23]；SFLA 算法具有遗传算法和 PSO 算法的优势,在数据降维中具有重要的作用^[24]。

1.5 识别模型建立

1.5.1 基于特征波段的支持向量机(SVM)分类识别建模

SVM 模型的核心思想是通过找出边际最大的决策边界 $\omega x + b = 0$,对数据进行分类。本研究中 SVM 算法选取的核函数为径向基核函数(Radical basis function, RBF)。研究通过调整惩罚参数 c 及核函数参数 g 来提升模型的泛化性能,从而使得模型有更好的预测效果。

具体步骤如下：

(1) 选定建模集与测试集

本试验所采集的样本分为两部分：①采集的样本包括炭疽病与黑斑病病害区域的平均光谱反射率,以这部分样本的高光谱数据作为测试集。②所获取的样本经过了从叶片健康状态到发病状态且病斑从小到大的连续动态过程,样本量大,以这部分采集的高光谱数据作为建模集。

(2) 数据归一化预处理

SVM 网络训练对建模集与测试集都进行归一化处理,以提高寻找最优解的效率,保证程序在运行时收敛加快。

(3) 参数寻优

本次建模试验以准确率为评价指标,通过十折交叉验证,选取建模效果最优的 c 和 g ,有效避免模型欠拟合现象的发生,同时保证了模型的泛化能力。

1.5.2 BP 神经网络分类识别建模

BP 神经网络由输入层、隐含层与输出层组成,各层有若干个节点,层与层之间直接通过权重来连接。核心步骤分为正向传播与反向传播。反向传播依靠学习率 η 和梯度更新权重。本研究搭建的 BP 神经网络隐含层设置为 1 层,激活函数为 tanh,输出层的激活函数为 softmax。

1.6 模型效果评测

用测试集验证模型在未知样本上的表现,所建的各 SVM 模型以分类准确率作为评价指标。对各 BP 模型以分类准确率、召回率、F1 值为评价指标进行综合评价。准确率即从模型预测角度出发,表示在预测为 $i(i = 1, 2, 3)$ 的样本中,预测正确的比率。召回率是从样本的角度出发,表示的是标签为 $i(i = 1, 2, 3)$ 的样本中,被正确预测的比率。F1 值为准确

率与召回率的调和平均值。

2 结果与分析

2.1 样本划分

样本的平均光谱反射曲线共 440 条,分别编号 1~440。具体划分结果如表 1 所示,用标签 1、2、3 分别代表健康叶片、炭疽病叶片和黑斑病叶片。类别 1 健康叶片所对应的编号为 1~140,建模集曲线 100 条,测试集曲线 40 条；类别 2 炭疽病叶片所对应的编号为 141~290,建模集曲线 110 条,测试集曲线 40 条；类别 3 黑斑病叶片所对应的编号为 291~440,建模集曲线 110 条,测试集曲线 40 条。

表 1 砀山酥梨叶片样本光谱划分

Tab.1 Spectral division of ‘Dangshan’ pear leaf samples

叶片状态	健康叶片	炭疽病叶片	黑斑病叶片
类别标签	1	2	3
建模集(样本数量)	100	110	110
测试集(样本数量)	40	40	40

2.2 梨叶片近红外特征波长提取结果

PCA、SPA、UAE、CARS、SFLA 法提取的特征波长如表 2 所示。

(1)PCA 法。对砀山酥梨健康叶片、炭疽病叶片和黑斑病叶片光谱数据求取各变量的标准差并进行标准差变换,然后计算各主成分的累计贡献率,并对加权系数加权平均,经过内部的交叉验证,将 339 个原始光谱波段压缩为 27 个主成分。

(2)SPA 法。原始光谱数据是一个 440×339 的光谱矩阵,将最小特征波长数设置为 15,任意选择一列向量,计算该列向量在剩余所有列向量上的投影,输出最大投影对应的列向量序号,当均方根误差(RMSECV)最小时,对应的波长即特征波长。

(3)UVE 法。经过 UVE 处理之后,共有 15 个被认为是有用信息的特征波长。

(4)CARS 法。CARS 是利用蒙特卡洛方法采样,当程序运行 100 次时,采样次数为 56,从中选择 44 个样本建立偏最小二乘模型并计算该模型的回归系数的绝对值与各回归系数的权重,当最优变量子集确定的十折交叉验证法均方根误差(RMSECV)最小值等于 1.27 时,提取的特征波长数有 26 个。

(5)SFLA 法。利用 SFLA 法可以计算出每个变量被选择的概率,经过之前对梨树叶片光谱数据的分析,本试验中将变量被选择的概率阈值设置为 0.8,经过程序运行,提取出的最优特征波长为 20 个。

5 种方法提取的特征波长多集中在 400 ~

表 2 不同算法提取的特征波长

Tab.2 Characteristic wavelength selected by different algorithms

算法	特征波长数	特征波长序号	对应的特征波长/nm
PCA	27	1,2,4,6,7,8,9,10,13,16,27,73,92,96,99,	401.77,403.54,407.08,410.62,412.39,414.16,415.93,417.75,
		102,104,110,119,131,141,167,179,196,	423.01,428.32,447.79,529.21,562.84,569.92,575.23,580.54,
SPA	12	202,221,230	594.75,610.63,631.87,649.57,695.59,716.83,746.92,757.54,
			791.17,792.94,826.57
UVE	15	1,2,5,8,10,13,15,18,19,27,32,280	401.77,403.54,408.85,414.16,417.7,423.01,426.55,431.86,
			433.63,447.79,456.64,895.6
CARS	26	80,96,104,111,114,119,146,189,204,231,	541.62,569.92,584.08,596.47,601.78,610.63,658.42,734.53,
		247,264,286,299,331	761.08,808.87,837.19,867.28,906.22,929.23,985.87
SFLA	20	5,14,30,31,35,48,50,68,71,84,85,86,91,	408.85,424.78,451.3,454.87,461.95484.96,488.5,520.36,
		169,174,176,237,279,284,295,310,317,	525.67,539.83,548.68,550.45,552.22,561.07,612.4,699.13,
		323,327,328,331	707.98,711.52,819.49,893.83,922.15,931,948.7,971.71,
			978.79,980.56,985.87
		50,67,74,76,77,84,94,101,111,114,116,	488.52,518.59,530.98,534.52,536.29,548.68,566.38,578.77,
		117,130,137,188,218,237,240,264,288	596.47,601.78,605.32,607.09,630.12,642.49,732.76,785.86,
			819.49,824.85,877.28,909.76

550 nm 间,表明病变叶片内部成分变化较为明显的多集中响应在此波段范围内。

2.3 基于特征波长的 SVM 模型分类结果

各模型分类准确率如表 3 所示。

表 3 各 SVM 模型分类预测准确率

Tab.3 Prediction results of SVM models

算法	特征波长数	准确率/%	
		建模集	测试集
PCA	27	93.71(300/320)	90.83(109/120)
CARS	26	91.43(292/320)	88.75(106/120)
SPA	12	94.80(303/320)	93.25(111/120)
UVE	15	87.14(279/320)	86.23(103/120)
SFLA	20	85.71(274/320)	84.26(100/120)

注:括号内斜线前数据为分类正确的样本数,斜线后数据为建模集或测试集样本总数。下同。

将 PCA、CARS、SPA、UVE 和 SFLA 算法所选取的特征波长作为 SVM 支持向量机的输入变量,本试验中 SVM 所应用的核函数为 RBF,其中参数 g 与 c 的最优取值均采用十折交叉验证法实现。参数 g 即函数自带参数 gamma,决定了数据映射到新的特征空间后的分布,gamma 值影响支持向量机的数量。参数 c 即惩罚参数,即对误差的宽容度。因此,PCA-SVM、CARS-SVM、SPA-SVM、UVE-SVM 和 SFLA-SVM 模型中的最优 c 值均为 16,最优 g 值均为 0.062 5。

(1)PCA-SVM 测试集的实际分类和预测分类结果见图 7a。将这 27 个特征波长作为输入变量,并对 c 和 g 划分网格进行搜索,采用十折交叉验证方法,得到最佳的 c 值为 16,最佳的 g 值为 0.062 5,最佳的 RMSECV,即得到的所有分类准确率的平均

数为 87.925 7。总支持向量数为 137,每类样本支持向量数为 115,建模集分类识别率为 93.71%,测试集样本分类识别率为 90.83%,均方根误差为 0.205 1,预测效果良好。

(2)SPA-SVM 测试集的实际分类和预测分类结果见图 7b。最佳的 RMSECV 为 87.735 8。总支持向量数为 146,每类样本的支持向量数为 123,建模集识别率为 94.80%,120 个预测样本中正确分类的个数为 111,测试集识别率为 93.25%,均方根误差为 0.154 7。建模集和测试集的识别率均达到了 90% 以上,说明 SPA-SVM 具有较好的预测效果。

(3)UVE-SVM 测试集的实际分类和预测分类结果见图 7c。得到惩罚参数 c 为 16,核函数参数 g 为 0.062 5,最佳的 RMSECV 为 80.625 0。总支持向量数为 236,每类样本的支持向量数为 198,建模集分类识别率为 87.14%,测试集样本分类识别率为 86.23%,均方根误差为 1.175 0。与 PCA-SVM 和 SPA-SVM 模型相比,UVE-SVM 模型的优点在于准确无误地将健康叶片样本识别出来。

(4)CARS-SVM 测试集的实际分类和预测分类结果见图 7d。得到惩罚参数 c 为 16,核函数参数 g 为 0.062 5,最佳的 RMSECV 为 86.525 6。总支持向量数为 155,每类样本支持向量数为 129。最终得到的建模集样本分类识别率为 87.14%,测试集样本分类识别率为 86.23%,均方根误差为 0.158 3。

(5)SFLA-SVM 测试集的实际分类和预测分类结果见图 7e。得到惩罚参数 c 为 16,核函数参数 g 为 0.062 5,最佳的 RMSECV 为 81.25。总支持向量数为 238,每类样本支持向量数为 203。建模集分类识别率为 87.14%。

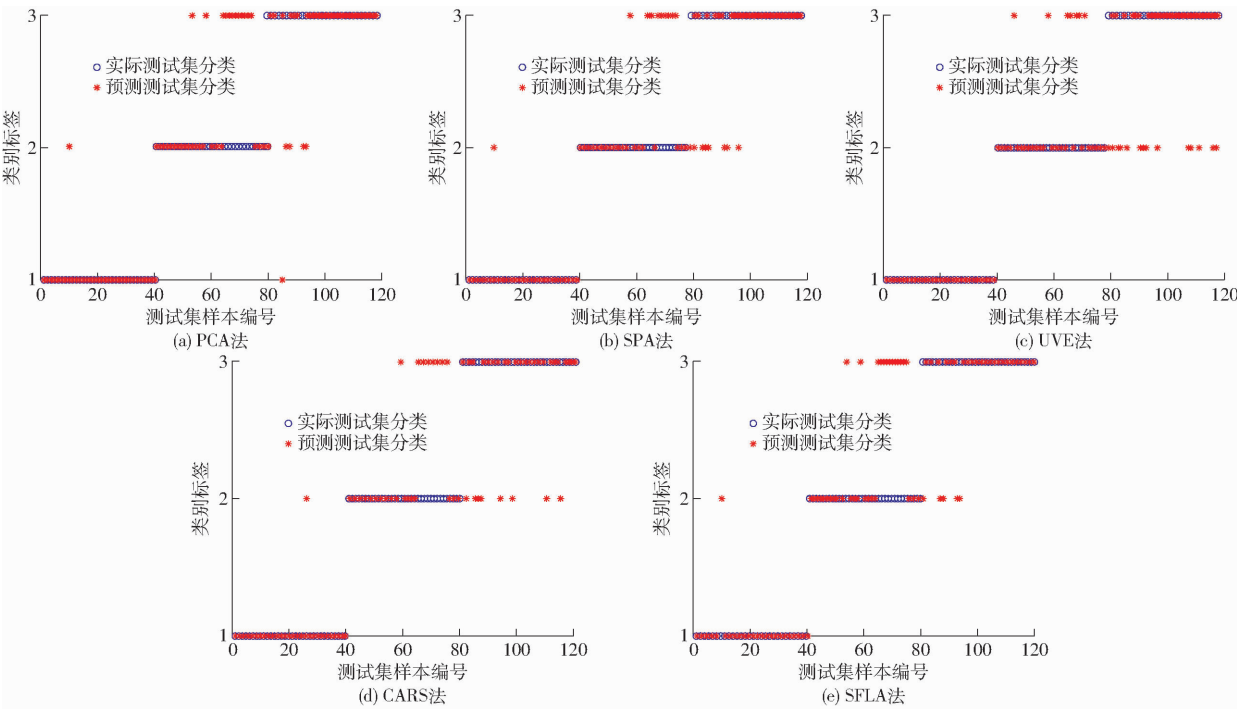


图 7 基于不同算法的 SVM 模型测试集分类结果

Fig. 7 Classification results of SVM model test set based on different algorithms

当实际测试集分类图例与预测测试集分类图例相重合时,分类结果即为正确,反之则说明模型分类错误。通过所有测试集的实际分类与预测分类图发现,SVM 支持向量机几乎可以识别所有的正常叶片,分类误差较大的为黑斑病叶片与炭疽病叶片。

2.4 基于特征波长的 BP 模型分类结果

各模型分类准确率如表 4 所示。

表 4 各 BP 模型分类预测准确率

Tab. 4 Prediction results of BP models

算法	特征波长数	准确率/%	
		建模集	测试集
PCA	27	85. 94 (275/320)	73. 33 (88/120)
CARS	26	84. 69 (271/320)	79. 17 (95/120)
SPA	12	86. 88 (278/320)	76. 67 (92/120)
UVE	15	78. 44 (251/320)	75. 83 (91/120)
SFLA	20	79. 69 (255/320)	75. 00 (90/120)

(1)PCA – BP 测试集的实际分类和预测分类结果见图 8a。此模型的学习率为 0. 1,网络结构为 27 – 20 – 3(输入层有 27 个节点,隐藏层有 20 个节点,输出层有 3 个节点)。此模型几乎可以识别所有的正常叶片,但在炭疽病识别上,表现能力欠佳,在测试集 40 个炭疽病叶片样本中,有 29 片被误判为黑斑病叶片。

(2)SPA – BP 测试集的实际分类和预测分类结果见图 8b。此模型学习率为 0. 1,网络结构为 12 – 50 – 3。SPA – BP 建模集准确率为 86. 88%,为 5 个

模型中最佳,且可以识别所有的正常叶片,误差集中在 2 处:将 16 片炭疽病叶误判成黑斑病叶片,将 9 片黑斑病叶片误判成炭疽病叶片。

(3)UVE – BP 测试集的实际分类和预测分类结果见图 8c。此模型学习率为 0. 2,网络结构为 15 – 16 – 3。UVE – BP 在识别正常叶片方面表现出色,38 片正常叶片分类正确,仅将 2 片正常叶片误判为黑斑病叶片。在识别炭疽病叶片方面,UVE – BP 将 28 片炭疽病叶片判定正确,但将余下 12 片误判为黑斑病叶片。在黑斑病叶片的识别上,25 片被判定正确,只有 1 片叶片被误判为正常叶片,还有 14 片叶片被误判为炭疽病叶片。

(4)CARS – BP 测试集的实际分类和预测分类结果见图 8d。此模型学习率为 0. 2,网络结构为 26 – 40 – 3。CARS – BP 测试集准确率为 79. 17%,为 5 个模型中最高,且识别出了大部分正常叶片,仅有 6 枚叶片被误判为炭疽病叶片。分类误差较大的集中在炭疽病叶片和黑斑病叶片的识别上,但 CARS – BP 可识别将近 75% 的炭疽病叶片,以及 80% 的黑斑病叶片。

(5)SFLA – BP 测试集的实际分类和预测分类结果见图 8e。此模型学习率为 0. 1,网络结构为 20 – 50 – 3。SFLA – BP 几乎可以识别所有的正常叶片,40 片正常叶片样本中仅有 2 片被误判。SFLA – BP 识别了 75% 的炭疽病叶片,余下的叶片样本被误判为黑斑病叶片。然而此模型在黑斑病叶片样本的识别方面误差较大,40 个样本中有 18 片叶片被

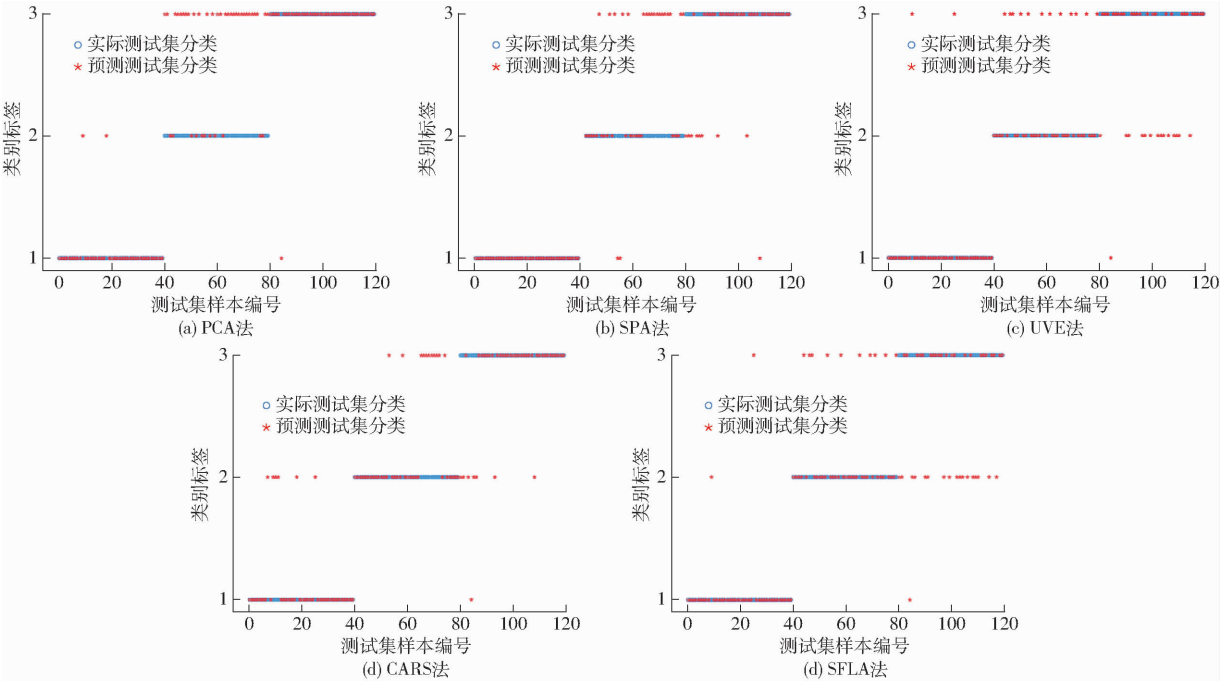


图 8 基于不同算法的 BP 模型测试集分类结果

Fig. 8 Classification results of BP model test set based on different algorithms

误判,且大多被误判为炭疽病叶片。片测试集的分类准确率、召回率、F1 值如表 5 所示,所建的各模型中,健康叶片、炭疽病、黑斑病叶建模集的各评价指标如表 6 所示。

表 5 BP 模型测试集分类预测结果

Tab. 5 Test set classification prediction results of BP models

评价指标	PCA - BP			SPA - BP			UVE - BP			CARS - BP			SFLA - BP		
	健康叶片	炭疽病叶片	黑斑病叶片	健康叶片	炭疽病叶片	黑斑病叶片	健康叶片	炭疽病叶片	黑斑病叶片	健康叶片	炭疽病叶片	黑斑病叶片	健康叶片	炭疽病叶片	黑斑病叶片
准确率/%	97	85	57	93	71	65	97	67	64	97	69	74	97	62	67
召回率/%	95	28	97	100	55	75	95	70	62	85	72	80	95	75	55
F1 值/%	96	42	72	96	62	70	96	68	63	91	71	77	96	68	60
样本量	40	40	40	40	40	40	40	40	40	40	40	40	40	40	40

表 6 BP 模型建模集分类预测结果

Tab. 6 Training set classification prediction results of BP models

评价指标	PCA - BP			SPA - BP			UVE - BP			CARS - BP			SFLA - BP		
	健康叶片	炭疽病叶片	黑斑病叶片	健康叶片	炭疽病叶片	黑斑病叶片	健康叶片	炭疽病叶片	黑斑病叶片	健康叶片	炭疽病叶片	黑斑病叶片	健康叶片	炭疽病叶片	黑斑病叶片
准确率/%	99	92	73	94	82	85	93	71	71	91	78	86	92	75	73
召回率/%	100	65	95	98	81	83	99	68	70	97	82	76	99	72	70
F1 值/%	100	76	83	96	81	84	96	70	70	94	80	81	95	73	71
样本量	100	110	110	100	110	110	100	110	110	100	110	110	100	110	110

2.5 模型效果评估对比

在所建的 5 个 SVM 模型中,测试集准确率大于 85% 的模型为 SPA - SVM、PCA - SVM、CARS - SVM、UVE - SVM。其中基于 SPA 算法优选的特征变量建立的 SVM 模型性能最好,在减少模型变量的同时提高了模型精度。SPA 算法将用于建模的特征波长由 339 个减少到了 12 个,测试集精度达到了 93. 25%,且相对于其它模型,SPA - SVM 可以准确

无误地将健康叶片样本识别出来,表明 SPA 法较其他 4 种特征波长提取方法,更大限度地剔除了噪声数据,降低了相关性小的波长的干扰,同时提高了光谱的表现能力。在所建的各 BP 模型中,建模集准确率最高的模型为 SPA - BP 模型,准确率为 86. 88%,测试集准确率最高的模型为 CARS - BP 模型,准确率为 79. 17%。综上所述,SPA - SVM 模型效果最佳。

经试验发现,所建的 10 个模型分别能识别大部分的梨炭疽病与黑斑病叶片,但由于 2 种病叶的表现症状过于相似,导致识别结果存在或多或少的偏差。在实际生产中,可以将模型结果与肉眼观察到的果实发病特征相结合进行对比分析,从而进一步提高 2 种叶片识别的准确率:梨炭疽病果实表面有褐色的病斑,明显下陷,软腐,中央有大量轮纹状排列隆起的黑色小粒点,即病菌分生孢子盘,且潮湿时呈绯红色黏液形式从中溢出;患有黑斑病的梨果果面通常出现一至数个黑色斑点,略凹陷,随着时间的推移,颜色变浅,形成浅褐至灰褐色圆形病斑,且发病后期病果畸形、龟裂,裂缝可深达果心,果面和裂缝内产生黑霉,并常常引起落果^[25-27]。

3 结论

(1)采用近红外高光谱成像技术,获取目标样本的光谱数据,用 ENVI 4.7 软件提取砀山酥梨病叶的兴趣区域并进行处理,再用 MSC、S-G 平滑法、SNV 分别对原始图像进行预处理。相较而言,基于 SNV 法的预处理效果最佳。

(2)基于 PCA、SPA、UVE、CARS 和 SFLA 分别提取了 27、12、15、26、20 条特征波长,5 种方法均有效剔除了与建模无相关性或相关性小的波段,提高了信噪比并提升了后期建模精度。

(3)在病害识别模型建立的方法上,一方面,基于特征波长并利用支持向量机进行数据建模,在基于 5 种方法所建分类模型中,SPA-SVM 模型识别效果最佳,其建模集准确率为 94.80%,测试集准确率为 93.25%,均方根误差为 0.1547。另一方面,BP 神经网络是完全不同于 SVM 的识别分类方式,CARS-BP 模型的测试集准确率最高(79.17%)。结果表明,近红外高光谱技术可以较好地应用在梨炭疽病和黑斑病叶片的识别检测中,且利用高光谱技术可以为识别砀山酥梨叶片黑斑病与炭疽病提供在线技术支持与理论依据。

(4)由于梨炭疽病叶片和黑斑病叶片内部物质元素的变化,借助近红外高光谱技术,2 种病叶能在相应波段呈现出不同的特点,因此 2 种病叶能被区分开,叶片识别率较人工分辨显著提高,大大改善了凭肉眼难以区分 2 种病叶的问题。

参 考 文 献

[1] 张绍铃,谢智华. 我国梨产业发展现状、趋势、存在问题与对策建议[J]. 果树学报,2019,36(8): 1067-1072.
ZHANG Shaoling, XIE Zhihua. Current status, trends, main problems and the suggestions on development of pear industry in China[J]. Journal of Fruit Science, 2019, 36(8): 1067-1072. (in Chinese)

[2] 吴良庆,朱立武,衡伟,等. 砀山梨炭疽病病原鉴定及其抑菌药剂筛选[J]. 中国农业科学,2010,43(18): 3750-3758.
WU Liangqing, ZHU Liwu, HENG Wei, et al. Identification of Dangshan pear anthracnose pathogen and screening fungicides against it[J]. Scientia Agricultura Sinica, 2010, 43(18): 3750-3758. (in Chinese)

[3] 王转卫,迟茜,郭文川,等. 基于近红外光谱技术的发育后期苹果内部品质检测[J]. 农业机械学报,2018,49(5): 348-354.
WANG Zhuanwei, CHI Qian, GUO Wenchuan, et al. Internal quality detection of apples during late developmental period based on near-infrared spectra technology[J]. Transactions of the Chinese Society for Agricultural Machinery, 2018, 49(5): 348-354. (in Chinese)

[4] 王凡,彭彦昆,汤修映,等. 樱桃番茄可溶性固形物含量的可见/近红外透射光谱无损检测[J]. 中国食品学报,2018,18(10): 235-240.
WANG Fan, PENG Yankun, TANG Xiuying, et al. Near infrared nondestructive testing of soluble solids content of cherry tomato[J]. Chinese Journal of Food Science, 2018, 18(10): 235-240. (in Chinese)

[5] POURDARBANI R, SABZI S, KALANTARI D, et al. Non-destructive visible and short-wave near-infrared spectroscopic data estimation of various physicochemical properties of Fuji apple (*Malus pumila*) fruits at different maturation stages[J]. Chemometrics and Intelligent Laboratory Systems, 2020, 206: 104-147.

[6] 刘鸿飞,黄敏敏,赵旭东,等. 基于机器视觉的温室番茄裂果检测[J]. 农业工程学报,2018,34(16): 170-176.
LIU Hongfei, HUANG Minmin, ZHAO Xudong, et al. Detection of cracking tomato based on machine vision in greenhouse[J]. Transactions of the CSAE, 2018, 34(16): 170-176. (in Chinese)

[7] DHAKSHINA K S, ESAKKIRAJAN S, BAMA S, et al. A microcontroller based machine vision approach for tomato grading and sorting using SVM classifier[J]. Microprocessors and Microsystems, 2020, 76: 103090.

[8] 邵园园,王永贤,玄冠涛,等. 基于高光谱成像的肥城桃品质可视化分析与成熟度检测[J]. 农业机械学报,2020,51(8): 344-350.
SHAO Yuanyuan, WANG Yongxian, XUAN Guantao, et al. Visual detection of SSC and firmness and maturity prediction for Feicheng peach by using hyperspectr imaging[J]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(8): 344-350. (in Chinese)

[9] GOWEN A A, O'DOMMELL C P, CULLEN P J, et al. Hyperspectral imaging-an emerging process analytical tool for food quality and safety control[J]. Trends in Food Science & Technology, 2007, 18(12): 590-598.

[10] 刘燕德,程梦杰,郝勇. 光谱诊断技术及其在农产品质量检测中的应用[J]. 华东交通大学学报,2018,35(4): 1-7.
LIU Yande, CHENG Mengjie, HAO Yong. Application of spectral diagnoses technology in determination of agricultural

- products quality[J]. Journal of East China Jiaotong University, 2018, 35(4): 1–7. (in Chinese)
- [11] SIRIPATRAWAN U, MAKINO Y. Simultaneous assessment of various quality attributes and shelf life of packaged bratwurst using hyperspectral imaging[J]. Meat Science, 2018, 146:26–33.
- [12] 於海明,徐佳琪,刘浩鲁,等. 基于高光谱和频谱特征的注水肉识别方法[J]. 农业机械学报,2019,50(11):367–372,366.
- YU Haiming, XU Jiaqi, LIU Haolu, et al. Identification of water injection meat based on hyperspectral technique and spectrum characteristics[J]. Transactions of the Chinese Society for Agricultural Machinery, 2019, 50(11): 367–372,366. (in Chinese)
- [13] 霍迎秋,张晨,李宇豪,等. 高光谱图像结合机器学习方法无损检测猕猴桃[J]. 中国农机化学报,2019,40(4): 71–77.
- HUO Yingqiu, ZHANG Chen, LI Yuhao, et al. Nondestructive detection for kiwifruit based on the hyperspectral technology and machine learning[J]. Journal of Chinese Agricultural Mechanization, 2019, 40(4): 71–77. (in Chinese)
- [14] 刘德华,张淑娟,王斌,等. 基于高光谱成像技术的山楂损伤和虫害缺陷识别研究[J]. 光谱学与光谱分析,2015,35(11): 3167–3171.
- LIU Dehua, ZHANG Shujuan, WANG Bin, et al. Detection of hawthorn fruit defects using hyperspectral imaging[J]. Spectroscopy and Spectral Analysis, 2015, 35(11): 3167–3171. (in Chinese)
- [15] 王新忠,卢青,张晓东,等. 基于高光谱图像的黄瓜种子活力无损检测[J]. 江苏农业学报,2019,35(5):1197–1202.
- WANG Xinzhong, LU Qing, ZHANG Xiaodong, et al. Non-destructive detection of cucumber seeds vigor based on hyperspectral imaging[J]. Journal of Jiangsu Agricultural Sciences, 2019, 35(5): 1197–1202. (in Chinese)
- [16] ELMASRY G, WANG N, VIGNEAULT C, et al. Early detection of apple bruises on different background colors using hyperspectral imaging[J]. LWT—Food Science and Technology, 2008, 41(2):337–345.
- [17] MEHL P M. Detection of defects on selected apple cultivars using hyperspectral and multispectral image analysis[J]. Applied Engineering in Agriculture, 2002, 18(2):219–226.
- [18] 吴龙国,王松磊,康宁波,等. 基于高光谱成像技术的灵武长枣缺陷识别[J]. 农业工程学报,2015,31(20): 281–286.
- WU Longguo, WANG Songlei, KANG Ningbo, et al. Non-destructive detection of defects in *Zizphus jujube* Mill cv. *Lingwu changzao* based on near-infrared hyperspectral imaging[J]. Transactions of the CSAE, 2015, 31(20): 281–286. (in Chinese)
- [19] 李勋兰,易时来,何绍兰,等. 高光谱成像技术的柚类品种鉴别研究[J]. 光谱学与光谱分析,2015,35(9): 2639–2643.
- LI Xunlan, YI Shilai, HE Shaolan, et al. Identification of pummelo cultivars based on hyperspectral imaging technology[J]. Spectroscopy and Spectral Analysis, 2015,35(9): 2639–2643. (in Chinese)
- [20] 陈华舟,潘涛,陈洁梅. 多元散射校正与 Savitzky–Golay 平滑模式的组合优选应用于土壤有机质的近红外光谱分析[J]. 计算机与应用化学,2011,28(5):518–522.
- CHEN Huazhou, PAN Tao, CHEN Jiemei. Combination optimization of multiple scatter correction and Savitzky–Golay smoothing modes applied to the near infrared spectroscopy analysis of soil organic matter[J]. Computers and Applied Chemistry, 2011, 28(5): 518–522. (in Chinese)
- [21] 罗耀华,郭科,赵仕波. 高光谱遥感主成分分析算法并行计算研究[J]. 国土资源科技管理,2012,29(6): 75–78.
- LUO Yaohua, GUO Ke, ZHAO Shibao. Principal component analysis of hyperspectral remote sensing in parallel computing based on CUDA[J]. Science and Technology Management of Land and Resources, 2012, 29(6): 75–78. (in Chinese)
- [22] 陈斌,孟祥龙,王豪. 连续投影算法在近红外光谱校正模型优化中的应用[J]. 分析测试学报,2007,26(1):66–69.
- CHEN Bin, MENG Xianglong, WANG Hao. Application of successive projections algorithm in optimizing near infrared spectroscopic calibration model[J]. Chinese Journal of Instrumental Analysis, 2007, 26(1):66–69. (in Chinese)
- [23] 王海龙,杨国国,张瑜,等. 竞争性自适应重加权算法和相关系数法提取特征波长检测番茄叶片真菌病害[J]. 光谱学与光谱分析,2017,37(7): 2115–2119.
- WANG Hailong, YANG Guoguo, ZHANG Yu, et al. Detection of fungal disease on tomato leaves with competitive adaptive reweighted sampling and correlation analysis methods[J]. Spectroscopy and Spectral Analysis, 2017, 37(7): 2115–2119. (in Chinese)
- [24] 章海亮,罗微,刘雪梅,等. 应用遗传算法结合连续投影算法近红外光谱检测土壤有机质研究[J]. 光谱学与光谱分析,2017,37(2):584–587.
- ZHANG Hailiang, LUO Wei, LIU Xuemei, et al. Measurement of soil organic matter with near infrared spectroscopy combined with genetic algorithm and successive projection algorithm[J]. Spectroscopy and Spectral Analysis, 2017, 37(2): 584–587. (in Chinese)
- [25] 冯国杰,李志涛,李宏,等. 梨树果实主要病害的发生与防治[J]. 现代农村科技,2013(11):24–25.
- FENG Guojie, LI Zhitao, LI Hong, et al. Occurrence and control of main diseases of pear fruit[J]. Modern Rural Science and Technology, 2013(11): 24–25. (in Chinese)
- [26] 高正辉,张金云,伊兴凯,等. 砀山酥梨炭疽病发生特征与防治技术[J]. 安徽农业科学,2010,38(19):10445–10446.
- GAO Zhenghui, ZHANG Jinyun, YI Xingkai, et al. Occurrence characteristics and mixing technology of anthracnose in Dangshan Su pear[J]. Anhui Agricultural Sciences, 2010, 38(19): 10445–10446. (in Chinese)
- [27] 张志铭,宋福,孙淑贞,等. 河北鸭梨黑斑病原菌的鉴定[J]. 植物检疫,2003,17(4):212–214.
- ZHANG Zhiming, SONG Fu, SUN Shuzhen, et al. Identification of the pathogen of black spot of Yali pear in Hebei[J]. Plant Quarantine, 2003, 17(4):212–214. (in Chinese)