

柑橘黄龙病光谱特征波段选择及光谱检测仪研制

兰玉彬^{1,2,3,4}, 王天伟¹, 郭雅琦¹, 杨冬子¹, 林韶明¹, 胡宇琦¹, 林晓晴¹,
邓小玲^{1,2,3,4*}

(1. 华南农业大学电子工程学院(人工智能学院), 广州 510642; 2. 国家精准农业航空施药技术国际联合研究中心, 广州 510642;
3. 岭南现代农业科学与技术广东省实验室, 广州 510642; 4. 广东省智慧农业工程技术中心, 广州 510642)

摘要: 黄龙病(Huanglongbing, HLB)被称为柑橘的癌症, 及早检测出患病植株可防止病情蔓延, 降低病情灾害程度。高光谱分析技术因其丰富的光谱信息, 成为近年来作物病害检测的研究热点。然而高光谱设备昂贵, 波段数较多, 计算量大, 在实际应用中尚未形成规模应用。使用合理的波段选择方法, 可以去掉冗余信息, 避免“维数灾难”, 减轻数据存储、计算与传输压力, 并降低设备成本。该研究利用地物谱仪获取了柑橘冠层叶片的高光谱信息, 提出一种基于典型成分分析(Exemplar Component Analysis, ECA)的柑橘黄龙病特征波段优选方法, 并与其他3种波段优选算法进行比较, 分别优选了7个光谱波段的组合。基于优选波段, 采用6种机器学习方法进行建模分类, 对4种波段选择方法的鲁棒性进行了分析。此外, 基于优选的特征波段设计了一款多光谱仪应用于柑橘黄龙病的检测。结果表明, 用ECA算法选择的特征波段, 其结合6种分类器在测试集上的准确率达到92%以上, 并具有较好的鲁棒性。自研基于特征波段的多光谱仪对于HLB的检测精确度最高可达95%。试验表明用少量特征波段表征HLB作为检测手段具有可行性, 合理的特征波段有助于降低专门农业病害光谱检测的设计成本, 提高果园病情防控精准度。

关键词: 高光谱; 遥感; 智能检测; 柑橘黄龙病; 光谱仪; 波段选择

doi: 10.11975/j.issn.1002-6819.2022.20.014

中图分类号: S127; S157.9

文献标志码: A

文章编号: 1002-6819(2022)-20-0119-10

兰玉彬, 王天伟, 郭雅琦, 等. 柑橘黄龙病光谱特征波段选择及光谱检测仪研制[J]. 农业工程学报, 2022, 38(20): 119-128. doi: 10.11975/j.issn.1002-6819.2022.20.014 http://www.tcsae.org

Lan Yubin, Wang Tianwei, Guo Yaqi, et al. Selection of spectral characteristic bands of HLB disease of citrus and spectrum detector development[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2022, 38(20): 119-128. (in Chinese with English abstract) doi: 10.11975/j.issn.1002-6819.2022.20.014 http://www.tcsae.org

0 引言

黄龙病(Citrus Huanglongbing, HLB)一直以来被视为柑橘的“癌症”, 患病植株表现为生长缓慢、低产、顶梢枯死、逐渐死亡^[1-4]。中国最早在1919年, 于广东潮汕地区发现有类似病害, 当地人观察到染病柑橘新梢及叶片黄化, 故称之为“黄龙病”, 由寄生在韧皮部内的革兰氏阴性细菌引起, 能够侵染柑橘属、枳属、金柑属和九里香等多种芸香科柑橘亚科的近缘属植物^[5]。1995年第13届国际柑橘病毒学家会议决定以柑橘黄龙病^[6]作为该病的正式名称。柑橘黄龙病目前尚无行之有效的药品用于应对, 也不能培养出抗病品种, 因此苗木一旦感染将无法救

治^[7], 感染枝干上的果实在正常采摘前就已脱落, 虽然有部分果实能够停留在枝上, 但果实品质已经发生了变化^[8], 农户往往选择在采摘季过后就将病树连根铲除, 以免病情扩散, 这严重阻碍了柑橘产业的可持续发展。

传统农业中农户依靠经验、目视来判定植株是否患病, 但柑橘黄龙病的症状较为复杂, 生产中常以柑橘树叶片黄化作为典型症状来加以判断, 主要包括均匀黄化、斑驳黄化和缺素黄化3种类型^[9]。但是涝害、环割过重、虫害等引起的生理性黄化亦可呈现均匀黄化; 缺素型黄化到底是由黄龙病引起的还是仅仅因为植株缺少氮、钙、镁、锌、铁、硼等矿质元素导致仅凭观察难以得出结论; 斑驳黄化较为典型, 叶片表现为左右不对称、叶基先黄化、斑块交界不明显、病叶脆而滑这4个特征, 可实际中, 在不同的季节黄龙病症状也会有所变化, 田间复杂的光照情况也会影响人的判断。实验室生化分析能够提供精确的诊断结果, 其中以分子检测技术效果尤为显著^[10-11], 但往往需要投入大量的人力和时间进行样本采集与标注, 后续分析也对相关人员的知识储备有较高要求, 整个流程周期较长, 难以在农户间推广。

光谱技术是一种可以反应物体微观差异的手段, 具有很高的灵敏度, 其检测精度不低于一般的化学方法, 而且在检测速度上比生化方法快很多^[12], 其中高光谱技术能够

收稿日期: 2021-10-28 修订日期: 2022-10-12

基金项目: 广东省引进领军人才项目(2016LJ06G689); 广东省重点领域研发计划项目(2019B020214003); 广州市重点研发计划项目(202103000090); 广东高校重点领域(人工智能)专项项目(2019KZDZX1012); 广东省科技创新战略专项资金项目(pdjh2020b0089); 岭南现代农业实验室科研项目(NT2021009); 国家自然科学基金面上项目(61675003)

作者简介: 兰玉彬, 博士, 教授, 研究方向为精细农业、农业航空施药、农业航空和地面遥感、农业航空导航、智能农业装备和新型传感器等。

Email: ylan@scau.edu.cn

*通信作者: 邓小玲, 博士, 副教授, 研究方向为智慧果园、农业人工智能和无人机遥感。Email: dengxl@scau.edu.cn

获取电磁波谱位于可见光-近红外 VIS-NIR (Visible-Near Infrared) 和短波红外 SWIR (Shortwave Infrared) 范围内的数百个波段光谱信息^[13-14], 其在农业应用中表现出的优秀性能已经得到很多研究者的证实^[15-17], 一些学者用高光谱技术检测柑橘黄龙病并取得了研究进展。Moriya 等^[18]通过轻型飞机搭载高光谱相机获取巴西柑橘种植园的遥感图像, 使用 Spectral Angle Mapper (SAM) 算法对数据进行分类, 检测感染 HLB 植株的准确率为 61.2%。Weng 等^[19]针对处于不同生长季节的柑橘, 提出一种基于高光谱成像结合碳水化合物代谢分析的 HLB 检测方法, 对健康、感染 HLB 和缺乏营养叶片的 3 类精度在凉季、热季和整个生长期分别达到了 90.2%、96.0% 和 92.6%。

由于高光谱具有极高的光谱分辨率, 其数据的众多波段之间, 尤其是相邻或相近波段存在严重的信息冗余, 波段选择^[20]作为一种重要的特征提取技术, 是解决这一问题的有效途径。肖怀春^[21]利用连续投影算法 (Successive Projections Algorithm, SPA) 和主成分分析 (Principal Component Analysis, PCA) 对轻度、中度、重度 3 类黄龙病叶片分别提取了特征波段, 并结合叶片纹理特征构建了黄龙病诊断模型, 其中, 使用 SPA 筛选了 27 个特征波段, 使用 PCA 筛选了 20 个特征波段。但 SPA 和 PCA 本质上是线性特征提取方法, 高光谱数据具有较高的维数, 在特征空间 (Feature space) 中可能具有复杂的非线性结构, 线性特征提取方法难以描述高光谱数据的非线性结构。王克健^[22]以鸡尾葡萄柚为试验对象, 通过分析 HLB 树和健康树的绿色叶片和黄化叶片的可见-近红外高光谱信息筛选出十个特征光谱, 构建了基于特征波段的最小二乘-支持向量机 HLB 预测模型, 对于建模集和预测集中罹病树叶片的识别精度分别达到 97.9% 和 96.6%。Deng 等^[23-24]使用地物光谱仪和无人机搭载高光谱相机, 分别获取了柑橘叶片高光谱数据和柑橘园遥感高光谱图像, 针对叶片光谱数据计算了所有波段的信息熵及其之间的熵距离, 利用后续搜索算法选择出了 13 个特征波段, 利用遗传算法 (Genetic Algorithm, GA) 对高光谱遥感图像进行特征波段选择, 提取了 10 个特征波段。这些试验都是基于专业仪器设备来获取柑橘叶片的光谱信息, 通常在预处理阶段需要手动剔除异常数据和噪声以克服方法本身鲁棒性较低的问题, 且采用的高光谱相机和地物光谱仪均价格不菲, 对技术落实到田间形成了阻碍。

本研究旨在探索一种高效且鲁棒性较好的柑橘黄龙病光谱特征波段选择方法; 并基于少量的优选波段, 研制低成本的柑橘黄龙病光谱检测仪, 以实现实际生产的大规模应用。

1 数据采集与预处理

1.1 试验基地情况

本试验数据主要采集于广东省惠州市博罗县杨村镇井水龙村柑橘黄龙病绿色防控与新栽培模式研发示范基地。该地属南亚热带海洋性季风气候, 生态环境保护完好, 上游有国家小二型的宝溪水库, 灌溉便利土壤肥沃,

适合柑橘果树的种植。基地内种植的柑橘品种为砂糖橘, 由专人对园区进行管理。本次试验采集的光谱采集自于 3 种情况的柑橘树: 健康 (healthy)、确诊 HLB、疑似症状 (suspected, 除去健康和确诊 HLB 之外的缺素但疑似黄龙病症状的大类)。

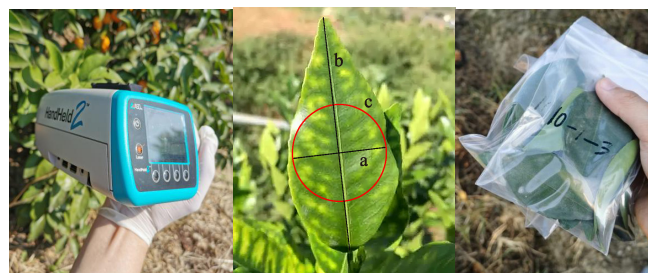


图 1 数据采集园区的柑橘图像

Fig.1 Citrus image of data collection area

1.2 数据采集

试验区域包括 healthy、HLB 和 suspected 3 种柑橘植株, 如图 1 所示。本研究使用英国 ASD FieldSpec HandHeld 2 (下文简称 HH2) 手持式地物光谱仪采集柑橘植株冠层的光谱信息。试验时间为 2021 年 1 月 15 日和 20 日, 当日天空为淡积云量, 无卷云、浓积云等, 光照稳定, 在 11:30—14:30 开展数据采集试验。试验人员身着深色衣物以防止光污染, 采集时避免有阴影落在 HH2 光学接收端 25° 视场角采集范围, HH2 光学接收端与目标叶片的距离为叶片较短直径 (如图 2 中 a, b 为叶片较长直径, c 为 HH2 视场) 的 2 倍, 如图 2 所示。每颗植株上随机选取 4 个点进行数据采集, 每采集完一棵树的光谱数据, 将地物光谱仪对准标准参考白板进行定标校准, 得到接近 100% 的基线, 然后再采集下一棵树, 依次进行试验。



a. 光谱数据采集 a. Spectral data collection b. HH2 使用说明 b. Instructions for use of HH2 c. 叶片样本采集 c. Leaf sample collection

图 2 数据采集过程

Fig.2 Data collection process

为确保试验样本真实病况与后期分析具有一致性, 在采集光谱数据后, 将对应区域的叶片进行采集放于样本袋中, 同时用记号笔在袋上标记, 使样本与光谱数据一一对应, 平均每棵树采集 10 片叶片。所采集的叶片样本存放在备有冰块盒子中, 试验结束后进行了聚合酶链反应 (Polymerase Chain Reaction, PCR) 检测确认。试验期间获取的 HLB、healthy、suspected 数据分别为 707、289 和 108 个。

1.3 数据预处理

通过 Viewspec Pro 软件观察到地物光谱仪所采集的数据在两端存在明显机器噪声，信噪比较低，如图 3 所示。因此从 HLB 数据集中剔除 325~349 nm 和 1 000~1 075 nm 的数据，采用 350~1 000 nm 的光谱数据进行特征波段分析。

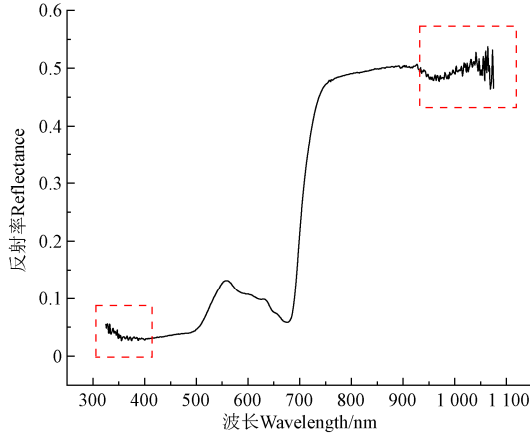


图 3 地物光谱仪机器噪声

Fig.3 Machine noise of ground object spectrometer

2 基于特征波段的柑橘黄龙病的检测方法

2.1 特征波段选择

在没有先验分析的情况下，在高光谱数据的众多波段中，通常无法确定应选择的波段数量作为感兴趣目标的特征波段数量。主流的波段选择算法往往也需要手动输入所需要的波段数量，基于聚类分析的波段选择方法可以解决这个问题。但是例如 K-means、DBSCAN 等聚类算法依赖于对其中某些参数的调整，需要依靠经验去不断尝试；层次聚类应对噪声表现不理想，奇异值会对效果产生很大影响，聚类结果可能形成链状，不利于特征波段的筛选。针对上述问题，本研究借助典型成分分析 (Exemplar Component Analysis, ECA) 的聚类思想^[25-26]，将其应用于柑橘黄龙病特征波段的提取，即从 651 个有效波段中，提取对柑橘黄龙病诊断最有效的特征波段，去除冗余波段信息。HLB 波段优选步骤如下：

首先计算各个波段的典型指数 (Exemplar Score, ES)，如式 (1) 所示，式中 ρ_i 为各波段的局部密度， δ_i 为第 i 点与具有比它大的密度的点的最近距离。

$$ES_i = \rho_i \cdot \delta_i \quad (1)$$

存在以下 4 种情况：

- 1) 当 ρ_i 较大， δ_i 较小时，说明该波段靠近聚类中心，但是有跟它属于同一类却比他更靠近聚类中心的点，因此该波段为聚类中心的可能较小
 - 2) 当 ρ_i 较小， δ_i 较小时，说明该波段处在所属类别的边缘位置。
 - 3) 当 ρ_i 较小， δ_i 较大时，说明该波段远离整个数据云团，因此极有可能是异常点。
 - 4) 当 ρ_i 较大， δ_i 较大时，说明该波段为聚类中心。
- 各波段的 ES 值如图 4a；然后将所有波段的 ES 按

照降序排列，如图 4b。从图 4b 中可以看到，ES 降序曲线存在一个明显的“拐点”，在拐点之前 ES 下降的很快，而在拐点之后，ES 下降的速度明显减慢。因此，ES 可以视为衡量波段重要性的指标，越重要的波段对应越大的 ES。

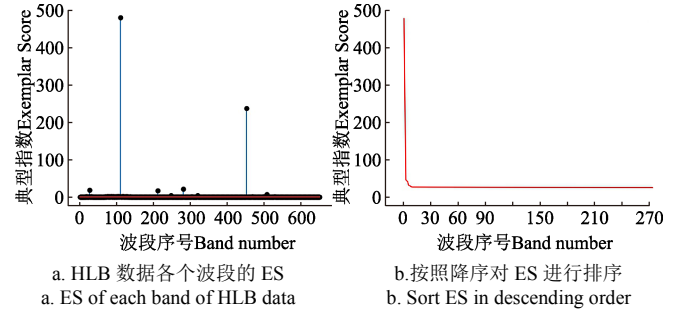


图 4 HLB 数据的 ES 分布图

Fig.4 ES distribution map of HLB data

将图 4b 中的拐点对应到图 4a 中的横线，并将图 4a 中 ES 密集区域局部放大，如图 5，可以看出对 HLB 数据而言，有 7 个波段的 ES 大于拐点（其中有 5 个 ES 的数值过大，在局部放大图中未能显示顶端圆点），这预示着 HLB 特征波段的数目可能为 7 个。通常，波长相近的波段往往具有较高的相关性，从图中可以看出，所选出的特征波段大致均匀分布在整個波段集合中，因此所选择的 HLB 特征波段具有较小的相关性，ECA 对波段相关性有较好的抑制作用。

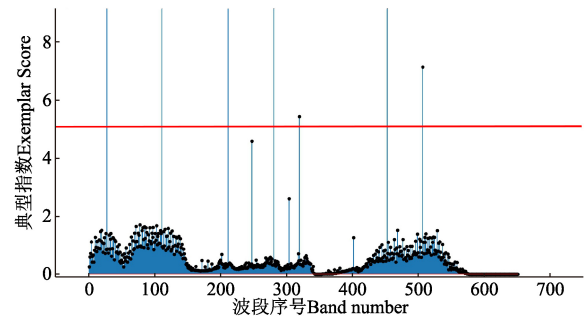


图 5 确定 HLB 特征波段数目

Fig.5 Determine the number of HLB characteristic bands

使用 ECA，已经确定 HLB 的特征波段数为 7，为了进一步确定 7 个波段可较好表征 HLB，本试验比较了其他 3 种波段选择方法：遗传算法 (Genetic Algorithm, GA)、最大椭球体积法 (Maximum Ellipsoid Volume, MEV) 和基于正交投影 (Orthogonal Projections, OP) 的波段选择方法 (OP-based Band selection, OPBS)，同样从 651 个波段中选择 7 个波段，然后分别将选出的 7 个波段使用多种方法进行分类。

2.2 模型建立方法

为了评价 4 种波段优选算法的有效性，本研究测试了噪声对波段算法的影响，数据集上分成了去除噪声的数据集和包含噪声的数据集。在模型训练中将数据集按 8:2 比例随机划分为训练集和测试集，并保证训练集和测试集中 3 类数据比例相同。为了全面评价优选特征波

段的鲁棒性,本研究采用了7种经典机器学习建模方法,对优选波段的有效性进行评估。尽管目前深度神经网络的建模效果普遍优于传统的机器学习算法,由于本研究样本集有限,采用深度神经网络容易产生过拟合,因此本文采用的算法主要是经典学习算法:逻辑回归(Logistic regression)、支持向量机(Support Vector Machine, SVM)和多个集成学习算法如:随机森林(Random forest)、引导聚类算法(Bagging)、迭代算法(Adaboost)、分布式梯度增强库(Xgboost)。然后针对训练集采用5折交叉验证的方法避免过拟合,使用Grid Search搜索方法进行超参数优化,获取最优参数后,用测试集对最终模型的泛化能力进行测试。如图6所示:

2.3 基于优选特征波段的柑橘黄龙病多光谱检测仪的研制

基于前面优选的7个特征波段,本研究比较了AMS光谱传感器的3个型号(AS7341、AS7421和AS7263)作为光探测器的选型,采用嵌入式技术研制低成本的多光谱仪,用于在田间检测柑橘黄龙病。AMS光谱传感器可采集400~1050nm范围的波段数据,基本覆盖了黄龙病的特征波段。通过软件可选择性采集指定特征波段的光谱数据^[27-30]。3个传感器的通道数量和中心波长如表1所示,其中AS7421有64个通道(中心波长从750nm到1050nm,每隔5nm取值),表中仅列出了所使用到8

个通道。测量原理表明,芯片测量某波长的线宽越窄,测量精度越高。所以,设计时应满足偏差尽量小的原则,偏差为测量波长与待测波长差值的绝对值,即所选特征波段应尽量靠近传感器通道中心波长。

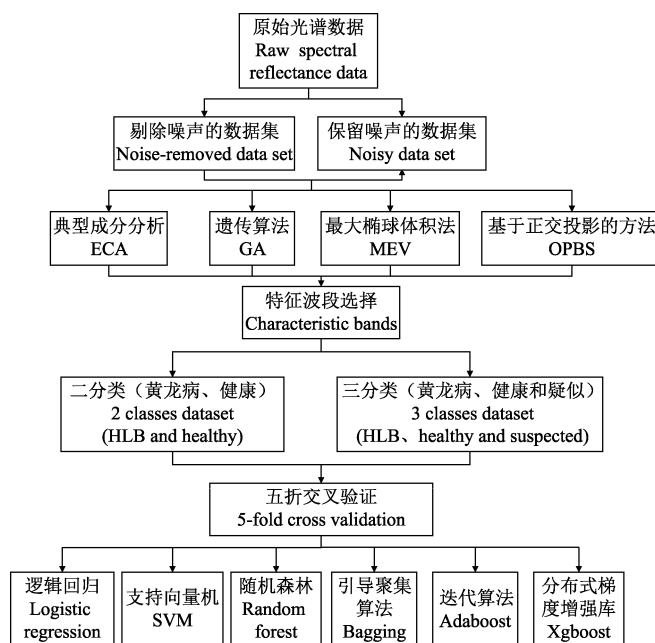


图6 数据处理步骤

Fig.6 Data processing procedure

表1 传感器通道中心波长与特征波段比较

Table 1 Comparison between the center wavelength of the sensor channel and the characteristic band

nm

传感器型号 Sensor model	AS7341			AS7263			AS7421		
	通道中心 Channel center	待测值 Value to be measured	偏差 Deviation	通道中心 Channel center	待测值 Value to be measured	偏差 Deviation	通道中心 Channel center	待测值 Value to be measured	偏差 Deviation
F1	400	377	23	610	-	-	800	802	2
	410	-	-						
	420	-	-						
F2	430	427	3	680	-	-	855	856	1
	440	-	-						
	450	436	4						
F3	460	460	0	730	707	23	935	934	1
	470	469	1					936	1
	480	-	-					-	-
F4	500	-	-	760	-	-	940	940	0
	510	-	-						
	520	521	1						
F5	540	-	-	810	-	-	945	944	1
	550	-	-						
	560	561	1						
F6	573	-	-	860	-	-	950	948	2
	583	-	-						
	593	-	-						
F7	610	-	-	-	-	-	955	953	2
	620	-	-						
	630	630	0						
F8	660	647	13	-	-	-	965	966	1
	670	669	1						
	680	-	-						

注:上述传感器都是一个 $m \times n$ 的光电二极管阵列,每个光电二极管都有自己的滤波器专用响应,即通道F,每个光电二极管之间的间距为 $200 \mu\text{m}$ 。因为不是所有 $m \times n$ 个通道都对应着待测波段,表中只列出了位于待测波段附近的通道。

Note: The above sensors features an $m \times n$ -photodiode arrays. Each photodiode has its own filter with a dedicated response, that is channel F. The pitch between each photodiode is $200 \mu\text{m}$. Because not all $m \times n$ channels correspond to the band to be measured, only the channels near the band to be measured are listed in the table.

通过图 7 可知, 采集的光谱曲线大致可以划分为 3 个区域: 350~680 nm 属于叶片的低反射率区 (A 区); 680~740 nm 属于快速变化区域 (B 区); 740~1 000 nm 属于叶片的高反射率区域 (C 区)。A 区和 C 区相对平缓, B 区快速变化, 波长偏移对试验结果影响较大, 所以应在该区域尽量减小单个波长测量的线宽。所选取的 3 个型号传感器参数如, 但相比于地物光谱仪, 所选取传感器的光谱分辨率较低, 本研究优选得到的特征波段只有 707 nm 位于 B 区。

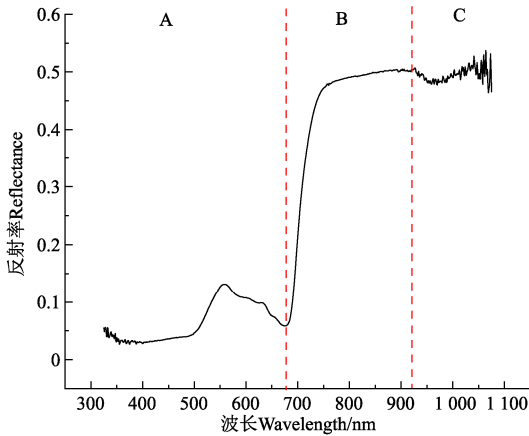


图 7 HLB 光谱的 3 个区域
Fig.7 Three zones of the entire spectrum

自研多光谱仪整个系统采用 STM32F407 作为主控制器, 使用 2 节 18650 锂电池串联供电, 通过 RT7C4337 时钟定时去控制 DC-DC 的使能脚 EN 来控制整个电路的电源。RT7C4337 的 I²C 不从单节电池端取电上拉而使用后端的 VDD_3V3 进行上拉, 以防止高电压倒灌给 STM32。AS7341、AS7421、AS7263 多光谱传感器模块测量作物的反射光谱, 通过 I2C 总线将数据传输给 STM32F407 处理器。数据经处理器处理后, 由 NB-IoT/4G 模块通过无线网络上传至服务器。STM32F407 处理器提供一个 Mini-USB 接口, 用于导出多光谱仪内存中的数据。通过预留的接口, 在多光谱仪外接手柄上扩展了操作按钮和一个 LCD 显示屏。通过外接手柄, 可以实现工作人员手持多光谱仪进行数据采集, 在未接手柄时, 多光谱仪可以通过三脚架部署在固定位置。自研多光谱仪的研发和调试由本研究团队与广州思林杰科技股份有限公司联合完成, 自研多光谱仪包含固定式和手持式两种, 其外观图和使用方式如图 8 所示。



图 8 固定式和手持式的自研多光谱仪
Fig.8 Fixed and handheld multi-spectrometer

3 结果与评价

3.1 特征波段优选结果和验证

传统的非监督高光谱波段选择技术中, 信息量大且相似性低的波段集合被认为是理想的波段选择对象。然而, 这些波段由于受到噪声或者异常的影响, 往往具有较低的鲁棒性和可靠性。为了进一步评估波段选择方法应用于柑橘黄龙病光谱特征波段选择的性能, 包含噪声的原始数据作为比较, 直接进入波段优选环节。该数据中 325~350 nm 和 1 000~1 075 nm 为机器噪声波段, 以下称为原始数据集。4 种方法选择的特征波段如表 2 所示。

表 2 所选波段子集的比较
Table 2 Comparison of selected band subsets

方法 Method	数据集 Data set	选择的波段 Selected bands/nm	特征波段/总数 Characteristic bands/total
典型成分分析 Exemplar Component Analysis (ECA)	剔除噪声 No noise	460, 802, 630, 377, 561, 856, 669	(7/651)
	包含噪声 Contains noise	460, 802, 630, 377, 562, 1 008, 856	(7/751)
遗传算法 Genetic Algorithm (GA)	剔除噪声 No noise	427, 436, 469, 521, 647, 944, 966	(7/651)
	包含噪声 Contains noise	358, 475, 571, 577, 629, 819, 1 019	(7/751)
最大椭球体积法 Maximum Ellipsoid Volume (MEV)	剔除噪声 No noise	707, 934, 936, 940, 945, 948, 953	(7/651)
	包含噪声 Contains noise	707, 934, 936, 940, 945, 1 016, 1 075	(7/751)
基于正交投影的方法 OP-based Band selection (OPBS)	剔除噪声 No noise	707, 934, 936, 940, 945, 948, 953	(7/651)
	包含噪声 Contains noise	707, 934, 936, 940, 945, 1 016, 1 075	(7/751)

由表 2 可知, 针对两种数据集, 使用 ECA 中的拐点选取的 7 个特征波段中, 有 5 个波段完全相同。有两个波段紧邻, 561 nm 和 562 nm。基于原始数据集选择出的特征波段里, 掺杂进来一个噪声波段 1 008 nm; 而在 GA 算法选择的特征波段差异较大, 未剔除噪声的数据集所选择的特征波段中掺杂有两个噪声波段; MEV 和 OPBS 在两种情况下所选择的特征波段表现出了一致性, 未剔除噪声时也有两个噪声波段被选择到。

除 ECA 外, 其他 3 种方法均在原始数据中选择到了 2 个噪声波段, 尤其是 GA 受噪声的干扰, 两次选择的特征波段有较大差异。事实上经过试验, 当所选取的特征波段超过 7 个时, 除 ECA 之外, 其他 3 种方法所选取的特征波段随着数量增加, 其中掺杂的噪声数目也会增加。

3.2 基于地物谱仪数据的模型效果比较

本试验针对地物谱仪 HH2 采集到的高光谱数据, 采用 7 种机器学习建模方法, 基于 4 种波段优选算法优选的 3 组特征波段组合, 表 3 展示了仅包含两类数据集 (HLB、healthy) 的分类效果, 表 4 展示了包含 3 类数据

集 (HLB、healthy 和 suspected) 3 类的分类效果。三分类中 suspected 类别包含了除健康和黄龙病确诊类别之外的所有缺素大类。表中给出了各模型的超参数设置以及分类准确率。

表 3 二分类时分类器的超参数和分类精度

Table 3 Hyperparameters and classification accuracy of classifier with two categories

方法 Method	分类器 Classifier	超参数 Hyperparameter	精确度 Accuracy/%
GA	Logistic regression	penalty='L2' class_weight='balanced'	84
	SVM	gamma=30 C=90	94
	Random forest	n_estimators=20 max_leaf_nodes=17	90
	Bagging	n_estimators=30 max_samples=10	82
	Adaboost	n_estimators=20 learning_rate=0.3	94
	Xgboost	max_depth=5 num_boost_round=50	96
MEV/ OPBS	Logistic regression	penalty='L2' class_weight='balanced'	86
	SVM	gamma=55 C=60	91
	Random forest	n_estimators=35 max_leaf_nodes=15	91
	Bagging	n_estimators=15 max_samples=8	83
	Adaboost	n_estimators=20 learning_rate=0.5	85
	Xgboost	max_depth=2 num_boost_round=55	90
ECA	Logistic regression	penalty='L2' class_weight='balanced'	100
	SVM	gamma=0.001 C=30	100
	Random forest	n_estimators=10 max_leaf_nodes=2	100
	Bagging	n_estimators=10 max_samples=8	100
	Adaboost	n_estimators=10 learning_rate=0.1	100
	Xgboost	max_depth=1 num_boost_round=10	100

为了直观对比性能, 将不同波段选择方法下各模型的精度绘制成条形图, 如图 9 所示。可以看出, 使用 ECA 选择的 7 个特征波段所建立的模型在各种情况下均表现优异, 在两类区分的情况下, 各模型更是达到了百分之百的准确率, 在 3 类区分的情况下, 各个模型在测试集的表现也优于其他 3 种波段选择方法。在同种波段选择方法下, 使用 Xgboost 作为分类器的模型相较于其他 5 种有更为稳定的表现。ECA 所选取的波段结合 Xgboost 分类, 平均准确率达到 92% 以上。

为了进一步定量确定优选的波段数量对柑橘黄龙病检测的影响, 本试验将选择的特征波段数从 7 个变化到 35 个, 间隔为 5, 如图 10 所示。针对包含 HLB 和 healthy 两类的数据, 将各方法选择的波段子集用 SVM 分类。各方法获得的整体分类精度如图 11 所示。从图可以看出, 无论波段选择数目为多少, ECA 选择的波段子集总能获

得较高的分类精度, 可以确定使用 ECA 筛选出的 7 个特征波段在具有良好检测效果的同时, 具有最大简约性, 这种效果符合设计光谱仪的要求。

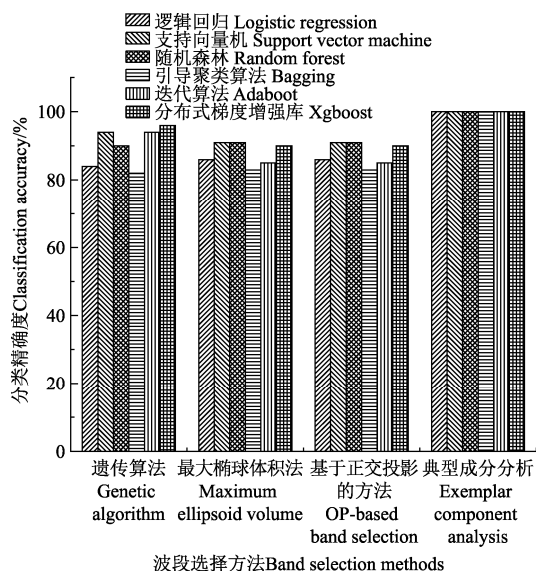
表 4 三分类时分类器的超参数和分类精度

Table 4 Hyperparameters and classification accuracy of classifier with three categories

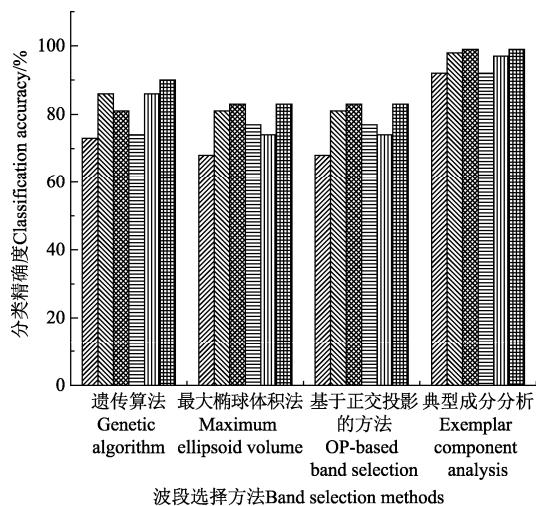
方法 Method	分类器 Classifier	超参数 Hyperparameter	精确度 Accuracy/%
GA	Logistic regression	penalty='L2' class_weight='balanced'	73
	SVM	gamma=20 C=80	86
	Random forest	n_estimators=55 max_leaf_nodes=20	81
	Bagging	n_estimators=10 max_samples=9	74
	Adaboost	n_estimators=40 learning_rate=0.5	86
	Xgboost	max_depth=6 num_boost_round=30	90
MEV OPBS	Logistic regression	penalty='L2' class_weight='balanced'	68
	SVM	gamma=40 C=10	81
	Random forest	n_estimators=95 max_leaf_nodes=20	83
	Bagging	n_estimators=30 max_samples=10	77
	Adaboost	n_estimators=70 learning_rate=0.2	74
	Xgboost	max_depth=17 num_boost_round=40	83
ECA	Logistic regression	penalty='L2' class_weight='balanced'	92
	SVM	gamma=45 C=55	98
	Random forest	n_estimators=15 max_leaf_nodes=20	99
	Bagging	n_estimators=65 max_samples=10	92
	Adaboost	n_estimators=10 learning_rate=0.2	97
	Xgboost	max_depth=3 num_boost_round=25	99

3.3 自研多光谱仪应用测试效果

采用自研多光谱仪在柑橘园中采集, 如图 11b, 每次对于同一目标分别获取 3 组 (MEV 和 OPBS 为同一组) 特征波段的光谱数据。采集规范、叶片样本获取及处理方法同使用地物光谱仪时一致。通过系统设置, 在存储的.txt 格式光谱反射率文件中可以只显示特征波段, 方便后续试验。经 PCR 验证叶片样本统计后, 本次获取的 HLB、healthy 和 suspected 样本个数分别为: 215、121、98。将 3 组特征波段下采集的数据使用之前训练好的模型进行测试, 将准确率结果绘制成条形图可以更直观了解, 如图 11 所示。相比图 9, 自研多光谱仪获取的特征波段数据输入模型后, 无论是二分类还是三分类, 相对于地物光谱仪数据选择的特征波段数据, 各模型的准确率略有下降。但基于 ECA 的波段优选方法和基于 Xgboost 的分类效果最佳, 二分类和三分类的准确率可以达到 95%。



a. 两类分类效果
a. Classification effect of the two types



b. 三类分类效果
b. Classification effect of the three types

图 9 基于地物谱仪数据的分类精度对比条形图
Fig.9 Bar graph of classification accuracy comparison using HH2 data

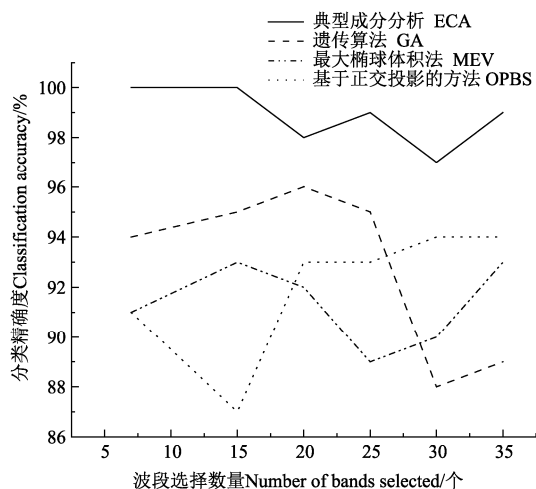
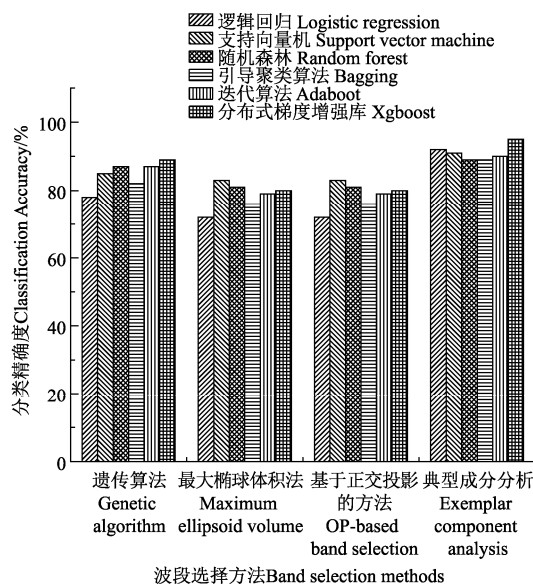
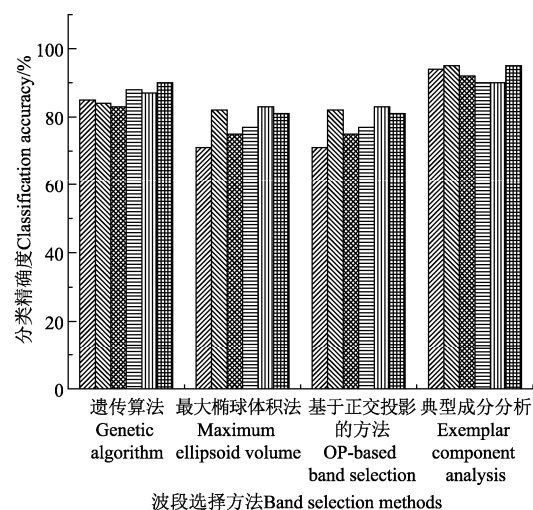


图 10 不同波段选择方法效果随子集波段数量的变化
Fig.10 The effect of different band selection methods varies with the number of subset bands



a. 两类分类效果
a. Classification effect of the two types



b. 三类分类效果
b. Classification effect of the three types

图 11 基于自研多光谱仪特征波段数据的分类精度对比条形图
Fig.11 Bar graph of classification accuracy comparison using data obtained by self-developed multi-spectrometer

4 讨论

在 4 种波段优选算法中, 当优选的特征波段数较少时, GA 具有较大的优势, 随着数量的增加, GA 选择的波段子集性能开始下降, 一个可能的原因是, 由于 GA 本身不是精确算法, 而是一种元启发式算法^[31-34], 只能搜索到近似解, 并有可能陷入局部最优, 无法保证是全局最优, 解的优劣和种群数量、迭代规则的设置有较大相关性, 而且每次计算, 会重新生成初始解, 最终结果可能会有较大偏差, 并且算法对初始种群的选择有一定的依赖性, 可结合其他算法进行改进。MEV 和 OPBS 的表现较为稳定, 且当波段数设为 7 时, 两种算法优选到的特征波段完全一致。随着选择波段数的增加, OPBS 选择的波段子集表现略有提升。然而, MEV 和 OPBS 两种算法都需要寻求使得某准则函数取得极值的波段子集^[35], 更倾向于选择集合中处于边缘位置的波段, 但这些波段

往往很不稳定, 极有可能受到环境干扰, 优选到的 7 个特征波段在整个光谱集合上呈现分布不均匀, 有些波段相邻紧密, 这对于设计一台低成本的多光谱仪不利, 因为低成本的光谱传感器其采集通道较宽, 不能满足紧邻波段的细致区分。因此, 利用自研多光谱仪采集的特征波段数据建模后, MEV 和 OPBS 的各个模型表现较差, 这也证实了这一分析。

本研究仍存在一些不足, 例如受限于发病季节性原因, 参与试验的样本偏少, 普适性有待增强; 相比于地物光谱仪, 自研多光谱仪所采集的特征波段光谱数据在各个模型上表现有所下降, 主要原因有 3 个方面: 首先, 所选取的 AMS 芯片光谱分辨率较低; 其次, 3 种型号的光谱传感器的光谱通道中心波段算法优选特征波段存在一定的偏差, 所以未能达到地物光谱仪数据的精度; 再者, 自研多光谱仪在光学结构的设计上仍有待改进。虽然多光谱仪的设计经历了多次迭代, 如在第 1 代设计基础上缩小了进光孔以减少杂光的干扰, 但与专业光学设备的采光孔径相比仍然较大, 没有镜头缺少了对入射光进行准直处理等环节, 光线在到达传感器前仍会有多次反射, 这都会降低光谱数据的质量。

然而, ECA 选择的特征波段在自研多光谱仪数据分析中仍获得了最佳的表现, 由前人的研究可知^[36-39], 染病显状和缺素疑似柑橘叶片在 450~700 nm 的光谱反射率明显高于健康叶片, ECA 所选择的 7 个波段中有 4 个位于此区间中; 感染黄龙病的柑橘叶片 (包含显状和未显状) 的光谱反射率在 730~1 023 nm 波长范围内呈现升高, 而单纯缺素疑似的叶片则呈现降低, 两者表现出相反的变化趋势, 所以此区间波段也有助于区分健康与染病植株, ECA 所选择的 7 个波段中有 2 个位于此区间。选择的 7 个黄龙病特征波段在整个光谱集合上分布均匀, 相关性不高, 并且每个波段与所选光谱传感器通道的中心波段偏差较小, 使得获取的数据质量较高。

本文研制的低成本光谱检测仪具有可拓展性, 不仅可以应用于本文的柑橘黄龙病检测, 也可以在其他作物病虫害和营养胁迫的特征波段基础上, 进行特定应用的波段定制, 从而实现多种作物多种病害、营养成分的分析和诊断。

5 结 论

本研究探索了柑橘黄龙病光谱特征波段提取以及光谱检测仪研制方法。针对柑橘黄龙病的检测, 提出了基于典型成分分析 (Exemplar Component Analysis, ECA) 的波段优选方法。ECA 算法优选了 7 个柑橘黄龙病特征波段, 分别为 377、460、561、630、669、802 nm 和 856 nm。基于以上优选的特征波段, 采用 AMS 光谱传感器和嵌入式技术设计了一款低成本的多光谱仪。得出主要结论如下:

1) 本文提出的基于 ECA 的柑橘黄龙病波段优选方法, 可以较好地获得柑橘黄龙病的特征波段, 鲁棒性高, 不仅可以区分健康叶片和黄龙病 (HLB) 叶片, 也可区分 HLB 叶片与因缺素表现为与黄龙病高度相似的叶片, 在包含机器噪声的情况下仍然可以较准确地获得较一致

的特征波段。

2) 相较于类似研究, 本文基于 ECA 的波段优选方法仅优选 7 个波段, 最大程度去掉了冗余的波段, 面对大量样本的应用场景, 可以减少数据传输与存储的压力。

3) 不同分类器的比较表明, 分布式梯度增强库在光谱数据方面表现最好, 并且效果较为稳定。基于特征波段, 其他分类器也实现了较高的检测准确率, 但在不同波段选择方法的表现并不稳定。

4) 设计了一款低成本多光谱仪。光谱仪数据结合基于特征波段搭建的模型能够获得不错的 HLB 检测效果, 虽然仍有一定的错误率, 但在三分类情况下最高识别准确率可达 95%, 相比于高精度、高价格的专业仪器, 其应用前景未来可期。

【参 考 文 献】

- [1] 袁亦文, 戈丽清, 王德善, 等. 柑橘黄龙病对柑橘产量和品质的影响[J]. 浙江农业科学, 2007(1): 87-88.
Yuan Yiwen, Ge Liqing, Wang Deshan, et al. Effect of citrus huanglong disease on citrus yield and quality[J]. Journal of Zhejiang Agricultural Sciences, 2007(1): 87-88. (in Chinese with English abstract)
- [2] 程春振, 曾继吾, 钟云, 等. 柑橘黄龙病研究进展[J]. 园艺学报, 2013, 40(9): 1656-1668.
Cheng Chunzhen, Zeng Jiwu, Zhong Yun, et al. Research progress on citrus huanglongbing disease[J]. Acta Horticulturae Sinica, 2013, 40(9): 1656-1668. (in Chinese with English abstract)
- [3] 柏自琴, 周常勇. 柑橘黄龙病病原分化及发生规律研究进展[J]. 中国农学通报, 2012, 28(1): 133-137.
Bai Ziqin, Zhou Changyong. The research progress of citrus huanglongbing on pathogen diversity and epidemiology [J]. Chinese Agricultural Science Bulletin, 2012, 28(1): 133-137. (in Chinese with English abstract)
- [4] Rosales R, Burns J K. Phytohormone changes and carbohydrate status in sweet orange fruit from Huanglongbing-infected trees[J]. Journal of Plant Growth Regulation, 2011, 30(3): 312-321.
- [5] Bové J M. Huanglongbing: A destructive, newly-emerging, century-old disease of citrus[J]. Journal of Plant Pathology, 2006, 88(1): 7-37.
- [6] 林孔湘. 柑桔黄梢(黄龙)病研究 II. 关于病原的探讨[J]. 植物病理学报, 1956(1): 13-42, 102-104.
Lin Kongxiang. Studies on citrus yellow shoot (Huang long) disease II discussion on pathogen[J]. Acta Phytopathologica Sinica, 1956(1): 13-42, 102-104. (in Chinese with English abstract)
- [7] 鹿连明, 范国成, 胡秀荣, 等. 田间柑橘植株不同部位黄龙病菌的 PCR 检测及发病原因分析[J]. 植物保护, 2011, 37(2): 45-49.
Lu Lianming, Fan Guocheng, Hu Xiurong, et al. PCR detection of huanglongbing pathogen in different parts of citrus plants in the field and analysis of the cause of the disease [J]. Plant Protection, 2011, 37(2): 45-49. (in Chinese with English abstract)
- [8] Bassanezi R B, Montesino L H, Stuchi E S. Effects of huanglongbing on fruit quality of sweet orange cultivars in Brazil[J]. European Journal of Plant Pathology, 2009, 125(4): 565-572.
- [9] 张超. 柑橘黄龙病病原亚洲种 Sec 依赖分泌蛋白的预测、鉴定及功能研究[D]. 北京: 中国农业科学院, 2020.
Zhang Chao. Prediction, Validation and Functional Study of Candidatus Liberibacter Asiaticus Sec Dependent Secretary

- Proteins[D]. Beijing: Chinese Academy of Agricultural Sciences, 2020. (in Chinese with English abstract)
- [10] 钱文娟. 基于核酸扩增技术的柑橘黄龙病检测方法研究[D]. 杭州: 浙江大学, 2018.
- Qian Wenjuan. Detection of Citrus Huanglongbing Based on Nucleic Acid Amplification Technique[D]. Hangzhou: Zhejiang University, 2018. (in Chinese with English abstract)
- [11] 王爱民, 邓晓玲. 柑桔黄龙病诊断技术研究进展[J]. 广东农业科学, 2008(6): 101-103.
- Wang Aimin, Deng Xiaoling. Research progress in diagnostic techniques of citrus huanglong disease[J]. Guangdong Agricultural Sciences, 2008(6): 101-103. (in Chinese with English abstract)
- [12] 袁波, 杨青. 光谱技术及应用[M]. 浙江: 浙江大学出版社, 2019: 44-79.
- [13] 张凝, 杨贵军, 赵春江, 等. 作物病虫害高光谱遥感进展与展望[J]. 遥感学报, 2021, 25(1): 403-422.
- Zhang Ning, Yang Guijun, Zhao Chunjiang, et al. Progress and prospects of hyperspectral remote sensing technology for crop diseases and pests[J]. National Remote Sensing Bulletin, 2021, 25(1): 403-422. (in Chinese with English abstract)
- [14] 童庆禧, 张兵, 张立福. 中国高光谱遥感的前沿进展[J]. 遥感学报, 2016, 20(5): 689-707.
- Tong Qingxi, Zhang Bing, Zhng Lifu. Current progress of hyperspectral remote sensing in China [J]. National Remote Sensing Bulletin, 2016, 20(5): 689-707. (in Chinese with English abstract)
- [15] 姚云军, 秦其明, 张自力, 等. 高光谱技术在农业遥感中的应用研究进展[J]. 农业工程学报, 2008, 24(7): 301-306.
- Yao Yunjun, Qin Qiming, Zhang Zili, et al. Research progress of hyperspectral technology applied in agricultural remote sensing [J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2008, 24(7): 301-306. (in Chinese with English abstract)
- [16] 桂江生, 费婧怡, 傅霞萍. 三维小样本元学习模型的大豆食心虫害高光谱检测[J]. 光谱学与光谱分析, 2021, 41(7): 2171-2174.
- Gui Jiangsheng, Fei Jingyi, Fu Xiaping. Hyperspectral imaging for detection of leguminivora glycinivorella based on 3D few-shot meta-learning model[J]. Spectroscopy and Spectral Analysis, 2021, 41(7): 2171-2174. (in Chinese with English abstract)
- [17] Pascucci S, Pignatti S, Casa R, et al. Special issue "hyperspectral remote sensing of agriculture and vegetation"[J]. Remote Sensing, 2020, 12(21): 3665-3665.
- [18] Moriya É A S, Imai N N, Tommaselli A M G, et al. Detecting citrus huanglongbing in Brazilian orchards using hyperspectral aerial images[J]. The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 2019, 42: 1881-1886.
- [19] Weng H, Lv J, Cen H, et al. Hyperspectral reflectance imaging combined with carbohydrate metabolism analysis for diagnosis of citrus Huanglongbing in different seasons and cultivars[J]. Sensors and Actuators B: Chemical, 2018, 275: 50-60.
- [20] 孙康, 陈金勇. 高光谱图像波段选择技术[M]. 北京: 电子工业出版社, 2020: 40.
- [21] 肖怀春. 基于高光谱成像技术的柑橘黄龙病诊断方法研究[D]. 上海: 华东交通大学, 2019.
- Xiao Huaichun. Study on Diagnosis Methods of Citrus Greening Based on Hyperspectral Imaging[D]. Shanghai: East China Jiaotong University, 2019. (in Chinese with English abstract)
- [22] 王克健. 柑橘黄龙病光谱检测与柑橘木虱 UAV 防控技术探索[D]. 重庆: 西南大学, 2018.
- Wang Kejian. Exploration of the Spectrum Monitoring Technology of Huanglongbing and the UAV Control Technology of Citrus Psyllid[D]. Chongqing: Southwest University, 2018. (in Chinese with English abstract)
- [23] Deng X, Huang Z, Zheng Z, et al. Field detection and classification of citrus Huanglongbing based on hyperspectral reflectance[J]. Computers and Electronics in Agriculture, 2019, 167: 105006.
- [24] Deng X, Zhu Z, Yang J, et al. Detection of citrus huanglongbing based on multi-input neural network model of UAV hyperspectral remote sensing[J]. Remote Sensing, 2020, 12(17): 2678.
- [25] Rodriguez A, Laio A. Clustering by fast search and find of density peaks[J]. Science, 2014, 344(6191): 1492-1496.
- [26] Sun K, Geng X, Ji L. Exemplar component analysis: A fast band selection method for hyperspectral imagery[J]. IEEE Geoscience and Remote Sensing Letters, 2014, 12(5): 998-1002.
- [27] 朱豪男, 胡孟晗, 张健, 等. 低成本便携式多光谱成像系统的研发及优化[J]. 中国图象图形学报, 2021, 26(8): 1796-1808.
- Zhu Haonan, Hu Menghan, Zhang Jian, et al. Development and optimization of a low-cost and portable multispectral imaging system [J]. Journal of Image and Graphics, 2021, 26(8): 1796-1808. (in Chinese with English abstract)
- [28] 万会江. 多光谱成像系统研究[D]. 杭州: 浙江大学, 2011.
- Wan Huijiang. Research of Multi-spectral Imaging System[D]. Hangzhou: Zhejiang University, 2011. (in Chinese with English abstract)
- [29] 乔鼎恒. 多光谱成像系统硬件设计与实现[D]. 北京: 北京工业大学, 2018.
- Qiao Dingheng. Multispectral Imaging System Hardware Design and Implementation[D]. Beijing: Beijing University of Technology, 2018. (in Chinese with English abstract)
- [30] 殷文鑫. 基于多旋翼无人机的多光谱成像遥感系统开发及应用[D]. 杭州: 浙江大学, 2018.
- Yin Wenxin. Development and Application of a Multispectral Remote Sensing System Based on the Multi-rotor Drone[D]. Hangzhou: Zhejiang University, 2018. (in Chinese with English abstract)
- [31] 席裕庚, 柴天佑, 恽为民. 遗传算法综述[J]. 控制理论与应用, 1996(6): 697-708.
- Xi Yugeng, Chai Tianyou, Yu Weimin. Survey on genetic algorithm[J]. Control Theory & Applications, 1996(6): 697-708. (in Chinese with English abstract)
- [32] 李浩光, 于云华, 逢燕, 等. 基于遗传算法的近红外光谱定性分析特征波长提取方法研究[J]. 光谱学与光谱分析, 2021, 41(8): 2437-2442.
- Li Haoguang, Yu Yunhua, Pang Yan, et al. Study on characteristic wavelength extractin method for near infrared spectroscopy identification based on genetic algorithm[J]. Spectroscopy and Spectral Analysis, 2021, 41(8): 2437-2442. (in Chinese with English abstract)
- [33] 赵冬, 赵光恒. 基于改进遗传算法的高光谱图像波段选择[J]. 中国科学院研究生院学报, 2009, 26(6): 795-802.
- Zhao Dong, Zhao Guangheng. Band selection of hyperspectral image based on improved genetic algorithm[J]. Journal of University of Chinese Academy of Sciences, 2009, 26(6): 795-802. (in Chinese with English abstract)
- [34] Lucas C, Leon N, Yiannis A, et al. A new visible band index (vNDVI) for estimating NDVI values on RGB images utilizing genetic algorithms[J]. Computers and Electronics in Agriculture, 2020, 172: 105334-105334.
- [35] Sheffield C. Selecting band combinations from multispectral data[J]. Photogrammetric Engineering and Remote Sensing, 1985, 51(6): 681-687.
- [36] 吴越, 苏华楠, 黄爱军, 等. 柑橘黄龙病菌侵染对甜橙叶片糖代谢的影响[J]. 中国农业科学, 2015, 48(1): 63-72.

- Wu Yue, Su Huanan, Huang Aijun, et al. Effect of candidatus *liberibacter asiaticus* infection on carbohydrate metabolism in citrus *sinensis*[J]. *Scientia Agricultura Sinica*, 2015, 48(1): 63-72. (in Chinese with English abstract)
- [37] 梅慧兰, 邓小玲, 洪添胜, 等. 柑橘黄龙病高光谱早期鉴别及病情分级[J]. *农业工程学报*, 2014, 30(9): 140-147.
- Mei Huilan, Deng Xiaoling, Hong Tinseng, et al. Early detection and grading of citrus huanglongbing using hyperspectral imaging technique[J]. *Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE)*, 2014, 30(9): 140-147.
- [38] 樊晶. 柑橘宿主对黄龙病病原菌侵染的应答机制[D]. 重庆: 重庆大学, 2010.
- Fan Jing. Response of Citrus to the Infection of Huanglongbing-associated Bacteria[D]. Chongqing: Chongqing University, 2010. (in Chinese with English abstract)
- [39] 翁海勇. 基于光学成像技术的柑橘黄龙病快速检测方法研究[D]. 浙江: 浙江大学, 2019.
- Weng Haiyong. Optical Imaging for Rapid Detection of Citrus Huanglongbing[D]. Zhe Jiang: Zhejiang University, 2019. (in Chinese with English abstract)

Selection of spectral characteristic bands of HLB disease of citrus and spectrum detector development

Lan Yubin^{1,2,3,4}, Wang Tianwei¹, Guo Yaqi¹, Yang Dongzi¹, Lin Shaoming¹, Hu Yuqi¹, Lin Xiaoqing¹,
Deng Xiaoling^{1,2,3,4*}

(1. College of Electronic Engineering, College of Artificial Intelligence, South China Agricultural University, Guangzhou 510642, China; 2. National Center for International Collaboration Research on Precision Agricultural Aviation Pesticide Spraying Technology, Guangzhou 510642, China; 3. Guangdong Laboratory for Lingnan Modern Agriculture, Guangzhou 510642, China; 4. Guangdong Engineering Technology Research Center of Smart Agriculture, Guangzhou 510642, China)

Abstract: Huanglongbing (HLB) is known as one of the most important diseases in citrus production. Early detection of diseased plants can be used to prevent the disease from the spreading, in order to reduce the severity of the disease. Hyperspectral analysis can be expected in the crop disease detection in recent years, due to the rich spectral information. However, the hyperspectral equipment is too expensive to apply in large scale, particularly with a large amount of calculation for a large number of bands. An accurate band selection can be utilized to remove the redundant information for the less data storage, calculation, and transmission, in order to avoid the "dimension disaster", and equipment costs. In this study, an Exemplar Component Analysis (ECA)-based method was proposed to optimize the characteristic wavebands of HLB diseases using a ground feature spectrometer. The spectrum detector was also selected the spectral characteristic bands from the hyperspectral information of citrus plants. Three algorithms were utilized to optimize the combination of sever spectral wavebands for comparison. The optimal bands were obtained to evaluate the robustness of the four band selection. Among them, the six machine learning methods were used for the modeling and classification. In addition, a multi-spectrometer was designed using the optimal characteristic band for the detection of citrus HLB. The results show that an accuracy of more than 92% was achieved in the test set of the feature band that selected by the ECA algorithm combined with the six classifiers, indicating the excellent robustness. The self-developed multi-spectrometer was detected up to 95% accuracy for the HLB using the characteristic band. Experiments show that the multiple spectrometer was feasible to characterize the HLB for a small number of characteristic bands. A reasonable characteristic band was greatly reduced the design cost of spectrum detection in the special agricultural disease, and then improved the accuracy of disease prevention and control in orchards. As such, the HLB spectral feature bands were extracted to develop a spectral detector. The proposed ECA-based band selection can be expected to obtain the characteristic bands of citrus HLB with the high robustness. Specifically, the healthy leaves were distinguished from the HLB leaves, particularly from the high similarity to the HLB disease due to deficiency. Moreover, the ECA-based band preference was only preferred seven bands. The redundant bands were removed to reduce the data transmission and storage in the application scenario of large number of samples. The comparison of different classifiers showed that the Xgboost performed the best, in terms of the stable spectral data. The high detection accuracy was also achieved in the rest classifiers using the feature bands. A low-cost multispectral instrument was also designed to promote the unstable performance in different band selection. Anyway, the accurate HLB detection was obtained using the spectrometer data with the feature bands model, where the highest recognition accuracy was close to 95% in the case of triple classification. The promising application prospect was given in the future, compared with the rest high precision and high price professional instruments. The data analysis and spectrometer can be affordable to promote the smart agriculture, compared with the generally expensive spectrometers on sale. Low-cost spectrometers are more likely to be popularized, in order to reduce the amount of generated data for the less hardware requirements. A balance between detection accuracy and production cost can lay a strong foundation for the wide popularization of intelligent equipment.

Keywords: hyperspectral; remote sensing; intelligent detection; citrus Huanglongbing; customized spectrometer; band selection