

# 基于荧光透射谱和高光谱图像纹理的茶叶病害预测研究

芦兵<sup>1,2</sup>, 孙俊<sup>1\*</sup>, 杨宁<sup>1</sup>, 武小红<sup>1</sup>, 周鑫<sup>1</sup>

1. 江苏大学电气信息工程学院, 江苏 镇江 212013

2. 江苏大学信息化中心, 江苏 镇江 212013

**摘要** 为了实现对茶叶病害的准确预测, 避免病害特征提取过程中对茶叶的二次破坏, 利用荧光透射技术对茶叶赤叶病叶片的荧光透射光谱特性展开研究。实验采集了健康茶叶叶片样本 45 个、赤叶病初期叶片样本 60 个及赤叶病中期叶片样本 60 个, 并按照 2:1 的比例划分成训练集和预测集样本数, 通过荧光透射手段利用高光谱仪器采集这些叶片的原始荧光透射光谱。通过对这 3 组叶片样本平均光谱强度曲线的分析, 证实了利用荧光透射光谱信息对这 3 种病害类型叶片进行分类的可行性。然后使用多项式平滑(savitzky-golay, S-G)方法对原始光谱进行平滑和降噪处理。最后采用竞争性自适应重加权抽样法(competitive adaptive reweighted sampling, CARS)对预处理后的光谱数据进行特征波长的选取。经过 50 次加权采样后, 最终选取 4 个特征波长, 分别为: 463, 512, 586 和 613 nm。为了最大化提取样本的病害特征信息, 强化分类器输入病害特征值的典型性, 使用高光谱反射技术, 采集 4 个特征波长下的高光谱图像, 分别使用 2 种不同的纹理提取算法提取病害叶片图像的纹理信息进行对比分析。首先利用灰度共生矩阵(GLCM)提取 4 幅图像的纹理信息, 分别计算 4 个方向的灰度共生矩阵(0°, 45°, 90°及 135°), 然后计算 5 个共生矩阵的均值和方差。为了提高鲁棒性, 取 4 幅图像纹理信息的平均值作为该叶片的纹理特征值, 最终得到 10 个特征值。利用 LBP(local binary patterns)算法获取特征波长下高光谱图像的纹理信息, 并使用 Uniform 模式对 LBP 模型进行降维, 最终每幅图像得到 944 个维度的 LBP 特征值, 同样取 4 幅图像的平均值作为该叶片的 LBP 纹理特征值。最后通过极限学习机(ELM)分别建立特征光谱联合灰度共生矩阵纹理信息及 LBP 算子纹理信息的预测模型, 由于模型的输入特征值不在一个量纲, 首先对输入特征值进行归一化处理, 然后再定义模型的输出标签, 即健康叶片的预测模型输出为 1, 赤叶病早期为 2, 中期为 3。实验测得基于 CARS-GLCM-ELM 模型的预测准确率为 81.82%, 基于 CARS-LBP-ELM 模型的预测准确率为 85.45%, 说明利用荧光透射光谱联合 LBP 算子纹理信息预测效果更好。由于没有达到预期效果, 利用 Softplus 函数对 ELM 的隐含层激活函数进行了优化, 替换掉原来的 Sigmoid 函数, 优化后的模型预测分类正确率达到 92.73%, 基本达到了预期效果。该研究将病害叶片的荧光光谱信息和对应特征波长下高光谱图像的纹理信息进行了融合, 研究结果可为茶叶病害的快速、准确预测提供一定的参考价值。

**关键词** 荧光透射; 高光谱图像; 特征融合; 灰度共生矩; LBP 算子; ELM

**中图分类号:** TP391.41 **文献标识码:** A **DOI:** 10.3964/j.issn.1000-0593(2019)08-2515-07

## 引言

茶叶是我国重要的经济作物, 对茶叶病害的及早发现与诊断, 有利于农业生产者及时采取有效的防护措施。如果能对病害所处周期进行准确判断, 将极大提高病害治疗方案的

精准性, 从而减少因欠治疗或过度治疗造成的不必要损失。如茵亚门真菌引起的茶赤叶斑病是发病率比较高的茶叶常见病之一, 其分生孢子器可在茶病叶组织里越冬, 靠风雨和水滴溅射传播, 危害性极高。对该类病害中前期可通过喷洒适量多菌灵可湿性粉剂或甲基硫菌灵悬浮剂进行预防和治疗。传统农业生产者多通过人工经验的方式对此类病害进行判

收稿日期: 2018-07-24, 修订日期: 2018-11-20

基金项目: 国家自然科学基金项目(31471413, 31701324), 江苏高校优势学科建设工程项目 PAPP[苏政办发(2011)6号], 江苏省六大人才高峰项目(ZBZZ-019), 常州市科学技术项目(社会发展)(CE20175042)资助

作者简介: 芦兵, 1983年生, 江苏大学电气信息工程学院博士研究生 e-mail: lubing@ujs.edu.cn

\* 通讯联系人 e-mail: sun2000jun@ujs.edu.cn

断,效率不高。叶类植物受到病害胁迫时叶片中的叶绿素组织往往会遭到破坏<sup>[1]</sup>,叶绿素荧光素有叶绿素“探针”之称,很多学者通过分析植物叶绿素荧光光谱和病害之间的关系建立了植物病害的预测模型,并取得了不错的效果。隋媛媛等<sup>[2]</sup>对黄瓜霜霉病害进行研究,使用激光激发叶绿素荧光提取叶片光谱信息,然后利用光谱特征值建立了病害特征指标用于对蔬菜病害进行分类预测,其预测能力达到 91.38%。以上研究说明利用荧光光谱信息对植物病害进行预测是完全可行的。但是在研究过程中发现,由于植物叶片在病害胁迫初期会产生活性酶抗体,这种抗体的存在干扰了叶绿素荧光强度的变化,特别是在防御活性酶和叶绿素双衰减的时候,很难判荧光强度的变化是由哪个因素引起的。为了强化病害的特征,有些学者提出在光谱的基础上加入部分病害叶片的图像特征,如 Xie 等<sup>[3]</sup>研究了茄子叶片早期枯萎病的光谱和纹理特征,将高光谱图像转换成 RGB, HSV 和 HLS 三种格式,并利用灰度共生矩阵 (GLCM) 提取各类图片的纹理信息,建立 K-近邻 (KNN) 和 AdaBoost 分类模型,用于检测健康和感染样本,识别率为 88.46%。张帅堂等<sup>[4]</sup>利用颜色矩提取了 4 种类别的茶叶病害颜色信息,联合光谱特征值通过 BP 神经网络对茶叶病害进行分类预测,病害识别率提高到 94.17%。但是以上研究多侧重不同病害类型的分类,对同

种病害,不同病期的病害特征关注较少。

以不同病期的茶叶赤叶病为研究对象(如图 1 所示),通过荧光透射技术采集荧光光谱信息,通过极限学习机 (ELM) 分别建立特征光谱联合灰度共生矩阵纹理及 LBP 算子纹理的预测模型,提高了不同病期病害的识别率。

## 1 实验部分

### 1.1 样本制备

实验所用病害茶叶样本均采摘自江苏省镇江市(位于:北纬 32°12', 东经 119°27')野鸡山茶种植与加工基地。在春雨过后 2 周左右,空气较为湿润的早晨进入茶叶种植基地,寻找发生病害的茶叶植株,选择大小 3 cm×6 cm 左右的叶片进行采摘,对采摘的病变叶片按病害周期进行分类后装入保鲜袋密封,并放入置有冰块的便携式保温箱中保存,然后迅速送往现代农业与装备重点实验室进行实验。按病害程度将叶片样本分成三类,分别为健康样本、赤叶初期样本和赤叶中期样本,通过筛选最终得到健康叶片样本 45 个、赤叶病初期及赤叶病中期叶片样本各 60 个,并按照 2:1 的比例划分各类训练集和预测集样本数。

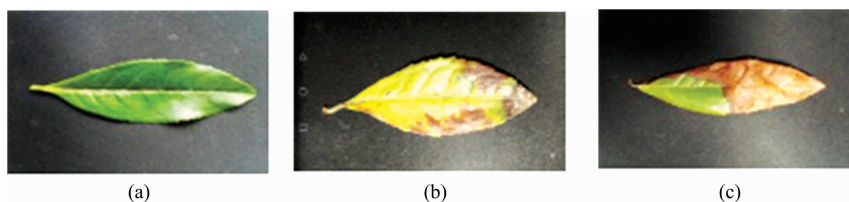


图 1 部分茶叶叶片样本

(a): 健康; (b): 赤叶初期; (c): 赤叶中期

Fig. 1 Some tea leaf samples

(a): Healthy leaf; (b): Early stage of red leaf disease; (c): Intermediate stage of red leaf disease

### 1.2 仪器设备

光谱图像的采集设备是由芬兰 Spectral Imaging 公司生产的 ImSpector V10E 型高精度光谱图像采集套件,分辨率 (30  $\mu\text{m}$  狭缝) 2.8 nm, 成像相机 (Zyla4. 4-6. 5  $\mu\text{m}$ -53FPS (1 024×1 024), Andor, Britain), LED 灯光源 (UBerLED100(340~800 nm), IT, USA), 卤素灯光纤光源 (3900E(400~2 000 nm), ISUZU OPTICS, 中国台湾), 电控位移台 (MSI300, 移动精度: 10  $\mu\text{m}$ , 行程: 30 cm, 中国台湾), 暗箱 (DC1300, ISUZU OPTICS, 中国台湾)。构建好的综合实验平台如图 2(a)所示。

在进行荧光光谱数据采集时,需将样本放置在带有透明玻璃的载物台上,透明玻璃由特殊材质制成,透光率大于 99%,且整个平面透光率绝对均匀,以降低对荧光激发光源的影响。为进一步降低激发光源和环境噪声对荧光光谱的干扰,在摄像头安装一个滤镜,滤镜的光特性如图 2(b)所示,从图中可以看出,470~630 nm 的光谱可以 100% 穿透滤镜,其余波段光谱则被滤镜抑制,无法被摄像头采集。在进行反射光谱图像采集时,需移除此玻璃载物台和滤镜,将

样本直接放置在黑色背景的电控位移台上。

### 1.3 光谱信息采集

#### 1.3.1 荧光光谱采集

将样本置于透明玻璃载物台上,设置物距 0.4 m,移动台速度 1.25 mm·s<sup>-1</sup>,使用中心波长 370 nm 的荧光激发光源对样品进行照射,由于受激发后的叶绿素荧光发射过程缓慢,且强度较低,如果曝光时间设置太短将无法捕捉到荧光光谱信息,经过反复尝试后得到最佳曝光时间为 220 ms,此时能够捕捉到较为明显的荧光光谱信号。最终完成对所有样本的荧光光谱采集后,得到三类样本的平均光谱,如图 3 所示。

当被感染叶片受到病菌孢子胁迫时,虽然叶片边缘的叶绿素组织已出现褐色病变,但叶片出于自身免疫力的保护作用,抗氧化酶活性增强,叶片整体细胞活性增强,从而导致荧光强度增强<sup>[5]</sup>,因此赤叶初期叶片的荧光波峰强度高于健康叶片。到了中期,活性酶在病菌胁迫下出现凋亡,叶片中多数叶绿素组织遭到破坏,荧光强度显著降低。通常,荧光光谱的波长要大于激发光的波长,那是因为处于激发态的叶

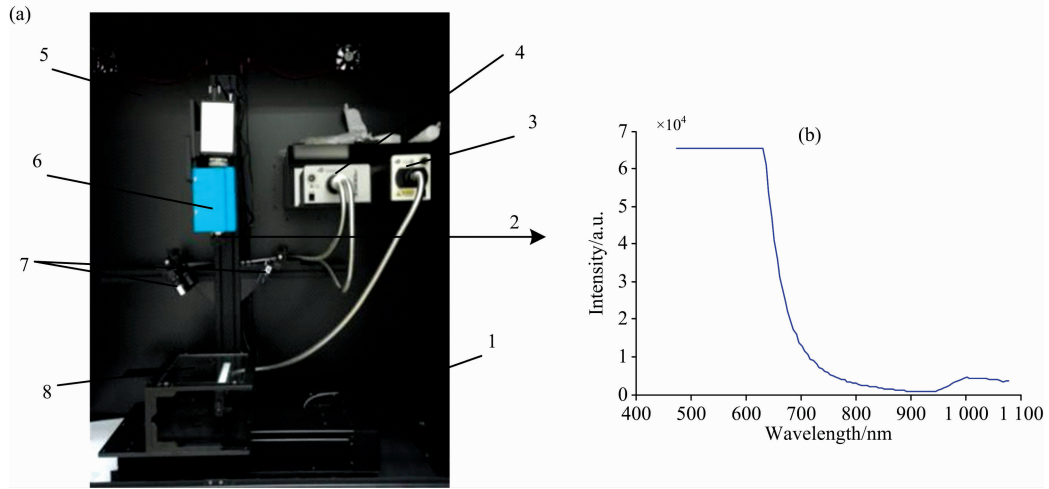


图 2 光谱数据采集综合平台

1: 载物台; 2: 滤镜; 3: 荧光电源; 4: 卤素光电源; 5: 暗箱; 6: CCD 摄像头; 7: 卤素光光源; 8: 荧光光源  
(a): 光谱采集套件; (b): 滤镜光特性

Fig. 2 Spectral data acquisition comprehensive platform

1: Carrier platform; 2: Filter; 3: Fluorescent power supply; 4: Halogen optical power supply;  
5: Dark box; 6: CCD camera; 7: Halogen light sources; 8: Fluorescent light source  
(a): Spectral acquisition Kit; (b): Optical character of filter

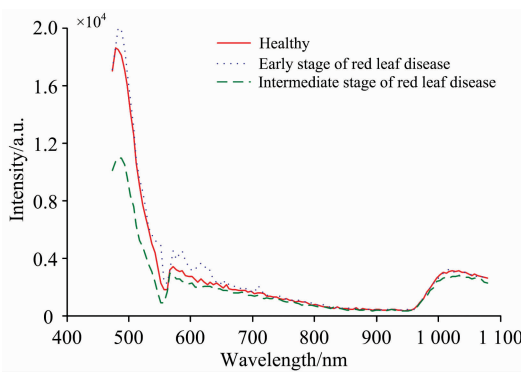


图 3 荧光平均光谱图

Fig. 3 Average spectrogram of fluorescence

绿素分子需要回归到稳定的基态, 在回归过程中有部分能量通过热能的形式散逸掉, 因此在位于 580 nm 处会产生一个小波峰。由于滤镜的通行带宽为 470~630 nm, 因此大于 630 nm 的光谱信息是无效信息。1 000 nm 处平均光谱和滤镜光特性分布曲线均出现了一个相似的波峰形态, 说明 630 nm 以后的光谱信息确实为滤镜的抑制噪声, 而非有效的荧光光谱信息。从图 3 中可以看出, 赤叶病害叶片在蓝绿光区存在一个明显增强的过程, 而在 600 nm 以后存在一个明显的下降趋势, 且不同病害程度的光谱强度在变化趋势中存在明显差异, 这一特性可作为不同病期茶叶赤叶病的诊断依据。经过数据处理最终得到 470~630 nm 范围内有效荧光光谱波长 72 个。

1. 3. 2 高光谱图像采集

在进行高光谱图像特征采集时, 需要关闭荧光激发电源, 开启载物台上方倾角 45° 的两个卤素光源, 将下方的玻

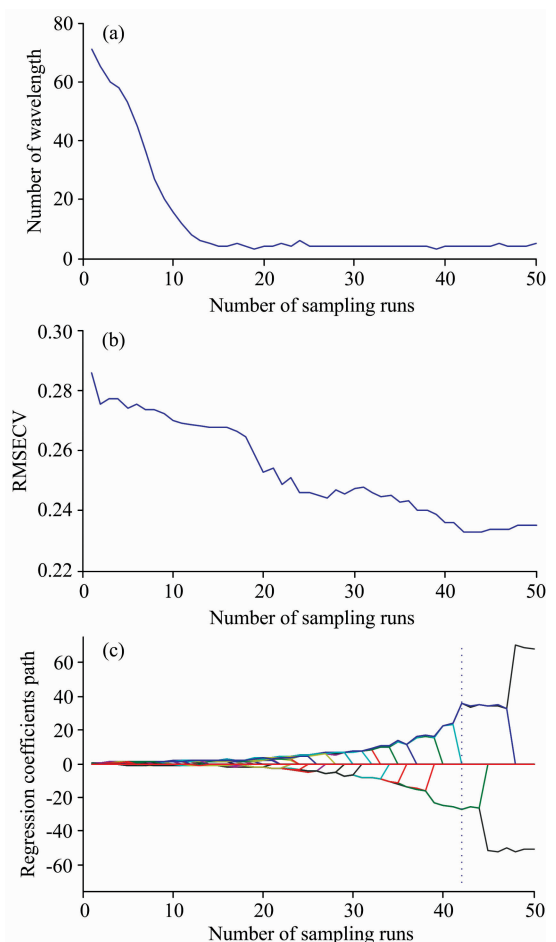


图 4 CARS 波长筛选过程

Fig. 4 The process of selecting wavelengths of CARS

璃载物台和荧光透射灯移除,重新设置物距为 0.5 m,曝光时间 25 ms,移动台速度  $1.75 \text{ mm} \cdot \text{s}^{-1}$ ,光谱范围 400~1 000 nm,然后将叶片样本直接摆放在黑色背景的移动载物台上进行高光谱图像的采集。

## 2 结果与讨论

### 2.1 光谱特征值选取

为了提高后期建模效率,需要从这些波段中选取最具代表性的特征波长<sup>[6]</sup>。本实验选用 CARS(competitive adaptive reweighted sampling)算法对特征波长进行采样选取,该算法已被证明在特征波长的选取上具有非常好的效果<sup>[7]</sup>。在 Matlab R2014a 中首先通过 S-G(Savitzky-Golay)方法对光谱数据进行平滑去噪,然后运行 CARS 算法进行特征波长的优选,选取过程如图 4 所示。

从图 4(a)中可以看出,在指数衰减函数<sup>[8]</sup>的作用下,经

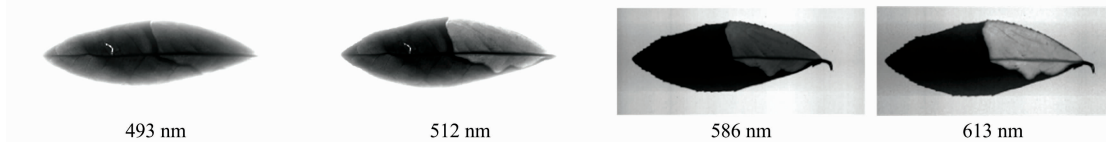


图 5 各特征波长下高光谱图像

Fig. 5 Hyperspectral images at each characteristic wavelength

#### 2.2.1 LBP 特征

LBP 算子被广泛用于叶片目标图像的纹理特征提取<sup>[9]</sup>。其计算方式为:设定某个像素点为中心点,以中心点半径范围为  $R$  的相邻像素点为研究对象,如果相邻点的像素强度大于中心点则标记为 1,否则标记为 0,然后从某一个角度开

前 15 次 MC 采样后波长数量有了明显的下降,之后逐渐趋于平缓,整个筛选过程先后经历了“粗选”和“精选”这两个阶段。图 4(b)为 5 折交叉验证均方根误差变化趋势图,在前 35 次采样中,误差逐步减小,35 次之后误差保持稳定,说明此时已没有冗余波长可筛除。图 4(c)是回归系数随采样次数变化图,当采样次数为 35 次时,该系数的回归趋势已经完全确立。因此以 35 次采样时的波长作为特征波长,最终得到 4 个特征波长分别为 493, 512, 586 和 613 nm。

### 2.2 高光谱图像特征提取

高光谱图像的空间分辨率较高,可较清晰的反应病害叶片的图像特征,通过提取病害叶片的图像特征并联合荧光光谱信息可以强化病害信息,提高预测模型的预测性能。为了提高特征提取的稳定性,本实验中以 4 个特征值下对应的高光谱图像为研究对象,分别提取 4 幅图像的纹理信息,然后取其平均值作为该样本最终的纹理特征值,以赤叶中期病害为例,4 个特征值下的高光谱图像如图 5 所示。

始,以下一个相邻点为中心继续计算 LBP 值。最常用的模式是  $P=8$ ,半径  $R=1$ , $P=8$  表示中心点相邻像素点的个数为 8, $R=1$  表示中心点到相邻点的像素距离为 1。标记完成后,按顺时针或者逆时针方向将相邻像素的二进制值进行组合,然后将二进制值转换成十进制值,其实现过程如图 6 所示。

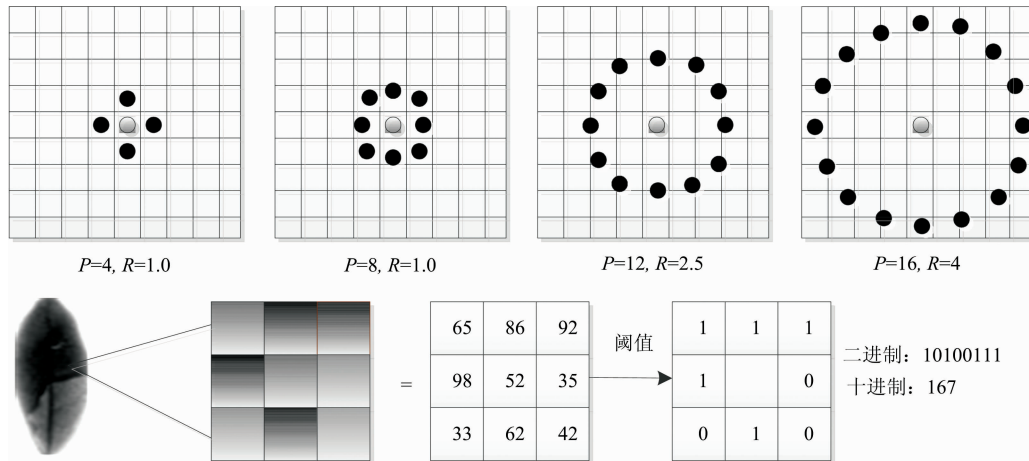


图 6 LBP 编码原理图

Fig. 6 Schematic diagram of LBP coding

图 6 上半部分为不同相邻点个数和半径下统计对象的分布示意图,从图中可以看出,相邻点数越多,其能量越强,半径选取越大,统计值之间的相关性越小,本实验中采用的模式是  $P=8$ , $R=1$ 。图中下半部分为图像中某像素点 LBP 值的编码过程,以像素点自身的像素值为参照,大于该像素

值的标记为 1,小于该像素值的标记为 0,按顺时针方向进行二进制编码得到 10100111,对应十进制值 167 即为该像素点 LBP 值。

图 7 是健康叶片 493nm 特征值下高光谱图像提取 LBP 值的过程,由于邻域有 8 个像素点的 LBP 算子会产生  $2^8 =$



64 种 LBP 值, 过多的表示维数容易产生信息冗余, 不利于纹理特征的表达<sup>[10]</sup>, 因此实验中使用 Uniform 模式对 LBP 模型进行降维。当对应于局部二进制模式的二进制数 0 和 1 之间的跳跃小于或者等于 2 次时, 对应的二进制被称为 Uni-

form 模式, 对于 8 个采集像素点, 二进制模式由原来的 256 个减少到 59 个。将特征提取图像分成 4×4 大小的 16 个区域, 对每个区域分别使用 Uniform 模式进行 LBP 值提取, 一共得到 16×59=944 个维度的 LBP 特征值。

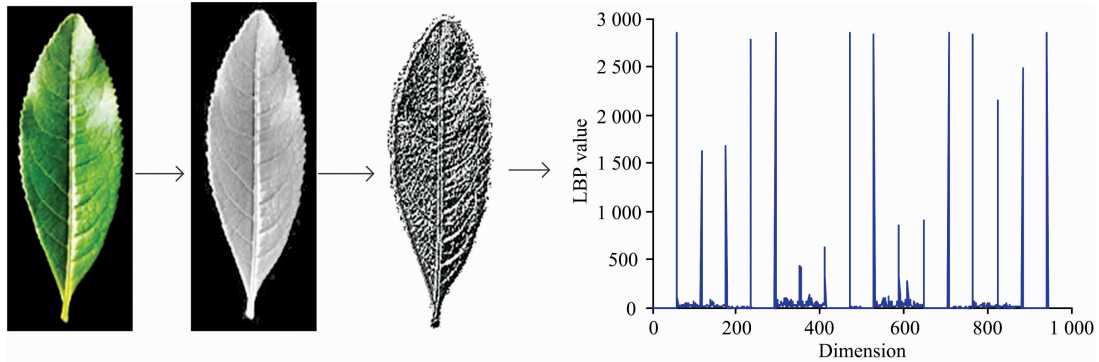


图 7 LBP 值提取过程

Fig. 7 Extraction process of LBP value

### 2.2.2 GLCM 特征

灰度共生矩阵 (GLCM) 用于计算相邻像素之间不同灰度值的组合数目, 以此反映在指定方向上的梯度信息<sup>[11]</sup>。在灰度梯度信息基础上可以进一步提取对比度、能量、逆差矩、相关性和熵信息作为图像纹理特征的典型性信息。这 5 个表征量对应的计算方法如式(1)~式(5)所示

$$f_1 = \sum_{x=y=0}^{L-1} (x-y)^2 \left\{ \sum_{x=0}^{L-1} \sum_{y=0}^{L-1} P(x, y) \right\} \quad (1)$$

$$f_2 = \sum_{x=0}^{L-1} \sum_{y=0}^{L-1} p(x, y)^2 \quad (2)$$

$$f_3 = \sum_{x=0}^{L-1} \sum_{y=0}^{L-1} \frac{p(x, y)}{1 + (x-y)^2} \quad (3)$$

$$f_4 = \frac{\sum_{x=0}^{L-1} \sum_{y=0}^{L-1} xyp(x, y) - u_1 u_2}{\delta_1^2 \delta_2^2} \quad (4)$$

$$f_5 = - \sum_{x=0}^{L-1} \sum_{y=0}^{L-1} p(x, y) \log p(x, y) \quad (5)$$

实验中将图片转换成 16 个灰度级, 先计算 4 个方向(0°, 45°, 90°及 135°)的灰度共生矩阵, 再计算 5 个共生矩阵的均值和方差, 最终得到 10 个特征值。

### 2.3 建模算法

单隐含层前馈神经网络以其良好的学习性能在众多领域得到广泛应用。然而传统学习算法(如 BP 算法等)大多采用梯度下降方法<sup>[12]</sup>, 普遍存在训练速度慢、无法达到全局最小、对学习率的选择敏感等缺点。极限学习机 (ELM) 用于输入层与隐含层之间连接的权值和阈值是随机产生的, 在整个训练过程中只需设定好隐含层的神经元个数即可, 不需要再对这两个参数进行动态调整就可以获得唯一最优解<sup>[13]</sup>。与传统分类方法相比, 学习速度快且泛化性能强, 因此, 本实验采用 ELM 算法作为多病期茶叶赤叶病的分类预测算法。由于模型的输入特征值不在一个量纲, 首先对输入特征值进行归一化处理, 然后再定义模型的输出标签, 即健康叶片的预测模型输出为 1, 赤叶早期为 2, 中期为 3。为了便于对实

验结果进行分析, 实验时将样本数据分成 3 个组合。组合 1 为光谱特征值, 组合 2 为光谱信息加 4 个特征波长下高光谱图像纹理信息的 LBP 均值, 组合 3 为光谱信息加高光谱图像的灰度共生矩特征值(5 个共生矩阵的均值和方差), 最终得到的分类效果如表 1 所示。

表 1 不同组合下分类效果对比

Table 1 Comparison of classification effects under different combinations

模型	病期	训练集	预测集	预测分类正确率/%
光谱特征值	健康	30/27	15/11	72.73
	初期	40/36	20/14	
	中期	40/35	20/15	
光谱特征值 LBP 特征值	健康	30/28	15/13	85.45
	初期	40/39	20/16	
	中期	40/38	20/18	
光谱特征值灰度 共生矩阵特征值	健康	30/27	15/12	81.82
	初期	40/37	20/16	
	中期	40/36	20/17	

从表 1 中可以看出基于光谱特征值的模型预测分类识别率最低, 只有 72.72%, 基于光谱特征值+LBP 特征值组合的模型预测分类正确率相对较高, 为 85.45%。从病期识别率上看, 组合 2 和组合 3 在各病期的识别率上均全面高于组合 1, 说明纹理信息的加入强化了病期特征, 有效提升了模型的预测性能。比较组合 2 和组合 3 的预测结果发现, 在健康叶片的预测上, 组合 2 和组合 3 的性能相当, 而在病早期和病中期的叶片分类预测上, 组合 2 的性能优于组合 1。因为实验时是以赤叶病纹占叶面比 1/3 作为病早期和病中期的划分界线, 说明 LBP 算子在病纹特征的提取上更为精确, 能够捕捉更多的局部纹理信息, 而基于灰度共生矩的病纹特征提取精度相对较低, 这可能和赤叶病害块状分布的特性有

关, 块状分布下局部特征的灰度梯度变化不是十分显著, 导致纹理信息提取量不够充分。

虽然组合 2 的预测分类正确率最优, 但是仍达不到预期的效果, 需要对模型进一步优化。实验中极限学习机隐含层采用的是默认 Sigmoid 函数作为激活函数, 为了更容易得到状态集合中每个状态的值, 该函数在做函数值判决时使用双侧抑制的机制<sup>[14]</sup>, 即设定在未达到目标点前, 环境反馈的赏金为负(正), 到达目标时为 0, 而本实验中分类问题的输出逼近值是单调的, 采用双侧抑制反而会增加废运算, 影响模型预测性能, 因此采用单边抑制机理的 Softplus 函数替代 Sigmoid 函数, 替换后主要产生如下变化: (1) 相对宽阔的兴

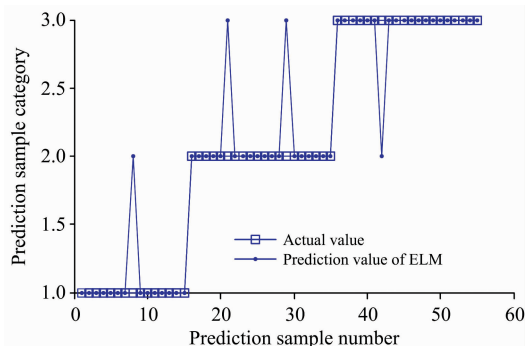


图 8 优化后预测集分类结果

Fig. 8 The classification result of prediction set after optimization

奋边界以及稀疏激活性。(2) 由于 Softplus 函数更接近生物激活模型, 使得模型网络的平均性能更好, 从而提高了泛化性能。优化后的模型预测准确率达到 92.73%, 如图 8 所示。

### 3 结 论

通过高光谱荧光透射技术对茶叶赤叶病不同病期的叶片样本进行了荧光光谱采集, 通过对平均光谱的形态分布和机理分析, 证明了使用荧光光谱对不同病期茶叶赤叶病害进行分类预测的可行性。为了提高后期建模的效率, 使用竞争自适应加权抽样(CARS)算法从全波长信息中优选出 4 个特征波长, 为进一步强化不同病期的病害特征, 采集了 4 个特征波长下反射光模式下病害叶片的高光谱图像, 分别使用 LBP、GLCM 算法提取图像中的纹理信息, 并取其平均值作为模型的输入特征值, 最终通过 ELM 建立预测分类模型。通过对比实验发现, LBP+ELM 预测模型性能最好, 分类正确率为 85.45%。由于没有达到预期效果, 利用 Softplus 函数对 ELM 的隐含层激活函数进行了优化, 替换掉原来的 Sigmoid 函数, 优化后的模型预测分类正确率达到 92.73%, 基本能够满足应用需求。说明利用 CARS-LBP-ELM 模型能够对不同病期的茶叶赤叶病进行有效的预测分类, 该方法可为类似叶类作物的病害预测及相关装备的研制提供一定的参考依据。

### References

- [1] Atta B M, Saleem M, Ali H, et al. *Laser Physics*, 2018, 28(6): 065607.
- [2] SUI Yuan-yuan, WANG Qing-yu, YU Hai-ye(隋媛媛, 王庆钰, 于海业). *Spectroscopy and Spectral Analysis(光谱学与光谱分析)*, 2016, 36(6): 1779.
- [3] Xie C, He Y. *Sensors*, 2016, 16(5): 676.
- [4] ZHANG Shuai-tang, WANG Zi-yan, ZOU Xiu-guo, et al(张帅堂, 王紫烟, 邹修国, 等). *Transactions of the Chinese Society of Agricultural Engineering(农业工程学报)*, 2017, 33(22): 200.
- [5] CHEN Bing, WANG Ke-ru, LI Shao-kun, et al(陈兵, 王克如, 李少昆, 等). *Transactions of the Chinese Society of Agricultural Engineering(农业工程学报)*, 2011, (9): 86.
- [6] ZHU Ya-xing, YU Lei, HONG Yong-sheng, et al(朱亚星, 于雷, 洪永胜, 等). *Scientia Agricultura Sinica(中国农业科学)*, 2017, 50(22): 4325.
- [7] SUN Jun, CONG Sun-li, MAO Han-ping, et al(孙俊, 丛孙丽, 毛罕平, 等). *Transactions of the Chinese Society of Agricultural Engineering(农业工程学报)*, 2017, 33(5): 178.
- [8] Wu L, Wang B, Yin Y, et al. *Analytical Letters*, 2016, 49(14): 2290.
- [9] ZHENG Yi-li, ZHONG Gang-liang, WANG Qiang, et al(郑一力, 钟刚亮, 王强, 等). *Transactions of the Chinese Society for Agricultural Machinery(农业机械学报)*, 2017, 48(3): 30.
- [10] SUN Jun, JIN Xia-ming, MAO Han-ping, et al(孙俊, 金夏明, 毛罕平, 等). *Transactions of the Chinese Society of Agricultural Engineering(农业工程学报)*, 2014, 30(10): 167.
- [11] Malegori C, Franzetti L, Guidetti R, et al. *Journal of Food Engineering*, 2016, 185: 48.
- [12] Azar Y, Cohen S. *Operations Research Letters*, 2018, 46(1): 128.
- [13] Bian Xihui, Li Shujuan, Fan Mengran, et al. *Analytical Methods*, 2016, 8(23): 4674.
- [14] XU Yuan, HUANG Bing-ming, HE Yan-lin(徐圆, 黄兵明, 贺彦林). *CIESC Journal(化工学报)*, 2017, 68(3): 916.

# Prediction of Tea Diseases Based on Fluorescence Transmission Spectrum and Texture of Hyperspectral Image

LU Bing<sup>1,2</sup>, SUN Jun<sup>1\*</sup>, YANG Ning<sup>1</sup>, WU Xiao-hong<sup>1</sup>, ZHOU Xin<sup>1</sup>

1. School of Electrical and Information Engineering, Jiangsu University, Zhenjiang 212013, China

2. Information Center, Jiangsu University, Zhenjiang 212013, China

**Abstract** In order to realize accurate prediction of tea disease and avoid secondary damage in the process of disease feature extraction, the fluorescence transmission technology was used to study the spectrum characteristics of tea red leaf disease. The total of 45 samples of healthy tea leaves, 60 samples of early stage of red leaf disease and 60 samples of intermediate stage of red leaf disease were collected in the experiment, and were divided to training set and prediction set according to the proportion of 2 : 1 for each kind. The original fluorescence transmission spectra of these leaves were collected using hyperspectral instrument by fluorescence transmission. Through the analysis of average spectral intensity curves of the three groups of leaves, the feasibility of using fluorescence transmission spectral information to classify the three types of leaves was confirmed. Then the polynomial smoothing (Savitzky-Golay, S-G) method was carried out for smoothing and noise reduction on the original spectral. Finally, competitive adaptive reweighted sampling (CARS) algorithm was used to select the characteristic wavelengths of the preprocessed spectral data. After 50 weighted samples, 4 characteristic wavelengths were selected finally, which were 463, 512, 586 and 613 nm respectively. In order to maximize the disease feature information of the samples and strengthen the typification of the classifier input value of disease feature, hyperspectral images were collected on 4 characteristic wavelengths respectively. Gray level co-occurrence matrix (GLCM) algorithm was used to extract image texture information, and 0°, 45°, 90° and 135° direction of the four gray level co-occurrence matrix were calculated. Then, the mean value and square error of the five symbiotic matrices were calculated, and the average value of the four image texture information was taken as the texture feature value of the leaf in order to enhance the recklessness. Finally, 10 feature values were obtained. The LBP (Local binary patterns) algorithm was used to extract the texture information from spectral image, and the uniform mode was used to reduce the dimension of LBP mode. Eventually, 944 dimension characteristic values of LBP were got from each image, similarly, the average value of the four images was taken as the characteristic value of LBP texture. Finally, the LBP eigenvalues of 944 dimensions were obtained for each image, and the average value of 4 images was also taken as the LBP texture feature value of the leaf. Finally, the prediction model was established under characteristic spectrum associated with the gray level co-occurrence matrix and the LBP operator respectively by using the extreme learning machine (ELM). As the input eigenvalues of the model were not in the same dimension, the input eigenvalues were normalized firstly, and then the output labels of the model were defined, that is, the output of the prediction model of healthy leaves was 1, the early stage of red leaf disease was 2, and the intermediate stage of red leaf disease was 3. The prediction accuracy based on CARS-GLCM-ELM model was 81.82%, and the prediction accuracy of CARS-LBP-ELM model was 85.45%. It showed that the effect of combining fluorescence transmission spectrum with LBP operator texture information was better. Due to the undesired results, the hidden layer activation function in ELM was optimized by using Softplus function instead of Sigmod function. The prediction accuracy of the optimized model was 92.73%. In this study, the fluorescence spectrum information of diseased leaves and texture information of hyperspectral images at corresponding characteristic wavelengths were fused, and the results can provide some reference for rapid and accurate prediction of tea diseases.

**Keywords** Fluorescence transmission; Hyperspectral image; Feature fusion; Gray level co-occurrence matrix; LBP algorithm; ELM

(Received Jul. 24, 2018; accepted Nov. 20, 2018)

\* Corresponding author