

基于近红外光谱的 SSA-RELM 的菠萝 含水率快速检测

Rapid detection of moisture content of pineapple based on near infrared spectroscopy and SSA-RELM

赵艳莉¹ 赵 倩² 李志强³

ZHAO Yanli¹ ZHAO Qian² LI Zhiqiang³

(1. 郑州财税金融职业学院, 河南 郑州 450000; 2. 河南科技大学, 河南 洛阳 471000;

3. 郑州大学, 河南 郑州 450001)

(1. Zhengzhou Vocational College of Finance and Taxation, Zhengzhou, Henan 450000, China;

2. Henan University of Science and Technology, Luoyang, Henan 471000, China;

3. Zhengzhou University, Zhengzhou, Henan 450001, China)

摘要:目的: 建立快速无损检测菠萝含水率的方法。方法: 提出一种基于连续投影法的特征波长选择和麻雀搜索算法(SSA)优化正则化极限学习机(RELM)的菠萝含水率检测模型。针对菠萝近红外光谱数据具有维度高、冗余信息多的特点, 分别对比连续投影法、主成分分析法和全波段等筛选特征波长的结果, 确定菠萝近红外光谱特征波长筛选方法; 针对 RELM 模型性能受其输入层权值和隐含层偏置的影响, 运用麻雀搜索算法优化 RELM 模型的输入层权值和隐含层偏置, 提出一种基于麻雀搜索算法改进正则化极限学习机的菠萝含水率检测模型。结果: 与遗传算法改进正则化极限学习机(GA-RELM)、粒子群算法改进正则化极限学习机(PSO-RELM)和 RELM 相比, 基于麻雀算法改进正则化极限学习机(SSA-RELM)的菠萝含水率检测模型的检测精度最高。结论: 麻雀搜索算法优化 RELM 模型可以有效提高 RELM 模型的菠萝含水率检测精度。

关键词: 近红外光谱; 菠萝; 含水率; 正则化极限学习机; 麻雀搜索算法; 特征波长; 连续投影法

Abstract: Objective: A method for a fast and non-destructive detection of pineapple moisture content was established.

Methods: A novel detection model of pineapple moisture content was proposed based on continuous projection feature wavelength

selection and Sparrow search algorithm. Firstly, according to the characteristic of pineapple NIR data with high dimension and redundant information, the results of feature wavelength selection such as successive projections algorithm, principal component analysis and full-band were compared, the selection method of characteristic wavelength of pineapple near infrared spectrum was determined. Secondly, considering that the performance of RELM model was affected by the selection of input layer weight and hidden layer bias, the sparrow search algorithm was used to optimize the input layer weight and hidden layer bias of RELM model, a novel pineapple moisture content detection model based on RELM model improved by sparrow search algorithm was proposed. **Results:** compared with GA-RELM, PSO-RELM and RELM, the detection model based on SSA-RELM had the highest detection accuracy. **Conclusion:** RELM model is optimized by sparrow search algorithm can effectively improve the detection accuracy of RELM model.

Keywords: near infrared spectroscopy; pineapple; moisture content; regularized extreme learning machine; sparrow search algorithm; characteristic wavelength; successive projections algorithm

基金项目: 河南省重点研发与推广专项支持项目(编号: 20HN91405); 河南省自然科学基金项目(编号: 2210016)

作者简介: 赵艳莉(1965—), 女, 郑州财税金融职业学院副教授。
E-mail: meetzyl@163.com

收稿日期: 2023-03-16 **改回日期:** 2023-10-31

菠萝, 又称凤梨, 属于热带水果之一。近年来, 芒果干、菠萝干等水果干已成为家家户户喜爱的零食之一, 其营养成分和口感不亚于新鲜水果。水果干制时, 水分与糖分含量是评价其品质的重要指标之一, 所以在收购和存贮果干过程中快速检测其含水率非常重要。传统的化学检测方法存在工作繁琐、耗费时间很长、成本高的缺点, 虽然这些检测方法的准确度基本上达到了要求, 但检

测的准确性与周围环境有很大关系,已经不能满足现代市场经济发展的需求。近红外光谱技术是近几年发展起来的一项新技术,被广泛应用于各个领域^[1-3]。相对于常规的含水率检测方法,近红外光谱检测技术有操作简单、成本低、仪器占地面积小等优点,而且还可以实现无前处理、无损快速、无污染、可在线检测和多组分同时检测。

为了减少近红外光谱中大量冗余信息,降低复杂度以及提高菠萝含水率检测模型的精度,研究拟提出一种基于连续投影法和麻雀搜索算法^[4](sparrow search algorithm, SSA)优化正则化极限学习机^[5](regularized extreme learning machine, RELM)的菠萝含水率检测模型。

1 材料与方法

1.1 试验材料与仪器

试验所用菠萝于 2023 年 3 月份采自广东徐闻,切成厚度为 9, 10 mm 的菠萝片,立即用真空包装袋封存带回

实验室存于冰箱中。试验时,取出菠萝样本于室温(20℃左右)下静置,待菠萝温度达到室温时,并用干净的抹布或纸巾将其表面水分擦拭干净。其中厚度为 9 mm 的菠萝采样 110 份,厚度为 10 mm 的菠萝采样 109 份,共取得 219 份样品。将样品的 70% 用作训练集,剩下的 30% 用作测试集。

1.2 试验仪器

手持便携式反射型光谱仪: NIR-S-G1 型,布鲁克光谱(BRUKER OPTICS)有限公司;

卤素水分测定仪: VM-E01 型,江苏维科特仪器仪表有限公司。

1.3 方法

1.3.1 近红外光谱分析 近红外光波长介于可见光和中红外之间,范围处于 780~2 526 nm,光谱分布区间如图 1 所示^[6-7]。近红外光谱区与含氢基团(O—H、N—H、C—H)振动的合频和各级倍频的吸收区一致,因此扫描样品的近红外光谱可以得到样品的结构信息和组成信息。

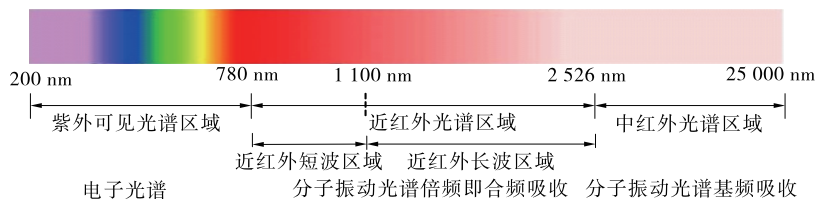


图 1 光谱分布区间

Figure 1 Spectral distribution interval

光谱分析时,光源和检测器放置于同一侧,光源发出光照射于待检测样品,光经过待检测样品返回检测器,计算出接收光与发射光的比值,即反射率。由于待测样品的大小基本不会影响反射率的大小,因此近红外光谱检测易实现在线检测。

1.3.2 建模思路 基于近红外光谱的麻雀搜索算法改进正则化极限学习机(SSA-RELM)的菠萝含水率快速检测的建模思路:

(1) 选择试验所用的菠萝样品。

(2) 菠萝样品切片,制成一定厚度的菠萝片,运用 NIR-S-G1 手持便携式反射型光谱仪采集菠萝样品的近红外光谱数据。

(3) 对菠萝样品的近红外光谱数据进行数据预处理(多元散射校正以及特征波长筛选)。

(4) 建立菠萝含水率的 SSA-RELM 快速检测模型。

(5) 菠萝含水率快速检测模型的验证。菠萝含水率快速检测建模流程图如图 2 所示。

1.3.3 SSA 算法 在标准 SSA 中, N 只麻雀组成的种群矩阵为:

$$\mathbf{X} = [x_1, x_2, \dots, x_N]^T, x_i = [x_{i,1}, x_{i,2}, \dots, x_{i,d}], i = 1, 2, \dots, N, \quad (1)$$

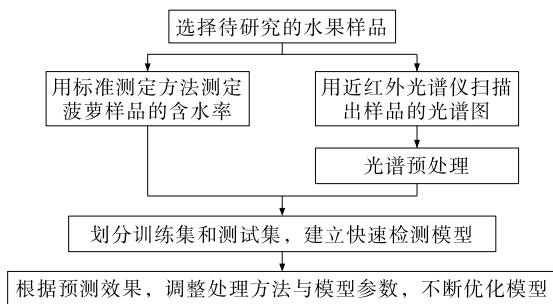


图 2 菠萝含水率快速检测建模流程图

Figure 2 Flow chart of rapid detection of moisture content of pineapple

式中:

$\mathbf{X}$ ——麻雀种群位置向量矩阵;

x_i ——第 i 个麻雀个体位置;

T ——矩阵转置符号;

d ——优化问题的维数;

N ——麻雀的种群大小。

麻雀种群的适应度值矩阵 $\mathbf{F}_x$ 可表示为:

$$\mathbf{F}_x = [f(x_1), f(x_2), \dots, f(x_N)]^T, f(x_i) = [f(x_{i,1}), f(x_{i,2}), \dots, f(x_{i,d})], \quad (2)$$

式中:

$f(x_i)$ ——第 i 只麻雀的适应度值。

麻雀个体的适应度值越高,其获取食物的优先级就越高。适应度高的麻雀个体会被当作发现者,带领麻雀群体靠近食物源。发现者的位置更新数学公式可表示为^[8]:

$$X_{i,j}^{t+1} = \begin{cases} X_{i,j}^t \times \exp\left(-\frac{i}{\alpha \times T_{\max}}\right), R_2 < S_T \\ X_{i,j}^t + Q \times L, R_2 \geq S_T \end{cases}, \quad (3)$$

式中:

t ——麻雀搜索算法的当前迭代次数;

$T_{\max}$ ——麻雀搜索算法的最大迭代次数;

$X_{i,j}^t$ ——第 t 次迭代时第 i 只麻雀的第 j 维位置

信息;

α ——随机数, $\alpha \in (0, 1)$;

R_2 ——预警值, $R_2 \in [0, 1]$;

S_T ——安全值, $S_T \in [0.5, 1]$;

Q ——正态分布的随机数,且 $Q \in [0, 1]$;

L ——元素全为 1 的 $1 \times d$ 的矩阵。

当 $R_2 < S_T$ 时,觅食环境附近没有天敌(捕食者),此时发现者执行广泛搜索策略;当 $R_2 \geq S_T$ 时,有部分麻雀发现觅食环境附近有天敌(捕食者),发出预警,麻雀群体飞往安全区域觅食。

跟随者的位置更新数学公式可表示为^[9-10]:

$$X_{i,j}^{t+1} = \begin{cases} Q \times \exp\left(\frac{X_{\text{worst}}^t - X_{i,j}^t}{i^2}\right), i > \frac{N}{2} \\ X_p^{t+1} + |X_{i,j}^t - X_p^{t+1}| \times A^+ \times L, i \leq \frac{N}{2} \end{cases}, \quad (4)$$

式中:

X_{worst}^t ——第 t 次迭代时全局最差的位置;

X_p^{t+1} ——第 $t+1$ 次迭代时发现者的最佳位置;

A^+ —— $1 \times d$ 的矩阵,矩阵中的元素随机赋值为 1 或 -1,其中 $A^+ = A^T (AA^T)^{-1}$ 。

当 $i > \frac{N}{2}$ 时,适应度值较差的第 i 只跟随者没有得到食物,将飞往其他地方进行觅食。

在标准的 SSA 算法中,一般随机选择占麻雀总数量的 10%~20% 的 F 只麻雀作为预警者,预警者的位置更新数学公式可表示为:

$$X_{i,j}^{t+1} = \begin{cases} X_{\text{best}}^t + \beta \times |X_{i,j}^t - X_{\text{best}}^t|, f_i > f_g \\ X_{i,j}^t + K \times \left(\frac{X_{i,j}^t - X_{\text{worst}}^t}{(f_i - f_w) + \epsilon} \right), f_i = f_g \end{cases}, \quad (5)$$

式中:

X_{best}^t ——迭代次数 t 时的全局最佳位置;

β ——步长控制参数, $\beta \in (0, 1)$;

K ——运动方位控制参数, $K \in [-1, 1]$;

ϵ ——最小常数,防止分母出现 0 的情况;

f_i 、 f_g 和 f_w ——当算法迭代次数到第 i 次的适应度

值、最优适应度值和最差适应度值。

当 $f_i > f_g$ 时,麻雀个体位于麻雀种群的边缘,易被捕食者攻击;当 $f_i = f_g$ 时,处于种群中心位置的麻雀个体意识到危险,向其他区域的麻雀个体靠近。

1.3.4 RELM 模型 极限学习机(extreme learning machine, ELM)的数学模型为^[11]:

$$f_H(x) = \sum_{i=1}^H \beta_i G(a_i, b_i, x), \quad (6)$$

式中:

H ——隐含层神经元数;

a_i, b_i ——第 i 个隐含层神经元的输入权重和偏置,

其中 $a_i \in R^n, b_i \in R$;

$G(a_i, b_i, x)$ ——第 i 个隐含层神经元的输出函数;

β_i ——第 i 个隐含层神经元对于输出神经元的权重系数,其中 $\beta_i \in R$ 。

$G(a_i, b_i, x)$ 输出函数可表示为:

$$G(a_i, b_i, x) = g(a_i \cdot x + b_i), \quad (7)$$

式中:

$g(\cdot)$ ——激活函数。

假设 N 组训练集数据 $\{(x_l, y_l)\}_{l=1}^N$, x_l, y_l 分别为输入变量和输出变量,且 $x_l \in R^n, y_l \in R^m$, 则:

$$y_l = \sum_{i=1}^H \beta_i G(a_i, b_i, x_l), l = 1, 2, \dots, N. \quad (8)$$

式(8)的矩阵形式为:

$$W\beta = Y, \quad (9)$$

式中:

$$W =$$

$$\begin{bmatrix} G(a_1, b_1, x_1) & \cdots & G(a_L, b_L, x_1) \\ \vdots & & \vdots \\ G(a_1, b_1, x_N) & \cdots & G(a_L, b_L, x_N) \end{bmatrix} = \begin{bmatrix} h(x_1) \\ \vdots \\ h(x_N) \end{bmatrix};$$

$$\beta = \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_L \end{bmatrix}_{L \times m};$$

$$Y = \begin{bmatrix} y_1 \\ \vdots \\ y_N \end{bmatrix}_{N \times m}.$$

因为 a_i, b_i 是随机生成的,只有 β 由 $\{(x_l, y_l)\}_{l=1}^N$ 计算获得的,为提高 ELM 模型的性能,将正则化因子 λ 引入 ELM 模型,这样求解 β 的问题就转化为求解式(10)的优化问题^[12]:

$$\min_{\beta} V_{\text{ELM}} = \frac{1}{2} \|\beta\|^2 + \frac{\lambda}{2} \sum_{l=1}^N \|\epsilon_l\|^2$$

$$s.t.: h(x_l)\beta = y_l - \epsilon_l, l = 1, 2, \dots, N, \quad (10)$$

式中:

ϵ_l ——训练偏差。

根据 Karush-Kuhn-Tucker 条件,式(10)可转化为式(11)的双重优化问题^[13]:

$$V_{\text{ELM}} = \frac{1}{2} \|\beta\|^2 + \frac{\lambda}{2} \sum_{l=1}^N \|\varepsilon_l\|^2 - \sum_{l=1}^N \sum_{j=1}^m a_{l,j} [h(x_l) \beta_j - y_{l,j} + \varepsilon_{l,j}]. \quad (11)$$

此时 β 的最优解为:

$$\hat{\beta} = W^T \left(\frac{I}{\lambda} + WW^T \right)^{-1} Y. \quad (12)$$

式(12)代入式(8)可得, RELM 模型的拟合回归值为:

$$\hat{y} = \sum_{i=1}^H \hat{\beta}_i G(a_i, b_i, x_l), l = 1, 2, \dots, N. \quad (13)$$

RELM 模型的拟合回归值主要取决于式(12)中的正则化因子 λ 、矩阵 W 和 Y 。

1.3.5 基于 SSA-RELM 的菠萝含水率快速检测模型

(1) 适应度函数: RELM 模型的性能受初始参数(输入权值 a_i 和隐含层偏置 b_i)的影响。为改善 RELM 模型的性能,将 SSA 算法应用于 RELM 模型参数选择,选择式(14)作为 SSA 算法优化 RELM 模型的适应度函数:

$$\min f(a_i, b_i) = \sqrt{\frac{1}{n} \sum_{i=1}^n (x(i) - x_p(i))^2} \\ s.t. \ a_i \in [a_{i\min}, a_{i\max}] \\ b_i \in [b_{i\min}, b_{i\max}], \quad (14)$$

式中:

n ——训练样本数量;

$x(i), x_p(i)$ ——第 i 个样本的实际值和预测值;

$[a_{i\min}, a_{i\max}], [b_{i\min}, b_{i\max}]$ ——第 i 个初始输入权值 a 和隐含层偏置 b 的寻优区间,且 $a \in [-1, 1]$ 和 $b \in [-1, 1]$ 。

(2) 算法步骤:为实现菠萝含水率的快速检测,基于近红外光谱的菠萝含水率 SSA-RELM 快速检测模型的算法步骤可具体描述为:

Step1:读取菠萝近红外光谱数据和含水率,光谱数据预处理以及光谱特征波长选择,归一化处理,并划分训练集和测试集。

Step2:设定 RELM 模型参数和麻雀搜索算法参数并随机初始化种群:发现者比例 P_D 、预警值 S_T 、侦察者比例 S_D 、种群规模 N 、最大迭代次数 $T_{\max}$,假设 RELM 模型的隐含层神经元数为 L ,输入层神经元数为 n ,SSA 算法的初始种群维数 $\dim = L(n+1)$ 。

Step3:计算每个麻雀个体的适应度。将训练集代入 RELM 模型,按式(14)计算每个麻雀个体的适应度并排序,找到当前最优适应度值和最优位置以及最差适应度值和最差位置。

Step4:选择发现者和跟随者,并分别按式(3)和式(4)更新发现者位置和跟随者位置。

Step5:随机选择预警者,并按式(5)更新预警者的位置。

Step6:判断算法终止条件。若满足终止条件,则输出 SSA 算法的最优适应度和最优位置,即对应 RELM 模型的输入层权值和隐含层偏置,将寻优获取的最佳参数代入 RELM 模型训练和测试;否则,返回 Step2。基于近红外光谱的 SSA-RELM 的菠萝含水率快速检测流程如图 3 所示。

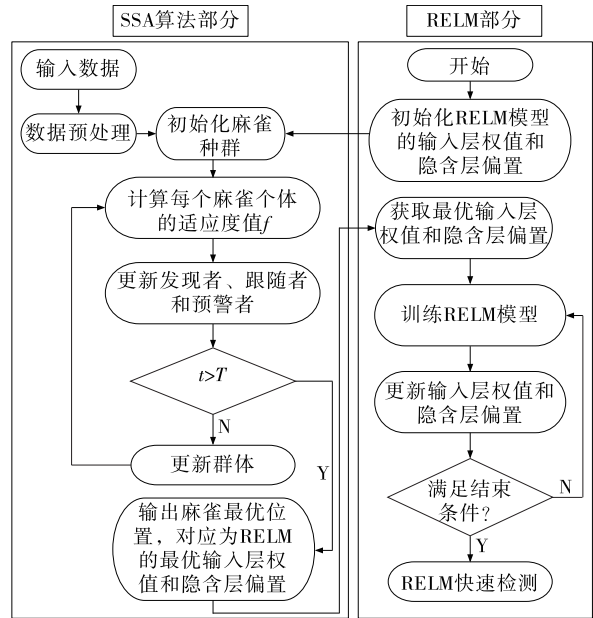


图 3 基于近红外光谱的 SSA-RELM 的菠萝含水率快速检测流程图

Figure 3 Flow chart of rapid detection of moisture content of pineapple based on near infrared spectroscopy and SSA-RELM

2 实证分析

2.1 评价指标

建立模型时用到的模型性能评价指标一般包括训练集的相关系数(R_c)和测试集的相关系数(R_p),训练集的均方根误差(R_{MSEC})和测试集的均方根误差(R_{MSEP})。相关系数 R_c 和 R_p 的值越大,越接近 1,则预测含水率的值与测量含水率的值之间的相关性越好。 R_{MSEC} 用于测试模型对训练集样本的预测能力, R_{MSEC} 值越小,则说明建立的模型的预测效果越好; R_{MSEP} 用于评估模型预测样品目标成分含量的能力,同理 R_{MSEP} 越小,说明该模型对目标信息的预测效果越好。相关系数 R_c 和 R_p ,均方根误差 R_{MSEC} 和 R_{MSEP} 的计算公式为^[14-15]:

$$R_c = \sqrt{1 - \frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}}, \quad (15)$$

$$R_{\text{MSEC}} = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n - f - 1}}, \quad (16)$$

$$R_p = \sqrt{1 - \frac{\sum_{i=1}^m (\hat{y}_i - y_i)^2}{\sum_{i=1}^m (y_i - \bar{y})^2}}, \quad (17)$$

$$R_{MSEP} = \sqrt{\frac{\sum_{i=1}^m (y_i - \hat{y}_i)^2}{m-1}}, \quad (18)$$

式中:

n, m ——训练集和测试集的样本数;

i ——第 i 个样本;

y_i ——第 i 个样本用标准方法测定的含水率实际值;

$\hat{y}_i$ ——第 i 个样本的相应组分的含水率预测值;

$\bar{y}$ ——样本集相应组分的平均值;

f ——RELM 建模采用的主成分数。

2.2 数据预处理

选择多元散射校正(MSC)、移动平均平滑(Smooth)、一阶导数、二阶导数、标准正态变换(SNV)、归一化(Nor)、标准化等多种预处理方法^[12,16]进行对比,找到菠萝含水率最佳预处理方法。原始光谱图和不同预处理光谱图如图4和表1所示。

由表1可知,用原始光谱数据建立的菠萝含水率模型测试集的预测效果是比较好的,但相对于经过预处理之后的预测效果还是较差,其相关系数分别为0.9902和0.9864。由表1可以看出,经过SNV、MSC、归一化(Nor)、标准化和Smooth以及以上方法的组合这几种预处理方法的测试集模型预测效果得到了较高的提升,所

以经过数据预处理一般是能够提高模型预测精度的。其中经过Smooth+MSC共同处理后的水分模型精度最好, R_{MSEC} 由0.0377下降到0.0235,相关系数 R_c 由0.9902提高到0.9961, R_{MSEP} 由0.0388下降到0.0234,相关系数 R_p 由0.9864提高到0.9946。通过对比,经过Smooth+MSC预处理能够提高训练集和测试集模型预测的精度。

2.3 筛选特征波长

针对菠萝样品的近红外光谱数据存在维度高、冗余复杂的特点,在菠萝含水率快速检测模型建立之前,首先运用连续投影法(SPA)、主成分分析法(PCA)和全波段(FS)对近红外光谱数据进行对比分析^[17],从而确定菠萝含水率近红外光谱特征波长的筛选方法。同时为了确定RELM模型隐层神经元数,RELM模型的激活函数为sig函数,初始隐层神经元数为1,不断增加,直到隐层神经元数为50,不同隐层神经元数和不同特征波长筛选的菠萝含水率检测准确率如图5所示。

由图5可知,连续投影法—正则化极限学习机(SPA-RELM)模型的隐层神经元数为30时,菠萝含水率快速检测的准确率最高,故文中RELM模型的隐层神经元数设定为30,特征筛选方式为SPA。

由图6可知,菠萝片厚度为9,10 mm时,其均方根误差和相关系数取得综合最优。

2.4 结果与分析

为验证SSA-RELM进行菠萝含水率快速检测的有效性和可靠性,将SSA-RELM与遗传算法改进RELM(GA-RELM)、粒子群算法改进RELM(PSO-RELM)和

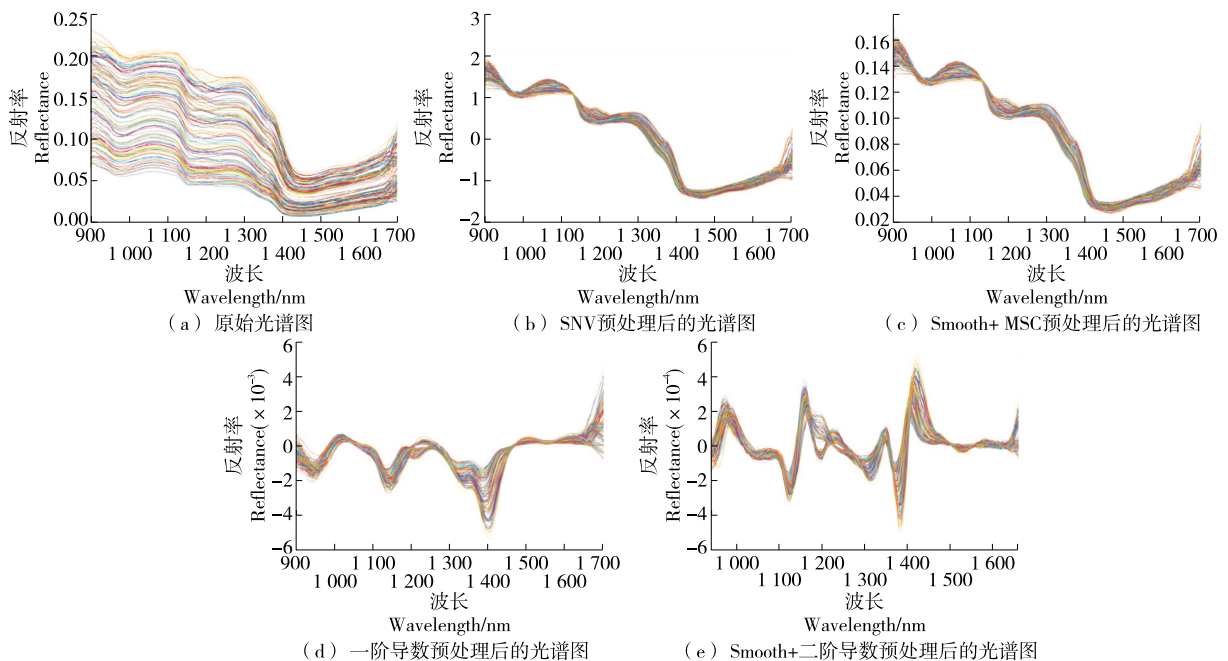


图4 原始光谱图和不同预处理方法处理后的光谱图

Figure 4 Original spectra and spectra after different pretreatment methods

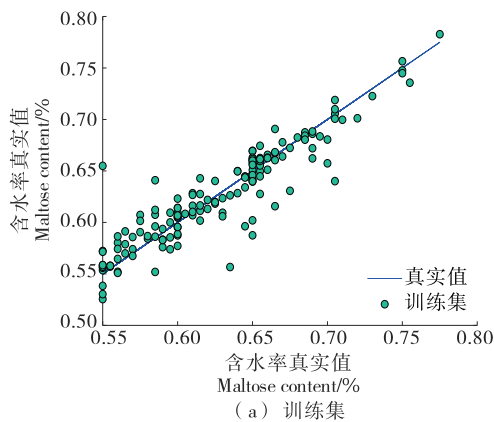
表 1 不同预处理方法的菠萝含水率 RELM 模型检测结果

Table 1 Test results of pineapple moisture content RELM model different from treatment methods

方法	训练集		测试集	
	R_c	R_{MSEC}	R_c	R_{MSEC}
原始光谱	0.990 2	0.037 7	0.986 4	0.038 8
SNV	0.995 8	0.022 5	0.994 6	0.022 5
MSC	0.995 9	0.022 1	0.993 9	0.024 0
Smooth	0.991 2	0.030 3	0.993 8	0.024 5
Nor	0.995 7	0.021 8	0.991 6	0.027 5
Normalize	0.995 8	0.021 7	0.991 6	0.027 5
一阶导数	0.991 8	0.029 8	0.976 6	0.026 1
二阶导数	0.994 6	0.025 2	0.985 4	0.024 6
Smooth+SNV	0.995 3	0.023 5	0.994 6	0.023 7
Smooth+MSC	0.996 1	0.023 5	0.994 7	0.023 4
Smooth+Nor	0.995 0	0.023 7	0.992 6	0.025 9
Smooth+标准化	0.995 1	0.023 6	0.992 5	0.025 8
Smooth+一阶导数	0.992 0	0.029 8	0.970 7	0.030 4
Smooth+二阶导数	0.994 3	0.025 8	0.983 2	0.029 8

RELM 进行比较。通用参数设定:种群规模 $N=10$,最大迭代次数 $T=100$;SSA 算法:发现者比例 $P_D=0.2$ 、预警值 $S_T=0.8$ 、侦察者比例 $S_D=0.2$;遗传算法^[18](genetic algorithm,GA):变异概率 $p_m=0.1$,交叉概率 $p_c=0.7$;粒子群算法^[19](particle swarm optimization algorithm, PSO):学习因子 $c_1=c_2=2$,惯性权重 $w=0.8$;RELM 模型^[20-21]:输入层神经元数量 $N_1=240$ 、隐含层神经元数量 $N_2=30$ 以及输出层神经元数量为 $N_3=1$ 。文中样本共 219 组,按照 70%:30%的比例对 219 份样本进行划分,划分后得到训练集 153 份样本,测试集 66 份样本。菠萝含水率快速检测结果如图 7~图 10 和表 2 所示。

由表 2 可知:



(a) 训练集

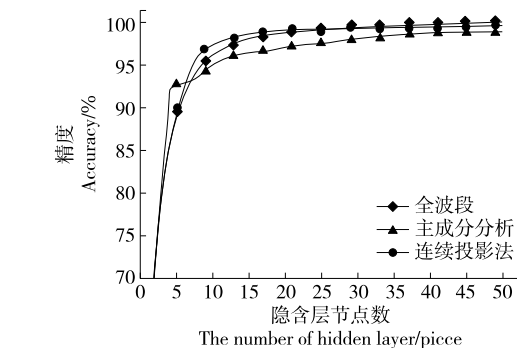


图 5 隐层神经元数与准确率关系图

Figure 5 The relationship between the number of hidden layer nodes and the accuracy

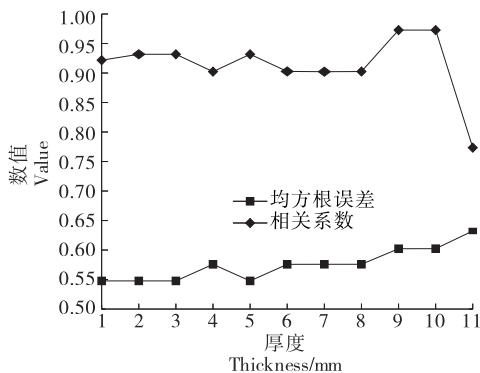
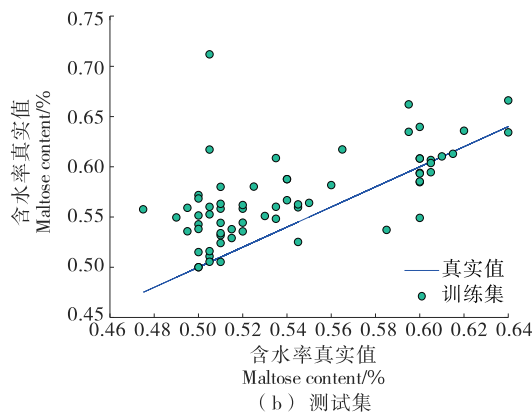


图 6 菠萝片厚度与精度的关系图

Figure 6 Relationship between pineapple slice thickness and precision

(1) 从训练集和测试集的评价指标来看,SSA-RELM 模型的菠萝含水率检测效果优于 GA-RELM、PSO-RELM 和 RELM 等模型,SSA-RELM 模型的 R_{MSEC} 最小且相关系数 R_c 和 R_p 最大,从而说明 SSA-RELM 模型的菠萝含水率检测精度最高,误差最小。

(2) 通过 SSA、GA 和 PSO 等算法对 RELM 模型的参数进行优化选择,可以有效提高 RELM 模型的菠萝含水率检测精度。



(b) 测试集

图 7 SSA-RELM 检测结果图

Figure 7 SSA-RELM test results diagram

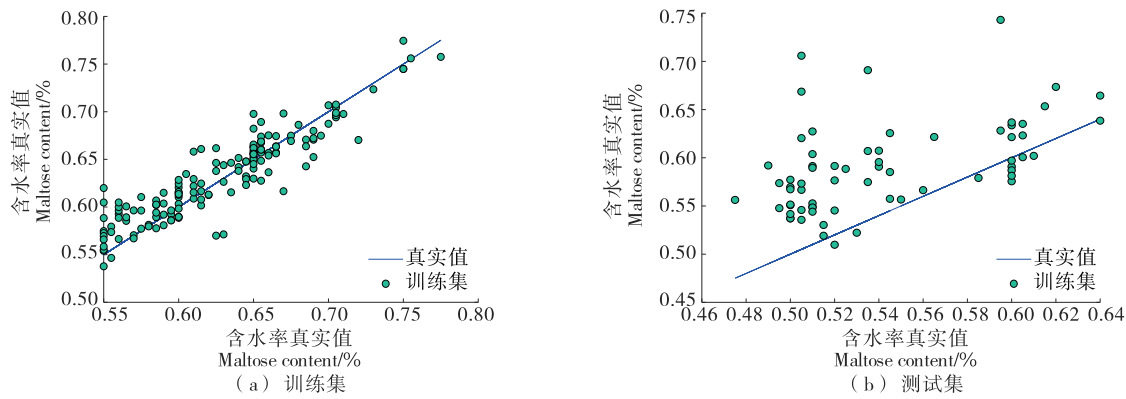


图 8 GA-RELM 检测结果图
Figure 8 GA-RELM test results diagram

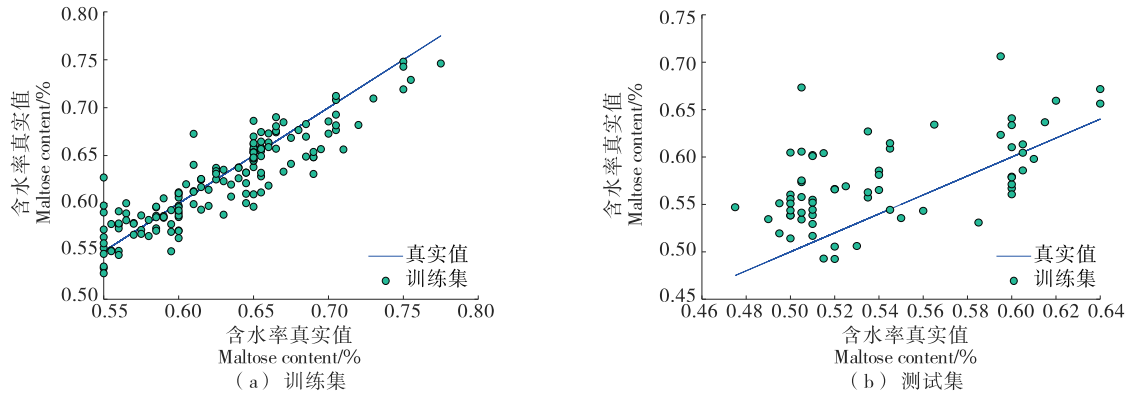


图 9 PSO-RELM 检测结果图
Figure 9 PSO-RELM test results diagram

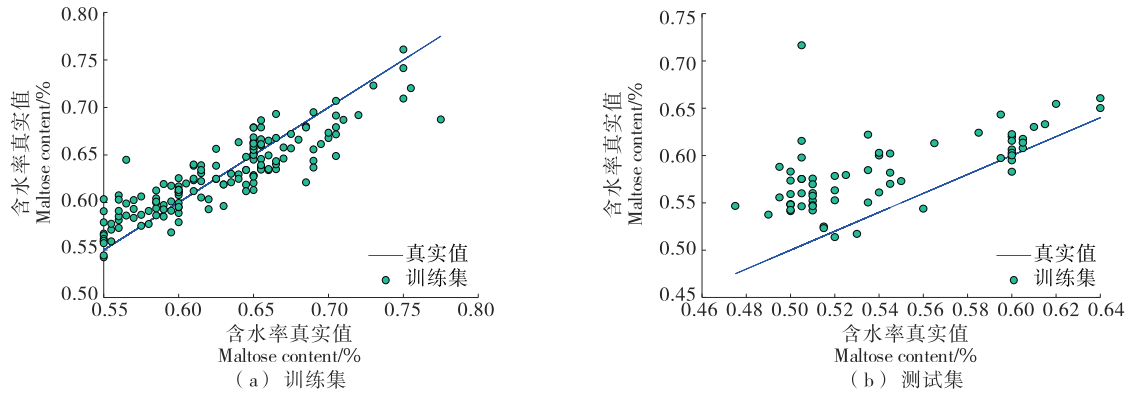


图 10 RELM 检测结果图
Figure 10 RELM test results diagram

表 2 菠萝含水率检测结果

Table 2 Results of moisture content of pineapple

方法	训练集		测试集	
	R_{MSEC}	R_c	R_{MSEP}	R_p
SSA-RELM	0.008 7	0.996 1	0.037 4	0.989 2
GA-RELM	0.016 3	0.978 5	0.039 3	0.971 2
PSO-RELM	0.016 9	0.976 4	0.045 8	0.964 6
RELM	0.017 7	0.965 1	0.047 6	0.945 7

3 结论

为提高菠萝含水率检测精度,提出了一种基于麻雀搜索算法改进正则化极限学习机(SSA-RELM)的菠萝含水率检测模型。研究表明,与遗传算法改进正则化极限学习机(GA-RELM)、粒子群算法改进正则化极限学习机(PSO-RELM)和正则化极限学习机(RELM)相比,SSA-RELM 的菠萝含水率检测模型的检测精度最高。然而通过智能算法优化正则化极限学习机模型的复杂程度

较高,对计算机的硬件配置要求较高,后续将从降低算法复杂度的角度去提升算法的计算效率和运行速度,从而提高算法的实时性,便于实际推广和工程应用。

参考文献

- [1] 刘云刚, 王伟. 基于 SFLA 优化的 BP 神经网络苹果鲜度气味识别系统[J]. 传感器与微系统, 2020, 39(8): 96-99.
LIU Y G, WANG W. Apple fresh odor recognition system based on SFLA optimized BP neural network[J]. Transducer and Microsystem Technologies, 2020, 39(8): 96-99.
- [2] HUANG Y, TIAN K, WU A Q, et al. Feature fusion methods research based on deep belief networks for speech emotion recognition under noise condition [J]. Journal of Ambient Intelligence and Humanized Computing, 2019, 10(5): 1 787-1 798.
- [3] 赵杰文, 张海东, 刘木华. 利用近红外漫反射光谱技术进行苹果糖度无损检测的研究[J]. 农业工程学报, 2005, 21(3): 162-165.
ZHAO J W, ZHANG H D, LIU M H. Non-destructive determination of sugar contents of apples using near infrared diffuse reflectance [J]. Transactions of the Chinese Society of Agricultural Engineering, 2005, 21(3): 162-165.
- [4] ESWARAMOORTHY S, SIVAKUMARAN N, SEKARAN S. Grey wolf optimization based parameter selection for support vector machines [J]. Compel International Journal for Computation & Mathematics in Electrical & Electronic Engineering, 2016, 35(5): 1 513-1 523.
- [5] IOSIFIDIS A, TEFAS A, PITAS I. Regularized extreme learning machine for multi-view semi-supervised action recognition [J]. Neurocomputing, 2014, 145(18): 250-262.
- [6] 介邓飞, 杨杰, 彭雅欣, 等. 基于高光谱技术的柑橘不同部位糖度预测模型研究[J]. 食品与机械, 2017, 33(3): 51-54.
JIE D F, YANG J, PENG Y X, et al. Research on the detection model of sugar content in different position of citrus based on thehyperspectral technology[J]. Food & Machinery, 2017, 33(3): 51-54.
- [7] 杨晓玉, 丁佳兴, 房盟盟, 等. 基于可见/近红外高光谱成像技术的鸡蛋新鲜度无损检测[J]. 食品与机械, 2017, 33(11): 131-136.
YANG X Y, DING J X, FANG M M, et al. Non-destructive determination of eggs freshness by Vis/NIR hyperspectral imaging technology[J]. Food & Machinery, 2017, 33(11): 131-136.
- [8] BORRAZ-MARTÍNEZ S, BOQUÉ R, SIMÓ J, et al. Development of a methodology to analyze leaves from Prunus dulcis varieties using near infrared spectroscopy[J]. Talanta, 2019, 204: 320-328.
- [9] NICOLA B M, THERON K I, LAMMERTYN J. Kernel PLS regression on wavelet transformed NIR spectra for prediction of sugar content of apple[J]. Chemometrics & Intelligent Laboratory Systems, 2007, 85(2): 243-252.
- [10] 郭志明, 赵春江, 黄文倩, 等. 苹果糖度高光谱图像可视化预测的光强度校正方法[J]. 农业机械学报, 2015, 46(7): 227-232.
GUO Z M, ZHAO C J, HUANG W Q, et al. Intensity correction of visualized prediction for sugar content in apple using hyperspectral imaging[J]. Transactions of the Chinese Society for Agricultural Machinery, 2015, 46(7): 227-232.
- [11] SAYED G I, DARWISH A, HASSANIEN A E. A new chaotic multi-verse optimization algorithm for solving engineering optimization problems[J]. Journal of Experimental & Theoretical Artificial Intelligence, 2018, 25(6): 1-25.
- [12] 刘燕德, 周延睿. 基于 GA-LSSVM 的苹果糖度近红外光谱检测[J]. 西北农林科技大学学报(自然科学版), 2013, 41(7): 229-234.
LIU Y D, ZHOU Y R. GA-LSSVM based near infrared spectroscopy detection of apple sugar content [J]. Journal of Northwest A & F University (Natural Science Edition), 2013, 41(7): 229-234.
- [13] DENG W, CHEN L. Color image watermarking using regularized extreme learning machine[J]. Neural Network World, 2010, 20(3): 317-330.
- [14] FANG L J, GUO W C. Nondestructive measurement of sugar content and firmness inkorlafragrant pears by using their dielectric spectra[J]. Modern Food Science and Technology, 2016, 32(5): 295-301.
- [15] 董学锋, 戴连奎, 黄承伟. 结合 PLS-DA 与 SVM 的近红外光谱软测量方法[J]. 浙江大学学报(工学版), 2012, 46(5): 824-829.
DONG X F, DAI L K, HUANG C W. Near-infrared spectroscopy soft-sensing method by combining partial least squares discriminant analysis and support vector machine[J]. Journal of Zhejiang University (Engineering Science), 2012, 46(5): 824-829.
- [16] ALCIN O F, SENGUR A, GHOFrani S, et al. GA-SELM: Greedy algorithms for sparse extreme learning machine[J]. Measurement, 2014, 55(3): 126-132.
- [17] 贺凯迅, 曹鹏飞. 基于智能优化算法的软测量模型建模样本优选及应用[J]. 化工进展, 2018, 37(7): 67-74.
HE K X, CAO P F. Training sample selection method based on intelligent optimization algorithms for soft sensor and its application[J]. Chemical Industry and Engineering Progress, 2018, 37(7): 67-74.
- [18] ZHANG H. Determination of tea polyphenols content in puer tea using near-infrared spectroscopy combined with extreme learning machine and GA-PLS algorithm [J]. Laser & Optoelectronics Progress, 2013, 50(4): 180-186.
- [19] ZHAO J W, OUYANG Q, CHEN Q S, et al. Simultaneous determination of amino acid nitrogen and total acid in soy sauce using near infrared spectroscopy combined with characteristic variables selection[J]. Food Science & Technology International, 2013, 19(4): 305-314.
- [20] 单亚锋, 高振彪. 基于双自适应 AIS-PSO 的瓦斯浓度软测量模型[J]. 计算机仿真, 2020, 37(1): 338-342.
SHAN Y F, GAO Z B. Study on double adaptive AIS-PSO based model for gas concentration soft-sensing[J]. Computer Simulation, 2020, 37(1): 338-342.
- [21] LI C H, LI L L, WU Y, et al. Apple variety identification using near-infrared spectroscopy[J]. Journal of Spectroscopy, 2018, 11(9): 1-7.