

doi:10.6041/j.issn.1000-1298.2021.10.028

基于组合色彩特征的苹果树叶片各生长期氮含量预测

王金星^{1,2} 刘雪梅¹ 刘双喜^{1,3} 权泽堃¹ 徐春保¹ 江浩¹

(1. 山东农业大学机械与电子工程学院, 泰安 271018; 2. 山东省农业装备智能化工程实验室, 泰安 271018;
3. 山东省园艺机械与装备重点实验室, 泰安 271018)

摘要: 为精准预测开花期、幼果期和果实膨大期苹果树叶片的氮含量, 提出一种基于组合色彩特征的苹果树叶片氮含量预测模型。首先, 获取苹果树叶片图像并提取 R 、 G 、 B 单色分量及 14 种色彩组合参数共计 17 种色彩特征, 通过主成分分析提取不同时期苹果树叶片氮含量关键影响因子, 消除原始变量之间的相关性, 降低模型输入向量维度; 其次, 对建立的 PCA-SVM、PCA-BP、PCA-ELM 预测模型在不同时期对苹果树叶片氮含量预测效果与精度进行对比, 得到不同时期最佳的预测模型; 最后, 利用最佳预测模型对不同时期苹果树叶片氮含量进行预测, 并通过自适应遗传算法对最佳预测模型参数进行优化。试验结果表明: 在不同生长期, PCA-SVM 模型的预测精度均高于 PCA-BP、PCA-ELM 模型; 优化后 PCA-SVM 预测模型在开花期、幼果期和果实膨大期的平均绝对误差分别为 0.640、0.558、0.544 g/kg, 平均绝对百分误差分别为 0.057、0.050、0.064 g/kg, 均方根误差分别为 0.800、0.747、0.737 g/kg, 优于优化前预测模型。该模型具有良好的预测性能和泛化能力, 可以为果园精准施肥管理、提升果品品质、避免资源浪费和环境污染提供理论依据。

关键词: 苹果树叶片; 氮含量预测; 色彩特征; 主成分分析; 支持向量机

中图分类号: S661.1 文献标识码: A 文章编号: 1000-1298(2021)10-0272-10

OSID:



Prediction of Nitrogen Content in Apple Leaves in Each Growth Period Based on Combined Color Characteristics

WANG Jinxing^{1,2} LIU Xuemei¹ LIU Shuangxi^{1,3} QUAN Zekun¹ XU Chunbao¹ JIANG Hao¹

(1. College of Mechanical and Electronic Engineering, Shandong Agricultural University, Taian 271018, China
2. Shandong Provincial Engineering Laboratory of Agricultural Equipment Intelligence, Taian 271018, China
3. Shandong Provincial Key Laboratory of Horticultural Machinery and Equipment, Taian 271018, China)

Abstract: In order to accurately predict the nitrogen content in different scales of apple leaves at flowering, young fruit and fruit expansion periods, a combined color characteristics based prediction model of apple leaf nitrogen content was proposed. Firstly, the image of apple leaves was obtained and 17 color features, including R , G , B monochromatic components and 14 color combination parameters were extracted, and the key influencing factors of nitrogen content of apple leaves in different periods were extracted by principal component analysis to eliminate the correlation between the original variables and reduce the input vector dimension of the model. Secondly, the PCA-SVM, PCA-BP and PCA-ELM prediction models were established in different periods, the prediction effect and accuracy of apple leaf nitrogen content were compared, and the best prediction model in different periods was obtained. Finally, the best prediction model was used to predict the nitrogen content of apple leaves in different periods, and the parameters of the best prediction model were optimized by adaptive genetic algorithm. The results showed that the prediction accuracy of PCA-SVM model was higher than that of PCA-BP and PCA-ELM model in different growth periods; the mean absolute error of PCA-SVM prediction model in flowering period, young fruit period and fruit expansion period was 0.640 g/kg, 0.558 g/kg and 0.544 g/kg, and mean absolute percentage error was 0.057 g/kg, 0.050 g/kg and 0.064 g/kg, and root mean square error was 0.800 g/kg, 0.747 g/kg and 0.737 g/kg, which was better than that of the prediction model before optimization. The model had good prediction performance and generalization

收稿日期: 2020-11-10 修回日期: 2021-03-31

基金项目: 国家重点研发计划项目(2016YFD0201104)和国家苹果产业技术体系项目(CARS-27)

作者简介: 王金星(1970—),男,教授,博士生导师,主要从事智能农业装备研究,E-mail: jinxingw@163.com

ability, which can provide theoretical basis for orchard precision fertilization management, improving fruit quality, avoiding resource waste and environmental pollution.

Key words: apple leaf; nitrogen content prediction; color features; PCA; SVM

0 引言

中国是世界上最大的苹果生产国和消费国, 总种植面积及总产量均占世界 50% 以上^[1-2]。氮元素是苹果树体内一种重要的必需元素, 它影响果树体内各种生理生化过程、果树抗性、果品产量、品质以及储藏性状^[3-4]。合理施用氮肥可以促进苹果树健康生长, 提高果实产量及果品品质, 减少农田环境污染与资源浪费。因此, 精准预测果树氮含量对指导果树健康成长有重要作用^[5-7]。叶片是整个树体上对土壤矿质营养反应最敏感的器官, 苹果叶片营养分析、诊断技术可以作为苹果树氮素营养盈亏并指导精准施肥的依据^[8-9]。传统的作物氮素检测方法主要有田间采样室内化学分析测定法和叶色卡法, 虽然结果准确可靠, 但分析成本高、检验周期长、时效性差^[10]。近年来, 机器学习及数字图像处理技术的快速发展为作物氮素营养快速、无损和准确监测提供了有效方法与手段^[11]。

目前, 国内外学者针对作物氮含量预测方法已经进行了大量的研究^[12-14], 取得了一定成果。近年一些学者尝试将支持向量机(SVM)方法^[15-17]应用到植被理化参数的反演中, 梁栋等^[18]利用 SVM 回归方法估测冬小麦叶面积指数 LAI, 且估测精度优于传统模型。何汝艳等^[19]联合应用连续小波变换与支持向量机方法对条锈病胁迫下冬小麦叶片全氮含量进行反演, 且具有较高的估测精度。

对作物的叶片氮含量进行的研究, 大部分是基于光谱数据, 其优点是预测精度较高, 但是成本昂贵、操作难度大, 难以推广使用。本研究通过提取图像特征结合机器学习对苹果树叶片进行氮素营养预测, 具有成本低、可操作性高、可实时监测的优点。本文以不同时期的田间试验为基础, 通过主成分分析提取不同时期的关键影响因子, 对比基于支持向量机、反向传播(Back propagation, BP)神经网络和极限学习机(Extreme learning machine, ELM)3 种方法建立的不同生长时期苹果树叶片氮含量预测模型, 以期建立不同时期不同因子的最佳苹果树叶片氮含量预测模型, 为苹果树营养诊断、精准施肥提供信息支持。

1 材料与方法

1.1 试验区域

试验样本采集自山东省万林农场试验园, 试验

品种为 4 年树龄的烟台红富士。试验园分为 A、B 两个区域, 区域 A 采集的样本作为测试集, 区域 B 采集的样本作为验证集。试验园位于山东省泰安市宁阳县葛石镇, 海拔 62.2 m, 位于东经 116°49′、北纬 35°45′, 属于温带大陆性半湿润季风气候, 年平均气温 13℃, 年平均降水量 697 mm。

1.2 数据采集

1.2.1 样本采集与叶片图像获取

分别于 4 月 3 日(开花期)、4 月 12 日(开花期)、5 月 18 日(幼果期)、6 月 3 日(幼果期)、6 月 22 日(果实膨大期)、7 月 9 日(果实膨大期)在区域 A 试验园随机选取 50 株苹果树作为试验本体, 标号 1~50, 并在每株试验本体的树冠外围新梢中间部位 4 个方位各摘取 5 片充分展开、无损伤、无病虫害的健康叶片作为试验样本^[20-21]。在 4 个不同方位选取平整的、便于色彩采集的 10 片叶片放置在光学标定板上, 将佳能 EOS 80D 型相机(焦距 35 mm)安装于深圳市优捷虎电子科技有限公司生产的卓美 z-888 型三脚架上, 利用水平仪调节相机, 使镜头与标定板垂直进行图像采集, 图像尺寸为 6 000 像素 × 4 000 像素。同一植株在不同时期采集的叶片图像如图 1 所示。



图 1 不同生长期苹果叶片原始图像

Fig. 1 Original images of apple leaves at flowering, young fruit and fruit expansion stages

1.2.2 样本处理与氮素营养测定

采用凯氏定氮法对苹果树叶片全氮含量进行测定。将带回实验室的叶片放在自来水下冲洗干净, 用 95:5 的盐酸溶液清洗 1~2 min, 再用无离子水冲洗 2 次, 放入 105℃ 高温烘箱杀青 30 min, 再调至 80℃ 持续干燥至样品质量恒定, 使用研钵将干燥样品研磨至粉末状, 称取 0.1 g 干样, 用 H₂SO₄-H₂O₂ 消解, 冷却后采用海能公司生产的 K9860 型全自动凯氏定氮仪测定苹果树叶片的全氮含量, 测量 3 次取平均值作为最终氮含量。

1.3 数据预处理

首先获取苹果树叶片原始图像。其次对单片叶片分割提取, 将图像背景标定板去掉后图像如图 2

所示,分割后得到单片叶片。

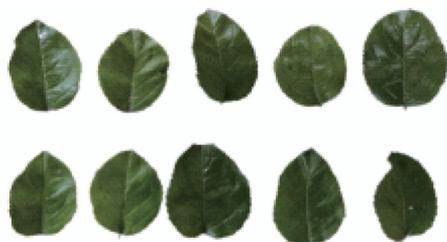


图2 去掉标定板后图像

Fig.2 Images after removing calibration board

对苹果树叶片根据多尺度 Retinex 理论 (Multi scale Retinex, MSR) 进行去光照处理, MSR 的优点是不仅可以增强图像对比度还可以对图像在动态范围内进行压缩, 并且, 在一定条件下, MSR 可以克服单尺度 Retinex 理论 (Single scale Retinex, SSR) 的缺点, 保持图像色彩恒常性的同时, 实现图像全局、部分动态范围压缩以及图像色彩的增强, MSR 计算公式为

$$r(x, y) = \sum_{k=1}^K W_k (\lg S(x, y) - \lg (F_k * S(x, y))) \quad (1)$$

式中 $r(x, y)$ ——输出图像

W_k ——第 k 个尺度的加权系数

$S(x, y)$ ——原始图像

F_k ——环绕函数

K ——高斯中心环绕函数的个数, 当 $K = 1$ 时, 只有一个中心环绕函数, MSR 理论退化为 SSR 理论

通常情况下, 为保证 MSR 算法同时包含 SSR 算法低、中、高 3 个尺度的优点, K 取值为 3, 并且 $W_1 = W_2 = W_3 = 1/3$ 。图 3 为图像处理前后对比。

最后, 在 RGB 空间下提取苹果树叶片各单色分



图3 图像处理前后对比

Fig.3 Comparison before and after image processing

量, 开花期、幼果期和果实膨大期苹果树叶片在 RGB 空间下各单色分量的分布范围。开花期, R 、 G 、 B 分量的最大值分别为 7.882、62.410、15.539, 最小值分别为 5.041、50.236、6.470, 平均值分别为 6.139、54.684、10.886; 在幼果期, R 、 G 、 B 分量的最大值分别为 7.867、53.602、23.589, 最小值分别为 2.764、43.132、13.694, 平均值分别为 5.231、47.449、18.838; 果实膨大期, R 、 G 、 B 分量的最大值分别为 9.143、49.821、28.361, 最小值分别为 2.333、40.828、22.515, 平均值分别为 5.552、45.118、25.447。

为避免图像的单色分量参数对氮素表征能力弱的问题, 进一步提升基于图像处理技术在 RGB 空间下对苹果树叶片氮含量监测的准确性, 引入 $R+G+B$ 、 $R-B$ 、 $R-G$ 、 $G-B$ 、 $R-G-B$ 、 $R/(R+G+B)$ 、 $G/(R+G+B)$ 、 $B/(R+B+G)$ 、 G/R 、 $R/(R+G-B)$ 、 $G/(R+G-B)$ 、 $B/(R+G-B)$ 、 $(R-B)/(R+G+B)$ 、 $(R-G)/(R+G+B)$ 共 14 个单色分量色彩组合参数, 共计 17 种图像色彩特征。因篇幅有限, 在开花期采集的部分原始数据, 如表 1 所示。

表 1 开花期采集的部分原始数据

Tab.1 Part of original data collected at flowering stage

序号	B	G	R	$R-B$	$R-G$	$R-G-B$	$B/(R+G+B)$	$(R-B)/(R+G+B)$	$(R-G)/(R+G+B)$	氮质量比/ ($g \cdot kg^{-1}$)
1	13.813	52.413	6.320	-7.493	-46.093	-59.906	0.308	-0.103	-0.635	9.281
2	15.266	53.438	6.152	-9.114	-47.281	-62.551	0.344	-0.122	-0.632	8.761
3	13.138	51.229	7.440	-5.697	-43.789	-56.927	0.289	-0.079	-0.610	8.215
4	12.177	52.758	5.041	-7.136	-47.717	-59.894	0.267	-0.102	-0.682	11.947
5	12.190	53.689	5.842	-6.348	-47.846	-60.037	0.258	-0.089	-0.667	13.943
6	9.342	55.439	6.482	-2.860	-48.956	-58.298	0.178	-0.040	-0.687	9.680
7	10.527	56.378	5.141	-5.386	-51.237	-61.764	0.206	-0.075	-0.711	12.385
8	9.701	55.499	6.081	-3.620	-49.418	-59.119	0.187	-0.051	-0.693	11.869
9	10.959	51.584	5.391	-5.568	-46.193	-57.152	0.238	-0.082	-0.680	13.159
10	9.3868	53.225	6.227	-3.160	-46.999	-56.385	0.187	-0.046	-0.683	11.493
11	11.940	54.181	5.801	-6.139	-48.380	-60.320	0.249	-0.085	-0.673	16.313
12	11.203	53.730	6.124	-5.078	-47.606	-58.808	0.230	-0.071	-0.670	11.682
13	7.140	51.352	5.340	-1.740	-45.952	-53.092	0.144	-0.027	-0.719	13.754
14	9.172	57.940	5.752	-3.420	-52.188	-61.360	0.168	-0.047	-0.716	12.670

2 预测模型构建

2.1 主成分分析法

主成分分析(Principal component analysis, PCA)是一种重要的数据统计方法,其能够从众多原始变量中找出几个综合性强的变量来替代原始变量,新的综合变量之间不仅互不相关,而且保留了原始变量的绝大部分信息,最终实现降维目的,使问题得到最佳综合^[22-24]。其计算步骤为:①预处理后的原始数据构造标准化矩阵。②对标准化矩阵计算相关系数矩阵 $R = (r_{ij})_{m \times n}$ 。③计算相关系数矩阵 R 的特征方程得到特征值和特征向量。④计算贡献率 τ_i 与累计贡献率 η_i 。累计贡献率越大说明数据信息损失越小。⑤计算主成分载荷矩阵以及主成分得分。其中载荷矩阵表示原始数据协方差矩阵的特征值。

$$l_{ij} = \sqrt{\lambda_i} e_{ij} \quad (i, j = 1, 2, \dots, p) \quad (2)$$

式中 l_{ij} ——主成分载荷矩阵

λ_i ——原始数据协方差矩阵的特征值

e_{ij} ——单位特征向量

最后,将原始样本数据代入主成分的表达式中即可计算出主成分得分。

2.2 对比模型

2.2.1 支持向量机

支持向量机是机器学习领域最重要的方法之一,是借助于统计与优化方法解决机器学习问题的强有力工具。针对苹果树叶片在 RGB 空间下各单色分量及各单色分量多种色彩组合参数检测过程的随机波动性,采用在小样本、非线性问题的拟合上有一定优势的 SVM 回归模型来构建苹果树叶片在 RGB 空间下 17 种色彩特征与苹果树叶片氮含量的预测模型。苹果树叶片在 RGB 空间下各单色分量及各单色分量多种色彩组合参数检测过程中,利用每组试验的输入变量 x_i 和输出量(苹果树叶片氮含量) s_i 构建样本空间 $\{(x_i, s_i), i = 1, 2, \dots, n\}$ 。 s_i 可表示为非线性函数模型

$$f_s(x) = \omega^T \varphi(x) + b \quad (3)$$

式中 $f_s(x)$ ——非线性函数模型

ω ——自变函数系数

$\varphi(x)$ ——输入空间 x 的非线性映射

b ——偏置

为最大可能减少训练误差的经验风险, ω 和 b 通过 $R(f)$ 模型进行评估

$$R(f) = \min \left(\frac{1}{2} \|\omega\|^2 + \frac{C}{n} \sum_{j=1}^n L(y_j, f(x_j)) \right) \quad (4)$$

式中 $R(f)$ ——考虑最少错分样本的广义最优分

类面函数

C ——惩罚因子 L ——损失函数

y_i ——真实值 $f(x_i)$ ——预测值

n ——样本个数

随后引入不敏感损失函数 ε 对结构风险最小化进行评估, ξ_i, ξ_i^* 定义为松弛变量,优化目标可转换为

$$\begin{cases} \min \left(\frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^n (\xi_i + \xi_i^*) \right) \\ \text{s. t.} \begin{cases} y_i - \omega^T \varphi(x) - b \leq \xi_i + \varepsilon \\ \omega^T \varphi(x) + b - y_i \leq \xi_i^* + \varepsilon \\ \xi_i \geq 0 \\ \xi_i^* \geq 0 \end{cases} \end{cases} \quad (5)$$

引入 Lagrange 方程

$$L = \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^n (\xi_i + \xi_i^*) - \sum_{i=1}^n \alpha_i (\varepsilon + \xi_i^* - y_i + \omega^T \varphi(x_i) + b) - \sum_{i=1}^n \alpha_i^* (\varepsilon + \xi_i - y_i + \omega^T \varphi(x_i) + b) - \sum_{i=1}^n (\eta_i \xi_i + \eta_i^* \xi_i^*) \quad (6)$$

式中 $\alpha_i, \alpha_i^*, \xi_i, \xi_i^*$ ——Lagrange 乘子

分别对 $\omega, b, \xi_i, \xi_i^*$ 使用 Lagrange 方程求偏导,并代入式(6),转化为对偶优化问题,求解获得支持向量回归机的回归函数。

$$f_s(x) = \sum_{i=1}^n (\alpha_i - \alpha_i^*) K \langle x_i, x_j \rangle + b \quad (j = 1, 2, \dots, n) \quad (7)$$

其中 $K \langle x_i, x_j \rangle = \varphi(x_i) \varphi(x_j)$

式中 $K \langle x_i, x_j \rangle$ ——核函数

2.2.2 BP 神经网络

BP 神经网络在复杂的非线性系统中具有较高的建模能力,并对数据具有良好的拟合能力,在预测方面应用广泛。研究表明,这种简单的网络结构可以逼近任意非线性函数,是一种典型函数映射关系。

2.2.3 极限学习机

极限学习机(ELM)是一种简单易用、有效的单隐层前馈神经网络。与传统的前馈神经网络训练速度慢、容易陷入极小值点、学习率选择敏感等缺点相比,ELM 算法随机产生输入层与隐含层的连接权值 w 及隐含层神经元的阈值 t ,且在训练过程中无需调整。只要设置隐含层神经元的个数,便可以获得唯一的最优解,具有学习速度快、泛化能力强等优点。

2.3 基于组合色彩特征的苹果树叶片氮含量预测模型

基于上述算法,为了提高预测模型精度,本文将

多种算法相结合,构建了基于组合色彩特征的苹果树叶氮含量预测模型。其预测流程如图4所示,预测步骤如下:

(1)调研国内外数字图像处理与机器学习技术在农业方面的应用,尤其是对作物氮素营养预测方面的研究。

(2)采集开花期、幼果期、果实膨大期3个不同尺度苹果树叶片样本和图像。

(3)利用化学方法对叶片进行氮素营养测定,对叶片图像进行中值滤波、去光照等预处理,在RGB空间下提取苹果树叶片17种色彩特征。

(4)利用主成分分析对苹果树叶片图形色彩特征进行优化,提取不同尺度关键影响因子。

(5)建立PCA-SVM、PCA-BP、PCA-ELM预测模型,对比筛选出不同尺度最佳预测模型;调整优化最佳模型参数,构建基于组合色彩特征的苹果树叶氮含量预测模型;通过田间试验验证苹果树叶氮含量预测模型性能,实现对不同生长时期苹果树叶氮含量的精准预测。

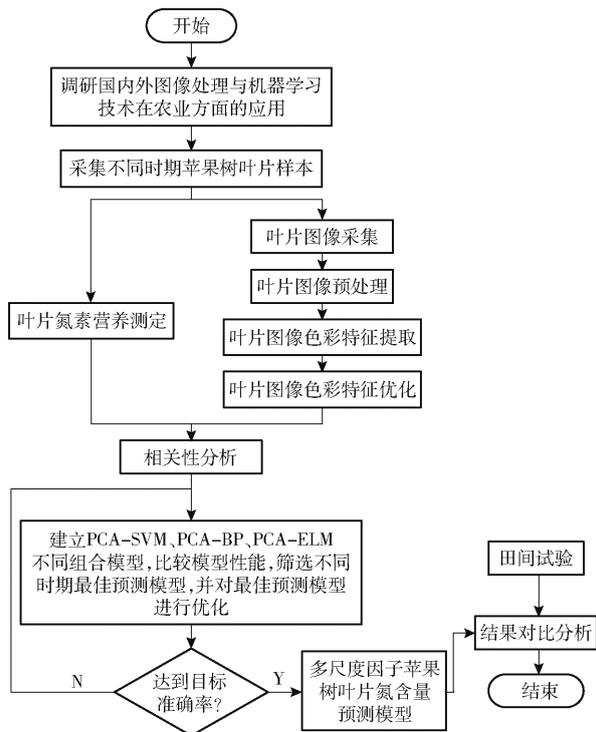


图4 基于组合色彩特征的苹果树叶氮含量预测流程图

Fig. 4 Prediction flow chart of apple leaf nitrogen content based on multi-scale factors

2.4 模型评价指标

使用预处理后的数据作为模型的输入,对基于组合色彩特征的预测模型进行训练和测试。本文选用的预测模型性能评价指标有平均绝对误差 (Mean absolute error, MAE)、均方根误差 (Root mean square error, RMSE) 和平均绝对百分误差 (Mean absolute

percentage error, MAPE)。

3 试验结果与分析

3.1 数据集

以山东省泰安市宁阳县万林农场试验园区域A中苹果树叶氮含量为研究对象。采用已经预处理过的试验数据,主要包括RGB空间中 R 、 G 、 B 和 $R+G+B$ 、 $R-B$ 、 $R-G$ 、 $G-B$ 、 $R-G-B$ 、 $R/(R+G+B)$ 、 $G/(R+G+B)$ 、 $B/(R+B+G)$ 、 G/R 、 $R/(R+G-B)$ 、 $G/(R+G-B)$ 、 $B/(R+G-B)$ 、 $(R-B)/(R+G+B)$ 、 $(R-G)/(R+G+B)$ 14个单色分量色彩组合参数,最终实现开花期、幼果期和果实膨大期苹果树叶氮含量的精准预测。

3.2 平台和环境

试验所使用计算机的配置如下:处理器为Intel (R) Xeon (R) Silver 4210R CPU @ 2.40 Hz;机带内存为64 GB;操作系统为Windows 10 (64位);图像分割、图像去光照处理与获取苹果树叶RGB空间下各单色分量及单色分量多种色彩组合参数由Matlab实现;SVM、BP神经网络和ELM回归模型均在Matlab中完成。

3.3 试验结果分析

3.3.1 模型分析

由于图像多种色彩特征之间存在一定的相关性,直接作为模型的输入会使数据量过于庞大,使用主成分分析法对所获取的数据降维和分析,得到不相关的新变量,最终提取出多尺度苹果树叶氮含量的关键影响因子。本文利用SPSS软件实现对苹果树叶氮含量关键影响因子的筛选,不同生长时期的方差及主成分贡献率如表2所示。

选择“特征值”大于1作为提取主成分的原则,由表2可知,开花期、幼果期、果实膨大期中前3个因子的特征值均大于1,因此3个生长时期均选择前3个因子代替原变量。采用具有Kaiser标准化的正交旋转法得到各个颜色因子对不同主因子的因子载荷,得到的不同时期成分矩阵如图5所示。图中图像色彩特征1~17分别表示成分 B 、 G 、 R 、 $R-B$ 、 $R-G$ 、 $G-B$ 、 $R+G+B$ 、 $R-G-B$ 、 $R/(R+G+B)$ 、 $G/(R+G+B)$ 、 $B/(R+G+B)$ 、 G/R 、 $R/(R+G-B)$ 、 $G/(R+G-B)$ 、 $B/(R+G-B)$ 、 $(R-B)/(R+G+B)$ 、 $(R-G)/(R+G+B)$ 。

由图5可以看出,开花期,对成分1贡献最大的有 B 、 $R-B$ 、 $G-B$ 、 $G/(R+G+B)$ 、 $B/(R+G-B)$ 、 $G/(R+G-B)$ 、 $B/(R+G-B)$ 、 $(R-B)/(R+G+B)$ 8种色彩特征,对成分2贡献较大的是 $R-G$ 、 $R/(R+G+B)$ 和 G/R 、 $R+G+B$ 对成分3贡献较

表 2 开花期、幼果期、果实膨大期方差及主成分贡献率

Tab. 2 Variance and principal component contribution rate in flowering, young fruit and fruit expansion stages

时期	成分	初始特征值			提取载荷平方和载入			旋转载荷平方和载入		
		特征值	方差百分比/%	累计贡献率/%	特征值	方差百分比/%	累计贡献率/%	特征值	方差百分比/%	累计贡献率/%
开花期	1	8.813	51.844	51.844	8.813	51.844	51.844	8.765	51.507	51.507
	2	5.767	33.925	85.769	5.767	33.925	85.769	4.500	26.468	77.975
	3	2.390	14.056	99.825	2.930	14.056	99.825	3.714	21.850	99.825
	4	0.013	0.079	99.904						
	∴	∴	∴	∴						
	17	-2.075×10^{-15}	-1.220×10^{-14}	100						
幼果期	1	8.476	49.861	49.861	8.476	49.861	49.861	7.719	45.409	45.409
	2	5.764	33.905	83.766	5.764	33.905	83.766	5.476	32.213	77.622
	3	2.065	12.147	95.913	2.065	12.147	95.913	3.110	18.292	95.914
	4	0.480	2.823	98.736						
	∴	∴	∴	∴						
	17	-5.481×10^{-16}	-3.224×10^{-15}	100						
果实膨大期	1	8.394	49.374	49.374	8.394	49.374	49.374	7.869	46.287	46.287
	2	6.371	37.477	86.851	6.371	37.477	86.851	6.167	36.276	82.563
	3	2.008	12.280	99.131	2.008	12.280	99.131	2.817	16.568	99.131
	4	0.113	0.667	99.798						
	∴	∴	∴	∴						
	17	-1.747×10^{-15}	-1.028×10^{-14}	100						

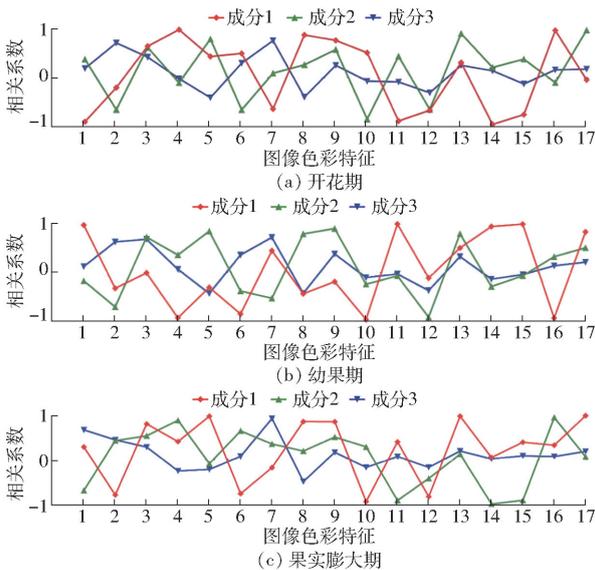


图 5 开花期、幼果期、果实膨大期成分矩阵

Fig. 5 Composition matrix in flowering, young fruit and expanding fruit stages

大;幼果期,对成分 1 贡献较大的有 $R - B$ 、 $G/(R + G + B)$ 、 $(R - B)/(R + G + B)$,对成分 2 贡献较大的有 $R/(R + G - B)$ 、 $(R - G)/(R + G + B)$,对成分 3 贡献较大的有 G 、 $R + G + B$;果实膨大期,对成分 1 贡献最大的有 $R - G$ 、 $G/(R + G + B)$ 、 $R/(R + G - B)$ 、 $(R - G)/(R + G + B)$,对成分 2 贡献较大的有 $B/(R + G + B)$ 、 $G/(R + G - B)$ 、 $B/(R + G - B)$ 、 $(R - B)/(R + G + B)$ 、 $R + G + B$ 对成分 3 贡献较大。

通过主成分分析方法对不同尺度苹果树叶片的

色彩特征进行降维处理,提取出前 3 个主成分因子,并建立各主成分的线性表达式为

$$F_{11} = 0.11u_1 - 0.038u_2 - 0.031u_3 + \dots + 0.079u_{17} \quad (8)$$

$$F_{12} = 0.023u_1 + 0.006u_2 + 0.276u_3 + \dots + 0.135u_{17} \quad (9)$$

$$F_{13} = 0.056u_1 + 0.283u_2 + 0.144u_3 + \dots + 0.012u_{17} \quad (10)$$

$$F_{21} = 0.115v_1 - 0.073v_2 - 0.033v_3 + \dots + 0.080v_{17} \quad (11)$$

$$F_{22} = 0.051v_1 + 0.057v_2 + 0.210v_3 + \dots + 0.172v_{17} \quad (12)$$

$$F_{23} = 0.093v_1 + 0.355v_2 + 0.123v_3 + \dots + 0.021v_{17} \quad (13)$$

$$F_{31} = 0.077w_1 - 0.018w_2 + 0.144w_3 + \dots + 0.132w_{17} \quad (14)$$

$$F_{32} = 0.193w_1 - 0.027w_2 - 0.023w_3 + \dots + 0.040w_{17} \quad (15)$$

$$F_{33} = 0.266w_1 + 0.238w_2 + 0.112w_3 + \dots + 0.049w_{17} \quad (16)$$

式中 F_{11} 、 F_{12} 、 F_{13} 、 F_{21} 、 F_{22} 、 F_{23} 、 F_{31} 、 F_{32} 、 F_{33} 表示不同尺度前 3 个主成分因子, u_1 、 u_2 、 \dots 、 u_{17} 、 v_1 、 v_2 、 \dots 、 v_{17} 、 w_1 、 w_2 、 \dots 、 w_{17} 表示不同尺度苹果树叶片的 17 种色彩特征。

因此本文根据尺度不同选取不同的关键影响因

子,开花期选用的关键影响因子为 B 、 $R - B$ 、 $G - B$ 、 $G/(R + G + B)$ 、 $B/(R + G - B)$ 、 $G/(R + G - B)$ 、 $B/(R + G - B)$ 、 $(R - B)/(R + G + B)$ 、 $R - G$ 、 $R/(R + G + B)$ 、 G/R 、 $R + G + B$; 幼果期选用的关键影响因子为 $R - B$ 、 $G/(R + G + B)$ 、 $(R - B)/(R + G + B)$ 、 $R/(R + G - B)$ 、 $(R - G)/(R + G + B)$ 、 G 、 $R + G + B$; 果实膨大期选用的关键影响因子为 $R - G$ 、 $G/(R + G + B)$ 、 $R/(R + G - B)$ 、 $(R - G)/(R + G + B)$ 、 $B/(R + G + B)$ 、 $G/(R + G - B)$ 、 $B/(R + G - B)$ 、 $(R - B)/(R + G + B)$ 、 $R + G + B$ 。选取不同的关键影响因子可降低数据维度和复杂度,并以此构建多尺度苹果树叶片氮含量预测模型的输入因子,输出参数为苹果树叶片氮含量。

3.3.2 模型选择

为了得到更加精准的预测效果,分别对 PCA - SVM 模型、PCA - BP 模型、PCA - ELM 模型进行对比,均以相同的样本进行训练。预测误差曲线和预测结果精度对比如图 6 和表 3 所示。

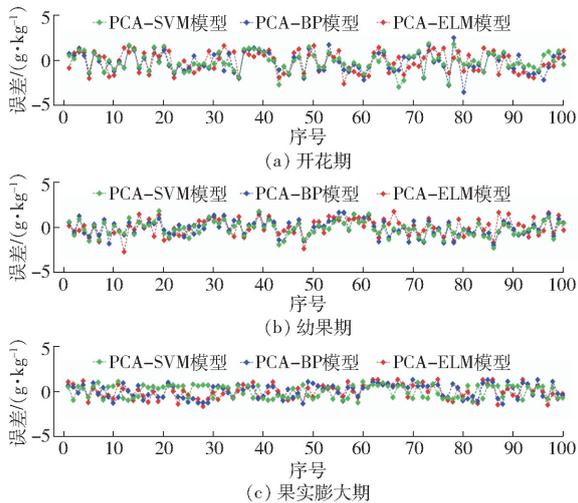


图 6 不同生长时期不同模型的预测误差曲线

Fig. 6 Prediction error curves of different models in different growth periods

由图 6 可以看出,开花期、幼果期和果实膨大期 PCA - SVM 模型的误差曲线在零值上下波动最小,其次是 PCA - BP 模型,3 个时期内误差曲线波动最大的是 PCA - ELM 模型。

由表 3 可以看出,开花期,PCA - SVM 模型的 MAE、MAPE 和 RMSE 均低于 PCA - ELM 和 PCA - BP 模型,PCA - SVM 模型的 MAE、MAPE 和 RMSE 分别为 1.072、0.099、1.035 g/kg,PCA - SVM 模型预测精度高于 PCA - BP、PCA - ELM 两种模型;幼果期,PCA - SVM 模型的 MAE、MAPE 和 RMSE 均低于 PCA - ELM 和 PCA - BP 模型,PCA - SVM 模型的 MAE、MAPE 和 RMSE 分别为 0.722、0.065、0.850 g/kg,PCA - SVM 模型预测精度高于 PCA -

表 3 不同生长时期各模型预测结果精度分析

Tab. 3 Accuracy analysis of prediction results of each model in different growth periods g/kg

时期	模型	MAE	MAPE	RMSE
开花期	PCA - BP	1.114	0.102	1.056
	PCA - ELM	1.745	0.157	1.321
	PCA - SVM	1.072	0.099	1.035
幼果期	PCA - BP	0.802	0.072	0.896
	PCA - ELM	1.147	0.102	1.071
	PCA - SVM	0.722	0.065	0.850
果实膨大期	PCA - BP	0.689	0.086	0.835
	PCA - ELM	1.311	0.157	1.145
	PCA - SVM	0.680	0.080	0.833

BP 和 PCA - ELM 模型;果实膨大期,PCA - SVM 模型的 MAE、MAPE 和 RMSE 也低于 PCA - ELM 和 PCA - BP 模型,PCA - SVM 模型的 MAE、MAPE 和 RMSE 分别为 0.680、0.080、0.833 g/kg,PCA - SVM 模型预测精度高于 PCA - BP 和 PCA - ELM 模型。对比结果表明,PCA - SVM 模型在开花期、幼果期、果实膨大期的预测精度均高于 PCA - BP、PCA - ELM 模型。因此本文选择 PCA - SVM 模型作为组合色彩特征苹果树叶片氮含量的预测模型。

3.3.3 模型优化

影响遗传算法^[25-26](Genetic algorithm, GA)性能和行为的关键因素为交叉概率(Crossover probability, P_c)和变异概率(Mutation probability, P_m),但要通过反复试验来确定这 2 个参数值,过程较为繁琐。自适应遗传算法(Adaptive genetic algorithm, AGA)^[27-28]可使 P_c 和 P_m 随适应度自动变更,即当种群中各个体的适应度趋于一定值或趋于局部最优时,增加 P_c 和 P_m ;相反,当适应度较发散时,减小 P_c 和 P_m 。对于种群适应度小于平均适应度的个体,采用较高的 P_c 和 P_m 淘汰该个体,而大于种群平均适应度的个体,采用较低的 P_c 和 P_m 使该个体得到保护进入下一代。但该方法较适于处于进化后期的群体,对初期进化不利。为了避免 P_c 和 P_m 取零值,防止其处于一种近似停滞不前的局面,保证每代的优良个体不被破坏,采用精英选择策略,使其直接复制到下一代中。优化后 P_c 和 P_m 计算公式分别为

$$P_c = \begin{cases} P_{c1} - \frac{P_{c1} - P_{c2} - (f' - f_{avg})}{f_{max} - f_{avg}} & (f' \geq f_{avg}) \\ P_{c2} & (f' < f_{avg}) \end{cases} \quad (17)$$

$$P_m = \begin{cases} P_{m1} - \frac{P_{m1} - P_{m2} - (f_{max} - f)}{f_{max} - f_{avg}} & (f' \geq f_{avg}) \\ P_{m2} & (f' < f_{avg}) \end{cases} \quad (18)$$

式中 f_{\max} ——群体中最大适应度
 f_{avg} ——每代群体的平均适应度
 f' ——交叉个体中较大的适应度
 f ——变异个体的适应度
 $P_{c1}、P_{c2}$ ——交叉概率,取 0.9、0.6
 $P_{m1}、P_{m2}$ ——变异概率,取 0.1、0.001
 惩罚因子 C 的范围为 $[0.1, 100]$, RBF 核函数宽度 σ 的范围为 $[0.01, 10]$, 交叉概率 P_c 的范围为 $[0.4, 0.9]$, 变异概率 P_m 的范围为 $[0.001, 0.1]$ 。支持向量机参数最优值如表 4 所示。

表 4 支持向量机参数最优值

Tab.4 Optimal parameters of support vector machine

参数		开花期	幼果期	果实膨大期
C	优化前	10.000	10.000	10.000
	优化后	7.000	0.150	80.000
σ	优化前	0.200	0.200	0.200
	优化后	0.056	0.100	0.300

把优化后的 C 和 σ 应用到组合色彩特征苹果树叶片氮含量预测模型中,与优化前的预测模型相比较,结果如图 7 所示。

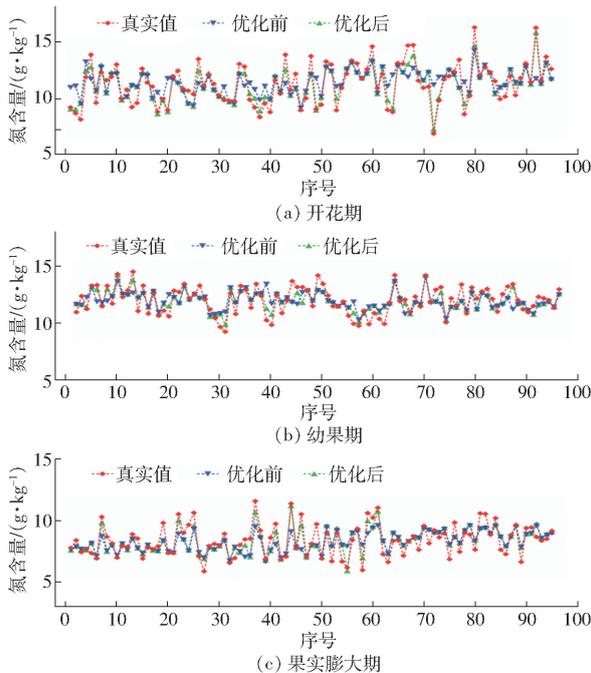


图 7 不同生长期优化前后预测模型氮含量预测结果对比
 Fig.7 Comparisons of nitrogen content predicted by prediction models before and after optimization in different growth stages

从图 7 可以看出,在开花期、幼果期和果实膨大期 3 个尺度中优化后的模型所得到的预测曲线均与实际曲线更加接近,说明优化后 PCA-SVM 苹果树叶片氮含量预测模型的拟合效果优于优化前。表 5 为优化后的 PCA-SVM 模型在开花期、幼果期和果实膨大期试验得到的 MAE、MAPE 和 RMSE。

表 5 不同生长期优化后 PCA-SVM 预测模型的精度分析

Tab.5 Accuracy analysis of PCA-SVM prediction model optimized in different growth periods

生育期	MAE	MAPE	RMSE
开花期	0.640	0.057	0.800
幼果期	0.558	0.050	0.747
果实膨大期	0.544	0.064	0.737

从表 5 可以看出,相对于开花期和幼果期,在果实膨大期取得的预测性能最好。此外,优化后的 PCA-SVM 模型的整体预测精度在不同时期均有所提高,开花期,与优化前相比 MAE、MAPE 和 RMSE 分别降低了 40.3%、42%、22.7%;幼果期,与优化前相比 MAE、MAPE 和 RMSE 分别降低了 22.7%、23.1%、12.1%;果实膨大期,与优化前相比 MAE、MAPE 和 RMSE 分别降低 20%、20%、11.5%,可以看出开花期组合色彩特征预测模型的预测精度提高最为明显,其次是幼果期和果实膨大期。试验结果表明,PCA-SVM 模型预测精度较高,泛化能力强,可以较好地拟合不同尺度 RGB 空间下各单色分量和各单色分量多种色彩组合参数与苹果树叶片氮含量之间的复杂关系,可以为苹果园精准施肥提供决策依据。

3.3.4 试验验证

试验验证样本于 2020 年 4—7 月在山东省泰安市宁阳县万林农场试验园区域 B 采集。在开花期、幼果期和果实膨大期随机抽样 50 株果树作为验证样本^[29],通过凯氏定氮法获得苹果树叶片的实际氮含量,利用组合色彩特征模型预测苹果树叶片的氮含量。通过相对误差评定苹果树叶片氮含量预测模型的准确性,开花期、幼果期和果实膨大期的苹果树叶片氮含量及相对误差分布如图 8 所示。

试验结果表明,开花期,苹果树叶片氮含量预测值与实测值的相对误差在 -6%~9.84% 之间;幼果期,苹果树叶片氮含量预测值与实测值的相对误差在 -9%~9.51% 之间;果实膨大期,苹果树叶片氮含量预测值与实测值的相对误差在 -10%~9.86% 之间。说明开花期和幼果期基于组合色彩特征模型预测苹果树叶片氮含量相对误差较小,总体来说,模型预测值与试验结果基本符合,不同时期苹果树叶片氮含量实测值与预测值相对误差在 10% 以内,验证了基于组合色彩特征的苹果树叶片氮含量预测模型的准确性。通过机器视觉与图像处理的方式获取苹果树叶片在 RGB 空间下各单色分量及各单色分量多种色彩组合参数,预测不同生长期苹果树叶片氮含量,与传统方法相比,本文

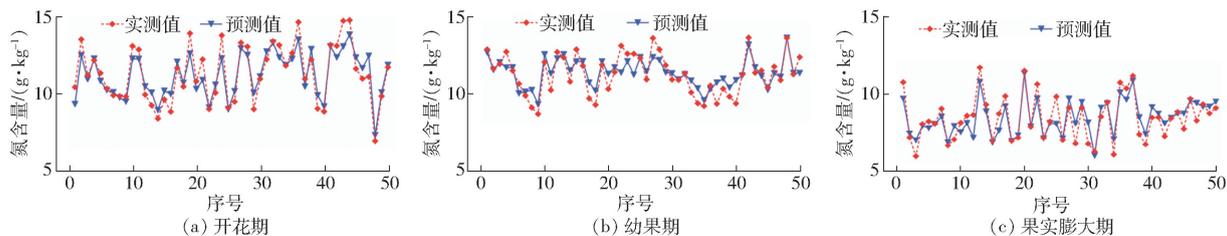


图8 不同生长时期苹果树叶片氮含量分布

Fig. 8 Distribution of nitrogen content in apple leaves at different growth stages

方法具有无损伤、效率高、成本低、操作简单的优势。准确地预测苹果树叶片氮含量也可为果园精准化管理,合理施肥、定时定量施肥提供理论依据。

4 结论

(1)采集不同生长时期的苹果树叶片并通过图像处理获取RGB空间下各单色分量与14种色彩组合参数,采用PCA筛选出不同尺度的苹果树叶片氮含量不同的关键影响因子,开花期选用的关键影响因子为 B 、 $R-B$ 、 $G-B$ 、 $G/(R+G+B)$ 、 $B/(R+G-B)$ 、 $G/(R+G-B)$ 、 $B/(R+G-B)$ 、 $(R-B)/(R+G+B)$ 、 $R-G$ 、 $R/(R+G+B)$ 、 G/R 、 $R+G+B$,幼果期选用的关键影响因子为 $R-B$ 、 $G/(R+G+B)$ 、 $(R-B)/(R+G+B)$ 、 $R/(R+G-B)$ 、 $(R-G)/(R+G+B)$ 、 G 、 $R+G+B$,果实膨大期选用的关键影响因子为 $R-G$ 、 $G/(R+G+B)$ 、 $R/(R+G-B)$ 、 $(R-G)/(R+G+B)$ 、 $B/(R+G+B)$ 、 $G/(R+G-B)$ 、 $B/(R+G-B)$ 、 $(R-B)/(R+G+B)$ 、 $R+G+B$,以此构建苹果

树叶片氮含量预测模型的输入样本,输出参数为苹果树叶片氮含量,实现了模型输入的降维,消除了变量之间的冗余性和相关性。

(2)提出PCA-SVM、PCA-ELM、PCA-BP共3种预测模型,在不同生长时期,由MAE、RMSE和MAPE作为模型评价指标确定多尺度预测模型。

(3)利用自适应遗传算法对组合色彩特征模型中惩罚参数 C 和RBF核函数宽度 σ 进行优化,优化后PCA-SVM预测模型在开花期、幼果期和果实膨大期的平均绝对误差分别为0.640、0.558、0.544 g/kg,平均绝对百分误差分别为0.057、0.050、0.064 g/kg,均方根误差分别为0.800、0.747、0.737 g/kg,优于优化前预测模型,能够拟合苹果树叶片氮含量与RGB空间下各单色分量及各单色分量多种色彩组合参数之间的非线性关系,解决了传统预测方法预测精度低、鲁棒性差等问题,该模型具有良好的预测性能和泛化能力,可以为果园精准施肥管理、提升果品品质、避免资源浪费和环境污染提供理论依据。

参考文献

- [1] MA W, ABDULAI A. Does cooperative membership improve household welfare? Evidence from apple farmers in China[J]. Food Policy, 2016(58): 94-102.
- [2] 王璇,刘军弟,邵砾群,等. 我国苹果产业年度发展状况及其趋势与建议[J]. 中国果树,2008,4(3):101-104,108. WANG Xuan, LIU Jundi, SHAO Liqun, et al. Annual development status, trend and suggestions of apple industry in China [J]. China Fruits, 2008, 4(3): 101-104, 108. (in Chinese)
- [3] ANDREWS M, RAVEN J A, LEA P J. Do plants need nitrate? The mechanisms by which nitrogen form affects plants[J]. Annals of Applied Biology, 2013, 163(2):174-199.
- [4] 李文庆,张民,束怀瑞. 氮素在果树上的生理作用[J]. 山东农业大学学报(自然科学版),2002,4(1):96-100. LI Wenqing, ZHANG Min, SHU Huairui. Physiological effects of nitrogen on fruit trees[J]. Journal of Shandong Agricultural University (Natural Science Edition), 2002, 4(1): 96-100. (in Chinese)
- [5] 束怀瑞. 果树栽培生理学[M]. 北京:农业出版社,1993.
- [6] 顾曼如,张若杼,束怀瑞,等. 苹果氮素营养研究初报——植株中氮素营养的年周期变化特性[J]. 园艺学报,1981(4): 21-28. GU Manru, ZHANG Ruozhu, SHU Huairui, et al. Preliminary report on nitrogen nutrition of apples—annual variation characteristics of nitrogen nutrition in plants[J]. Acta Horticulturae Sinica, 1981(4): 21-28. (in Chinese)
- [7] 姜远茂,张宏彦,张福锁,等. 北方落叶果树养分资源综合管理理论与实践[M]. 北京:中国农业大学出版社,2007.
- [8] 范广元,李壮,李敏,等. 辽西红富士苹果叶片矿质营养分析及诊断研究[J]. 中国果树,2014,4(3):20-25. FAN Guangyuan, LI Zhuang, LI Min, et al. Analysis and diagnosis of mineral nutrition in leaves of Red Fuji apple in western Liaoning Province[J]. China Fruits, 2014, 4(3): 20-25. (in Chinese)
- [9] 中国农业百科全书总编辑委员会,畜牧业卷编辑委员会. 中国农业百科全书农业化学卷[M]. 北京:农业出版社,1996.
- [10] 李艳大,舒时富,陈立才,等. 基于便携式作物生长监测诊断仪的江西双季稻氮肥调控研究[J]. 农业工程学报,2019,

- 35(2):100-106.
- LI Yanda, SHU Shifu, CHEN Licai, et al. Nitrogen regulation of double cropping rice in Jiangxi Province based on portable crop growth monitoring and diagnosis instrument[J]. Transactions of the CSAE, 2019, 35(2): 100-106. (in Chinese)
- [11] 高翠玲. 图像处理技术在农业方面的应用[J]. 河南农业, 2017, 4(12):16-17.
- GAO Cuiling. Application of image processing technology in agriculture[J]. Henan Agriculture, 2017, 4(12): 16-17. (in Chinese)
- [12] 李艳大, 叶春, 曹中盛, 等. 基于作物生长监测诊断仪的双季稻叶片氮含量和氮积累量监测[J]. 应用生态学报, 2020, 31(9):3040-3050.
- LI Yanda, YE Chun, CAO Zhongsheng, et al. Monitoring of leaf nitrogen content and nitrogen accumulation in double cropping rice based on crop growth monitoring and diagnosis instrument[J]. Chinese Journal of Applied Ecology, 2020, 31(9):3040-3050. (in Chinese)
- [13] YANG J, DU L, CHENG Y J, et al. Assessing different regression algorithms for paddy rice leaf nitrogen concentration estimations from the first-derivative fluorescence spectrum[J]. Optics Express, 2020, 28(13):18728-18741.
- [14] RUBIO-DELGADO J, PÉREZ C J, VEGA-RODRÍGUEZ M A, et al. Predicting leaf nitrogen content in olive trees using hyperspectral data for precision agriculture[J]. Precision Agriculture, 2020, 4:1-21.
- [15] 程洪, DAMEROW L, BLANKE M, 等. 基于图像处理与支持向量机的树上苹果早期估产研究[J/OL]. 农业机械学报, 2015, 46(3):9-14, 22.
- CHENG Hong, DAMEROW L, BLANKE M, et al. Early yield estimation of apple trees based on image processing and support vector machines[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2015, 46(3): 9-14, 22. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?file_no=20150302&flag=1&journal_id=jcsam. DOI:10.6041/j.issn.1000-1298.2015.03.002. (in Chinese)
- [16] 王津京, 赵德安, 姬伟, 等. 采摘机器人基于支持向量机苹果识别方法[J]. 农业机械学报, 2009, 40(1):148-151, 147.
- WANG Jinjing, ZHAO Dean, JI Wei, et al. Apple recognition method based on support vector machine for picking robot[J]. Transactions of the Chinese Society for Agricultural Machinery, 2009, 40(1): 148-151, 147. (in Chinese)
- [17] 罗建军, 杨红云, 路艳, 等. 基于高光谱和参数优化支持向量机的水稻施氮水平分类研究[J]. 中国土壤与肥料, 2020, 4(5):250-257.
- LUO Jianjun, YANG Hongyun, LU Yan, et al. Classification of nitrogen application levels in rice based on hyperspectral and parameter optimized support vector machine[J]. Soil and Fertilizer Sciences in China, 2020, 4(5):250-257. (in Chinese)
- [18] 梁栋, 管青松, 黄文江, 等. 基于支持向量机回归的冬小麦叶面积指数遥感反演[J]. 农业工程学报, 2013, 29(7):117-123.
- LIANG Dong, GUAN Qingsong, HUANG Wengjiang, et al. Remote sensing inversion of leaf area index based on support vector machine regression in winter wheat[J]. Transactions of the CSAE, 2013, 29(7): 117-123. (in Chinese)
- [19] 何汝艳, 乔小军, 蒋金豹, 等. 小波法反演条锈病胁迫下冬小麦冠层叶片全氮含量[J]. 农业工程学报, 2015, 31(2):141-146.
- HE Ruyan, QIAO Xiaojun, JIANG Jinbao, et al. Inversion of total nitrogen content in winter wheat canopy leaves under stripe rust stress by wavelet transform[J]. Transactions of the CSAE, 2015, 31(2): 141-146. (in Chinese)
- [20] 安静, 姚国清, 朱希存. 苹果叶片氮素含量高光谱检测研究[J]. 国土资源遥感, 2016, 28(2):67-71.
- AN Jing, YAO Guoqing, ZHU Xicun. Study of hyperspectral detection for nitrogen content of apple leaves[J]. Remote Sensing for Land & Resources, 2016, 28(2):67-71. (in Chinese)
- [21] 杨福芹, 冯海宽, 李振海, 等. 基于可见光-近红外光谱特征参数的苹果叶片氮含量预测[J/OL]. 农业机械学报, 2017, 48(9):143-151.
- YANG Fuqin, FENG Haikuan, LI Zhenhai, et al. Prediction for nitrogen content of apple leaves using spectral features parameters from visible and near infrared lights[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2017, 48(9): 143-151. http://www.jcsam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=20170918&journal_id=jcsam. DOI:10.6041/j.issn.1000-1298.2017.09.018. (in Chinese)
- [22] 王鹏新, 冯明悦, 孙辉涛, 等. 基于主成分分析和 Copula 函数的干旱影响评估研究[J/OL]. 农业机械学报, 2016, 47(9):334-340.
- WANG Pengxin, FENG Mingyue, SUN Huitao, et al. Drought impact assessment based on principal component analysis and Copula function[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2016, 47(9): 334-340. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?file_no=20160945&flag=1&journal_id=jcsam. DOI:10.6041/j.issn.1000-1298.2016.09.045. (in Chinese)
- [23] 李顺峰, 张丽华, 刘兴华, 等. 基于主成分分析的苹果霉心病近红外漫反射光谱判别[J]. 农业机械学报, 2011, 42(10):158-161.
- LI Shunfeng, ZHANG Lihua, LIU Xinghua, et al. Discrimination of apple mould heart disease by near infrared diffuse reflectance spectroscopy based on principal component analysis[J]. Transactions of the Chinese Society for Agricultural Machinery, 2011, 42(10): 158-161. (in Chinese)
- [24] TANG Y M, FENG F, GUO Z Z, et al. Integrating principal component analysis with statistically-based models for analysis of causal factors and landslide susceptibility mapping: a comparative study from the loess plateau area in Shanxi (China)[J]. Journal of Cleaner Production, 2020, 277:124159.