

孙玉婷, 杨红云, 王映龙, 等. 基于支持向量机的水稻叶面积测定[J]. 江苏农业学报, 2018, 34(5): 1027-1035.
doi:10.3969/j.issn.1000-4440.2018.05.009

基于支持向量机的水稻叶面积测定

孙玉婷¹, 杨红云^{2,3}, 王映龙^{1,3}, 周 琼¹, 孙爱珍³, 杨文姬^{2,3}

(1. 江西农业大学计算机与信息工程学院, 江西 南昌 330045; 2. 江西农业大学软件学院, 江西 南昌 330045; 3. 江西省高等学校农业信息技术重点实验室, 江西 南昌 330045)

摘要: 本研究探讨了水稻叶长、叶宽与叶面积之间的关系。应用改进网格搜索算法参数优化后的支持向量机对计算机图像处理技术测量所得到的样本数据进行训练, 以水稻叶片长度(L)、叶片最大宽度(W)、叶面积回归拟合值(S)作为输入变量对水稻叶面积进行预测。结果表明, 在不同的水稻品种下, 相比于其他模型本研究提出的支持向量机Ⅲ模型预测结果的均方根误差和平均相对误差最低, 均方根误差分别为1.882 1(两优培9)、1.387 3(金优458)和1.348 2(中早35), 平均相对误差分别为2.901 4%(两优培9)、5.273 5%(金优458)和4.929 3%(中早35)。该模型能较真实地反映水稻叶面积的实际大小, 满足农业科学研究的需求, 为植物叶面积的预测提供了一种新的方法。

关键词: 叶长; 叶宽; 叶面积; 网格搜索算法; 支持向量机

中图分类号: TP751 **文献标识码:** A **文章编号:** 1000-4440(2018)05-1027-09

Determination of rice leaf area based on support vector machine

SUN Yu-ting¹, YANG Hong-yun^{2,3}, WANG Ying-long^{1,3}, ZHOU Qiong¹, SUN Ai-zhen³, YANG Wen-ji^{2,3}

(1. School of Computer and Information Engineering, Jiangxi Agricultural University, Nanchang 330045, China; 2. School of Software, Jiangxi Agricultural University, Nanchang 330045, China; 3. Key Laboratory of Agricultural Information Technology of Colleges and Universities in Jiangxi Province, Nanchang 330045, China)

Abstract: This study explored the relationship among rice leaf length, leaf width and leaf area. The support vector machine(SVM) based on improved grid search algorithm was used to train the sample data obtained by the computer image processing technique. Rice leaf area was predicted by using the leaf length (L), leaf maximum width (W) and leaf area regression fitting value (S) as input variables. The results showed that the root mean square error and average relative error of SVMⅢ model prediction results proposed in this study were the lowest compared with other models. The root mean square error were 1.882 1 (LYP9), 1.387 3 (JY458) and 1.348 2 (ZZ35), respectively. The average relative errors were 2.901 4% (LYP9), 5.273 5% (JY458) and 4.929 3% (ZZ35), respectively. The model can reflect the actual size of rice leaf area more faithfully and meet the needs of agricultural science research, which provides a new method for plant leaf area prediction.

Key words: leaf length; leaf width; leaf area; grid search algorithm; support vector machine

收稿日期: 2018-02-28

基金项目: 国家自然科学基金项目(61562039、61363041、61462038);

江西省教育厅科技项目(GJJ160374、GJJ170279)

作者简介: 孙玉婷(1995-), 女, 江西萍乡人, 硕士研究生, 研究方向为机器学习与农业信息技术。(E-mail) 18270826309@163.com

通讯作者: 杨红云, (E-mail) nc_yhy@163.com

植物叶片是植物进行光合作用和蒸腾作用的重要器官^[1], 其叶面积的大小决定了植物光合面积的大小, 对植物的生理生态、作物栽培、作物产量都具有十分重要的意义。传统的叶片面积的测量方法主要有网格法、复印称质量法、系数法和叶面积仪器测量法^[2-3]等。

但是上述方法存在破坏植株生长,费工费时或价格昂贵等缺点,不利于相关工作的开展^[4]。不少学者为能简便、准确、迅速地测量植物的叶面积做了大量研究,分别对蓝莓^[5]、番茄^[6]、香梨^[7]、木菠萝^[8]和节瓜^[9]等植物叶面积进行了回归分析,得到了相关叶面积的拟合方程。随着计算机技术的高速发展又衍生出了基于数字图像处理技术的测量方法,杨劲峰等^[10]通过平台扫描仪获得叶片数字图像并建立了运用数字图像处理测定蔬菜叶面积的方法,并与目前较为常用的叶面积测量方法进行比较分析,得到图像处理方法与上述方法测定结果显著的线性相关。赵英等^[11]利用数码相机获取芒果叶片的数字图像,并利用 Photoshop 图像处理软件计算其叶面积,测定结果与传统的方法结果差异不显著。高君亮等^[12]通过 3 种不同的方法(方格法、称质量法和数字图像法)对 3 种杨树(小叶杨、新疆杨和二白杨)的叶面积进行测定,分析得到数字图像法与其他 2 种方法的测定结果具有显著正线性相关。崔世钢等^[13]采用基于 Photoshop 图像法对数码相机所获得油菜叶片图像进行分析测定,将所得到的油菜叶面积与叶面积仪测定法、方格法的测量结果进行回归分析,结果存在极显著线性相关。虽已有人对水稻的叶面积指数进行了大量的研究^[14-16],但是利用机器学习方法预测单株水稻叶面积还鲜见报道。本研究通过机器学习的方法建立水稻叶长、叶宽及叶面积回归拟合值与叶面积的模型,探求一种方便、快捷、误差小且不影响水稻植株生产的测量方法,从而实现简便、快速、正确估算水稻叶面积,为研究水稻高产栽培技术和生产管理提供参考。

1 材料与方法

1.1 试验设计

水稻田间试验于 2015-2017 年在江西农业大学农学试验站和江西省成新农场进行,供试品种为籼型两系杂交稻品种两优培九(LYP9),籼型三系杂交水稻品种金优 458(JY458)、常规稻品种中早 35(ZZ35)。分别采集两优培九水稻品种 1 360 组叶片,金优 458 水稻品种 720 组叶片以及中早 35 水稻品种 680 组叶片作为观测样本。

1.2 数据获取与分析

1.2.1 水稻叶片数字图像获取 水稻取样后,扫描 4 种不同施氮水平下水稻样品的顶三叶的数字图像。选择平板扫描仪 MRS-9600TFU2,分辨率设置

为 600 dpi,缩放比例为 100%,RGB 格式进行扫描,并以 JPG 格式存储于计算机中。植株活体剪下叶片取样后,整齐地平放在扫描仪工作台上,尽快扫描避免水稻叶片因蒸腾水分变形造成误差。

1.2.2 图像预处理 由于多种因素的影响,图像在生成、传输和变换的过程中会造成图像质量下降,为了避免图像的分析误差,首先对图像进行增强处理,然后将其转化为二值化图像。采用开运算和闭运算去除二值化图像上由于光照过强所形成的光斑,最后对图像进行边缘提取,得到水稻叶片的轮廓。

1.2.3 水稻叶片长度和宽度的获取 采用 min-boundrect 函数可以得到水稻叶片最小外接矩形的 4 个顶点 $(c_1, r_1), (c_2, r_2), (c_3, r_3), (c_4, r_4)$,根据 2 点间的距离公式,可以获得水稻叶片的长 L_0 ,宽 W_0 。

$$L_0 = \sqrt{(c_2 - c_1)^2 + (r_2 - r_1)^2} \quad (1)$$

$$W_0 = \sqrt{(c_4 - c_1)^2 + (r_4 - r_1)^2} \quad (2)$$

将水稻叶片长、宽的像素值与真实值之间进行转换。转换公式如下。

$$L = \frac{L_0}{600} \times 2.54 \quad (3)$$

$$W = \frac{W_0}{600} \times 2.54 \quad (4)$$

公式(3)、(4)中, L_0 、 W_0 分别为叶片长、宽的像素值, L 、 W 分别为叶片长、宽的真实值。

1.2.4 水稻叶面积的获取 数字图像由许多像素点组成,每个像素点代表一定的实际面积值^[17],为了确定扫描仪采集的样本图片像素代表的实际面积大小,在扫描图片时引入了一个已知面积大小的参考物。

$$S_1 = S_2 \times \frac{P_1}{P_2} \quad (5)$$

公式(5)中 S_1 为水稻的叶面积, S_2 为参照物的实际面积(本研究中 S_2 定为 $2.5 \text{ cm} \times 2.5 \text{ cm}$), P_1 为水稻叶片图像所包含的像素点, P_2 为参照物图像所包含的像素点。

1.3 数据预处理

为了提高运行效率和建模的精度^[18],将不同量纲的数据归一化至区间 $[0, 1]$,采用的归一化原理如公式(6)。

$$y = \frac{(y_{\max} - y_{\min}) \cdot (x - x_{\min})}{(x_{\max} - x_{\min})} + y_{\min} \quad (6)$$

公式(6)中 x 表示样本数据, x_{\max} 和 x_{\min} 分别表

示样本数据中的最大值和最小值, y 表示归一化后的数据, y_{\max} 和 y_{\min} 分别表示归一化后数据的最大值和最小值。

1.4 改进网格搜索算法

支持向量机是建立在统计学习理论的 VC 维理论和结构风险最小原则基础上的新型机器学习方法^[19]。它在解决小样本、非线性和高维模式识别问题中表现出许多特有的优势^[20],并在一定程度上克服了“维数灾难”和“过学习”等问题。支持向量机应用于实际问题时能否取得良好的预测结果,取决于能否成功地设置该算法的关键参数^[21],其主要参数为惩罚系数 c 和核函数参数 g 。网格搜索算法的原理是让 c 和 g 在一定的范围划分网格并遍历网格内所有点进行取值,对于选定的 c 和 g 利用 K-CV 方法得到此组 c 和 g 的预测误差,最终选择使训练集交叉验证分类准确率最高的那组 c 和 g 作为最优参数组合。传统的网格搜索算法搜索对于较小的步

距会影响寻优效率,而对于较大的步距容易陷入局部最优。本研究采用改进的网格搜索算法优化支持向量机不仅具有良好的效果,还能大幅度减少参数的寻优时间。其算法原理是先在较大范围内采用大步距粗搜获得局部最优参数组合,再在这组参数附近选择一个小区间,采用传统的小步距进行二次精搜^[22],得到最后的最优参数组合。

2 结果与分析

2.1 水稻叶长、叶宽误差

以试验观测到的两优培九水稻数据为分析样本,随机选取 20 组,采用人工和计算机图像处理技术对水稻叶长、叶宽进行测量,其测量结果对比见表 1。

从表 1 可以看出,与人工测量相比,采用计算机图像处理测量水稻叶长最大相对误差为 0.97%,平均相对误差为 0.40%,而水稻叶宽最大相对误差为 5.88%,平均相对误差为 3.10%,能够满足本研究的要求。

表 1 水稻叶长叶宽误差分析

Table 1 The error analysis of rice leaf length and leaf width

| | 水稻叶长 | | | 水稻叶宽 | | |
|----|-----------|-----------|----------|-----------|-----------|----------|
| | 人工测量 (cm) | 图像处理 (cm) | 相对误差 (%) | 人工测量 (cm) | 图像处理 (cm) | 相对误差 (%) |
| 1 | 44.3 | 44.4 | 0.23 | 1.7 | 1.7 | 0 |
| 2 | 52.9 | 52.4 | 0.95 | 2.0 | 2.0 | 0 |
| 3 | 60.2 | 60.5 | 0.50 | 1.8 | 1.8 | 0 |
| 4 | 54.0 | 54.2 | 0.37 | 2.0 | 1.9 | 5.00 |
| 5 | 56.2 | 56.2 | 0 | 1.9 | 1.8 | 5.26 |
| 6 | 42.2 | 42.0 | 0.47 | 1.7 | 1.8 | 5.88 |
| 7 | 45.1 | 45.1 | 0 | 1.7 | 1.8 | 5.88 |
| 8 | 47.6 | 47.7 | 0.21 | 1.8 | 1.9 | 5.56 |
| 9 | 50.2 | 50.4 | 0.40 | 1.8 | 1.9 | 5.56 |
| 10 | 48.3 | 48.4 | 0.21 | 1.9 | 1.9 | 0 |
| 11 | 43.2 | 43.3 | 0.23 | 1.7 | 1.7 | 0 |
| 12 | 62.1 | 61.5 | 0.97 | 2.0 | 2.0 | 0 |
| 13 | 43.2 | 43.0 | 0.46 | 1.8 | 1.8 | 0 |
| 14 | 54.8 | 54.8 | 0 | 1.7 | 1.7 | 0 |
| 15 | 51.8 | 52.1 | 0.58 | 1.7 | 1.8 | 5.88 |
| 16 | 47.2 | 47.4 | 0.42 | 1.8 | 1.7 | 5.56 |
| 17 | 51.7 | 51.6 | 0.19 | 1.7 | 1.7 | 0 |
| 18 | 44.8 | 44.5 | 0.67 | 1.7 | 1.8 | 5.88 |
| 19 | 48.1 | 47.9 | 0.42 | 1.7 | 1.8 | 5.88 |
| 20 | 55.4 | 55.0 | 0.72 | 1.8 | 1.9 | 5.56 |

2.2 不同寻优方法优化支持向量机参数

以试验观测到的两优培九水稻的 1 360 组数据为分析样本,对样本数据进行归一化处理,选择其中 1 020 组数据为训练集,340 组数据为测试集,水稻的叶长和叶宽为模型的输入变量,水稻叶面积为模型的输出变量。

采用改进的网格搜索算法对惩罚系数 c 和 RBF 核参数 g 寻优,设定初始网格搜索 c 和 g 的范围均为 $[2^{-8}, 2^8]$,搜索步距均设为 1,采用 K-CV 方法对训练集进行测试,其中 $K=5$,得到局部最优参数组合,在得到的局部最优参数附近重新定义搜索范围和步距,其 c 和 g 搜索范围均为 $[2^{-2}, 2^2]$,搜索步距均为 0.1。二次寻优的结果如图 1、图 2 所示。

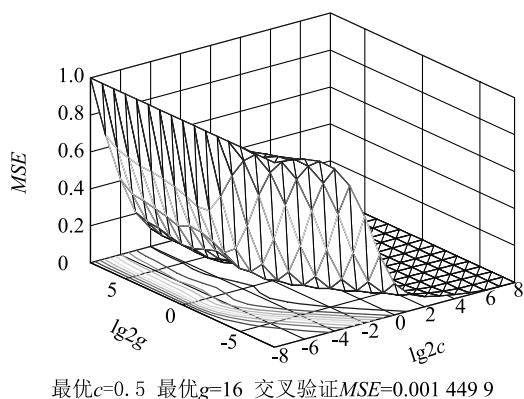


图 1 支持向量机 (SVM) 参数粗略选择结果

Fig.1 Rough selection of support vector machine (SVM) parameters

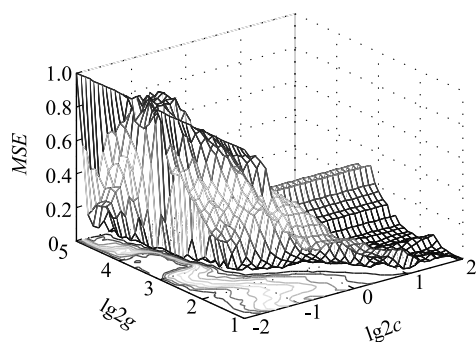


图 2 支持向量机 (SVM) 参数精细选择结果

Fig.2 Fine selection of SVM parameters

选择好最优参数组合 $[c, g]$ 后,对训练集进行训练,模型回归结果显示相关系数为 0.979 942,表明该模型的回归能力较好,为了进一步测试该模型

的精度,对测试集进行预测,本研究分别采用了传统网格搜索算法,与改进网格搜索算法进行比较,包括最终模型参数、均方根误差(反归一化后)以及寻优时间的相关系数,其对比结果见表 2。

由表 2 可知,采用改进的网格搜索算法寻优时间约为传统网格搜索算法的 1/53,但由于其在二次寻优的区间选择上含有较多的经验成分,所以均方根误差略低于传统网格搜索算法。牺牲了一点均方根误差而节省了大量的寻优时间是可以接受的,因此应用改进的网格搜索算法建立水稻叶面积模型是有效可行的。

表 2 不同网格搜索算法参数寻优的结果

Table 2 Comparison of optimization results of different grid search algorithms

| | 参数 c | 参数 g | 均方根误差 | 时间 (s) |
|----------|---------|--------|---------|-------------|
| 传统网格搜索算法 | 0.329 9 | 128 | 1.789 8 | 1 411.051 9 |
| 改进网格搜索算法 | 0.406 1 | 32 | 1.887 9 | 26.684 8 |

2.3 不同水稻品种叶面积的预测

本研究以试验观测到的 3 种不同水稻叶片数据为依据,采用参数优化后的支持向量机算法对水稻叶面积进行训练和预测。采用归一化法将观测数据处理到 $(0, 1)$ 区间,两优培九水稻品种选取 1 020 组数据为训练样本,340 组数据为预测样本。金优 458 水稻品种选取 540 组数据为训练样本,180 组数据为预测样本,中早 35 水稻品种选取 510 组数据为训练样本,170 组数据为预测样本。

2.3.1 两优培九水稻品种叶面积的预测 以试验观测到的两优培九水稻的 1 020 组数据为分析样本,以叶长、叶宽为自变量,以图像处理所得到的叶面积为因变量进行回归分析(图 3),回归方程达到显著水平 ($P < 0.05$),水稻的叶长、叶宽与叶面积呈线性关系,其回归方程为:

$$S_{LYP9} = -45.370\ 1 + 1.336\ 6L + 26.664\ 7W \quad (7)$$

公式(7)中, L 为叶长, W 为叶宽, S 为叶面积,回归方程的决定系数为 0.960 5。

将水稻叶长、叶宽的观测值代入公式(7)中,求出对应的叶面积回归拟合值(S),将水稻叶长(L)、叶宽(W)、叶面积回归拟合值(S)作为模型的三维输入变量,水稻叶面积作为模型的输出变量,其样本训练如图 4 所示。

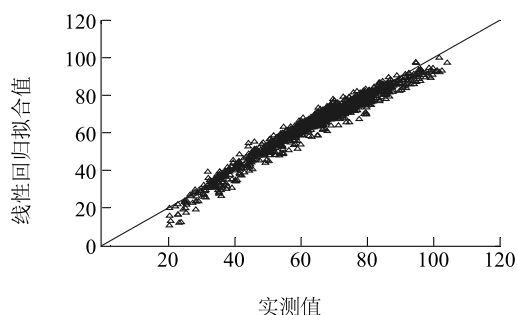


图3 两优培九水稻叶面积线性回归拟合

Fig.3 The linear regression fitting of rice leaf area

图4显示,水稻品种的模型在样本训练过程中表现出了良好的数据泛化能力,在训练集上验证模型的回归效果,结果显示平方相关系数为98.017%,根据训练后获得的最优模型参数,并对剩余的340组样本数据进行预测,其预测结果如图5所示,为了对比本研究提出的模型在叶面积测量上的效果,分别用其他3种不同的模型对相同的数据进行叶面积预测,其一以水稻叶长(L)、叶宽(W)为二维输入变量,叶面积为输出变量建立支持向量机模型,其二以水稻叶长(L)、叶宽(W)、叶长叶宽乘积($L \times W$)为三维输入变量,叶面积为输出变量,建立支持向量机模型,其三以水稻叶长(L)、叶宽(W)为二维输入变量,叶面积为输出变量,建立线性回归模型。为了定量分析4种模型的预测效果,本研究引入均方根误差和平均相对误差2个评价指标,评价结果如表3所示。

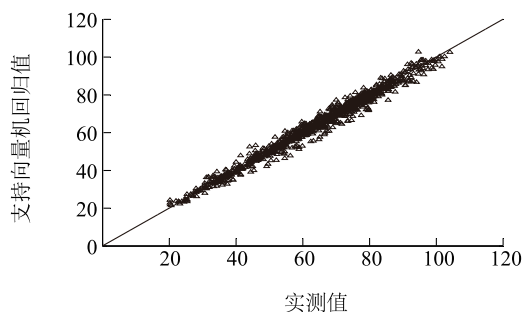


图4 两优培九水稻叶面积支持向量机训练结果

Fig.4 The training results based on SVM of rice leaf area

通过比较分析,4种模型的预测精度均比较高,本研究提出的支持向量机Ⅲ模型均方根误差和平均相对误差均小于其他3个模型,考虑到不完全抽样的性质,本研究需要进一步对差异进行显著性检验,原假设为配对样本无显著性差异,本研究利用SPSS软件对上述4种模型的预测结果进行Wilcoxon检验,得到结果如表4所示。

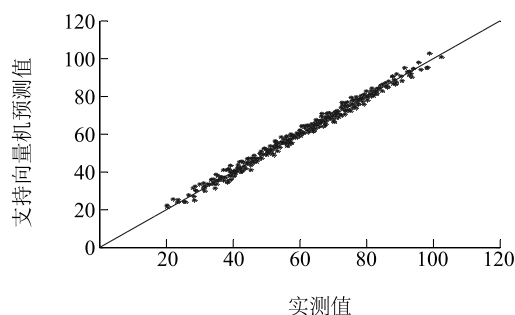


图5 两优培九水稻叶面积支持向量机预测结果

Fig.5 The prediction results based on SVM of rice area

表3 不同模型预测两优培九水稻叶面积结果评价

Table 3 The result of rice leaf area predicted by different models

| 模型 | 均方根误差 | 平均相对误差(%) |
|------------------------------------|---------|-----------|
| 线性回归 | 3.207 2 | 4.808 3 |
| 支持向量机Ⅰ(L 、 W) | 1.887 9 | 2.906 0 |
| 支持向量机Ⅱ(L 、 W 、 $L \times W$) | 1.893 7 | 2.913 6 |
| 支持向量机Ⅲ(L 、 W 、 S) | 1.882 1 | 2.901 4 |

从Wilcoxon检验结果来看,除了支持向量机Ⅰ(L 、 W)-支持向量机Ⅱ(L 、 W 、 $L \times W$)与支持向量机Ⅰ(L 、 W)-支持向量机Ⅲ(L 、 W 、 S),其他配对样本的检验结果显著性 P 值均小于给定的显著性水平0.05,拒绝原假设,判定模型之间存在显著性差异。

2.3.2 金优458、中早35水稻品种叶面积的预测

为了验证本研究提出的模型的普适性,采用另外2种不同的水稻品种金优458和中早35进行验证。分别取试验观测到的金优458的540组数据以及中早35的510组数据为分析样本,均以水稻叶长、叶宽为自变量,以图像处理所得到的叶面积为因变量进行回归分析(图6、图7),回归方程达到显著水平($P < 0.05$),水稻的叶长、叶宽与叶面积呈线性关系,其回归方程为:

$$S_{JY458} = -21.333\ 6 + 0.781\ 9L + 18.455\ 9W \quad (8)$$

$$S_{ZZ35} = -22.496\ 0 + 0.904\ 6L + 17.775\ 8W \quad (9)$$

公式(8)、公式(9)中, L 为叶长, W 为叶宽, S 为叶面积,JY458水稻品种回归方程的决定系数为0.951 6,ZZ35水稻品种回归方程的决定系数为0.967 5。将水稻叶长、叶宽的观测值代入公式(8)、公式(9)中,求出对应的叶面积回归拟合值(S),将水稻叶长(L)、叶宽(W)、叶面积回归拟合值(S)作

为模型的三维输入变量,水稻叶面积作为模型的输出变量,其样本训练如图 8、图 9 所示。

表 4 两优培九水稻叶面积的显著性检验结果

Table 4 Significant test results of rice (LYP9) rice leaf area

| 方法 | 组别 | Z 值 | P 值 |
|-------------|---|--------|-------|
| Wilcoxon 检验 | 线性回归-支持向量机 I (L, W) | -2.537 | 0.011 |
| | 线性回归-支持向量机 II ($L, W, L \times W$) | -2.527 | 0.012 |
| | 线性回归-支持向量机 III (L, W, S) | -2.505 | 0.012 |
| | 支持向量机 I (L, W)-支持向量机 II ($L, W, L \times W$) | -1.750 | 0.080 |
| | 支持向量机 I (L, W)-支持向量机 III (L, W, S) | -0.600 | 0.548 |
| | 支持向量机 II ($L, W, L \times W$)-支持向量机 III (L, W, S) | -3.493 | 0 |

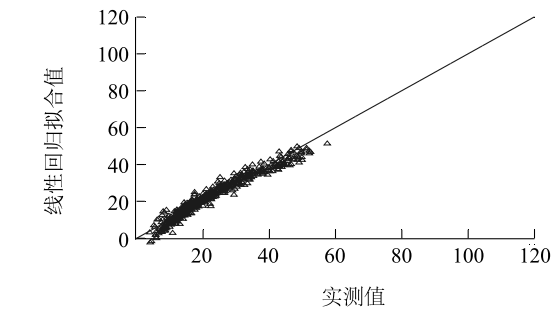


图 6 金优 458 水稻叶面积线性回归拟合
Fig.6 The linear regression fitting of rice (JY458) leaf area

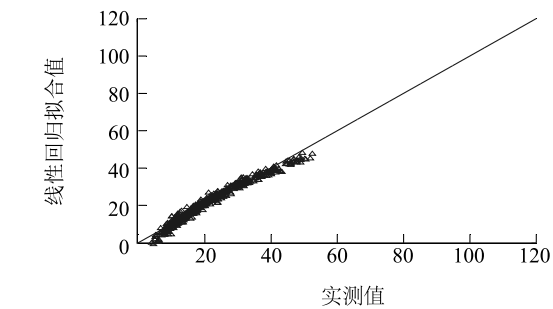


图 7 中早 35 水稻叶面积线性回归拟合
Fig.7 The linear regression fitting of rice (ZZ35) leaf area

从图 8、图 9 可以看出,2 种水稻品种的模型在样本训练过程中表现出了良好的数据泛化能力,2 种水稻品种的叶面积均稳定在 0~60,基于参数优化后的支持向量机回归对于 2 种不同水稻品种训练样本的平方相关系数分别为 97.15% (JY458) 和 98.62% (ZZ35),根据训练后获得的最优模型参数,对剩余的 180 组 (JY458)、170 组 (ZZ35) 样本数据进行预测,其预测结果如图 10、图 11 所示,为了对比本研究提出的支持向量机 III 模型在叶面积测量上的效果,分别用其他 3 种不同的模型对相同的数据

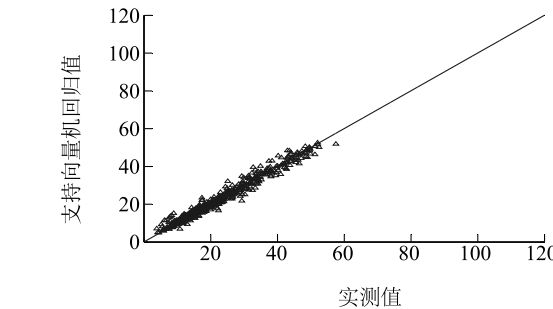


图 8 金优 458 水稻叶面积支持向量机训练结果
Fig.8 The training results based on SVM of rice (JY458) leaf area

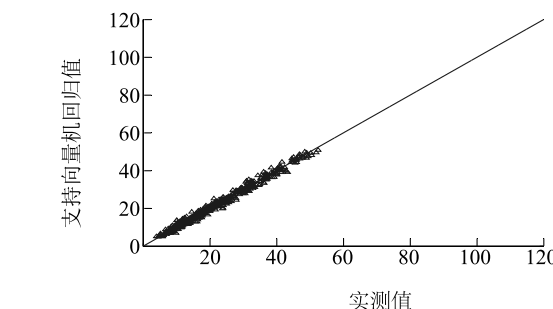


图 9 中早 35 水稻叶面积支持向量机训练结果
Fig.9 The training results based on SVM of rice (ZZ35) leaf area

进行叶面积预测,其一以水稻叶长 (L)、叶宽 (W) 为二维输入变量,以叶面积为输出变量建立支持向量机模型,其二以水稻叶长 (L)、叶宽 (W)、叶长叶宽乘积 ($L \times W$) 为三维输入变量,以叶面积为输出变量,建立支持向量机模型,其三以水稻叶长 (L)、叶宽 (W) 为二维输入变量,以叶面积为输出变量,建立线性回归模型。引入均方根误差和平均相对误差 2 个评价指标定量分析 4 种模型的预测效果,其评价结果如表 5 所示。

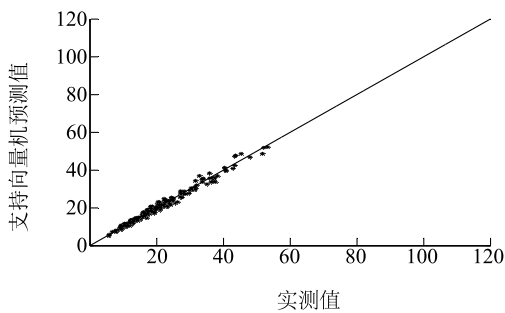


图 10 水稻叶面积支持向量机预测结果(金优 458)

Fig.10 The prediction results based on SVM of rice (JY458) leaf area

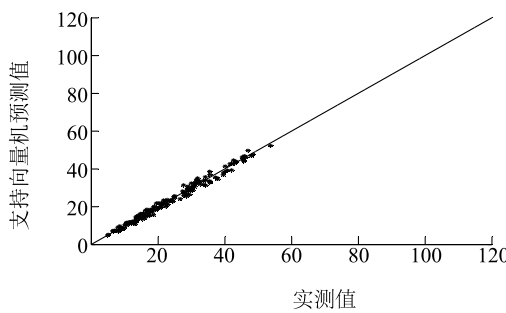


图 11 水稻叶面积支持向量机预测结果(中早 35)

Fig.11 The prediction results based on SVM of rice (ZZ35) leaf area

表 5 不同模型预测金优 458、中早 35 水稻叶面积结果评价

Table 5 The result of rice leaf area predicted by different models

| 模型 | 金优 458 | | 中早 35 | |
|--------------------|---------|------------|---------|------------|
| | 均方根误差 | 平均相对误差 (%) | 均方根误差 | 平均相对误差 (%) |
| 线性回归 | 1.927 1 | 9.073 0 | 1.978 2 | 8.994 5 |
| 支持向量机 I (L、W) | 1.548 4 | 5.865 2 | 1.353 6 | 4.969 5 |
| 支持向量机 II (L、W、L×W) | 1.425 3 | 5.389 1 | 1.365 8 | 4.992 1 |
| 支持向量机 III (L、W、S) | 1.387 3 | 5.273 5 | 1.348 2 | 4.929 3 |

通过比较分析,2 个水稻品种在本研究提出的支持向量机 III 模型预测下,其结果的均方根误差和平均相对误差均小于其他 3 个模型,考虑到不完全抽样的性质,本研究需要进一步对差异进行显著性

检验,原假设为配对样本无显著性差异,利用 SPSS 软件对上述 4 种模型的预测结果进行检验,2 种水稻的显著性检测结果如表 6、表 7 所示。

表 6 金优 458 水稻叶面积的显著性检验结果

Table 6 Significant test results of rice (JY458) rice leaf area

| 方法 | 组别 | Z 值 | P 值 |
|-------------|--------------------------------------|--------|-------|
| Wilcoxon 检验 | 线性回归-支持向量机 I (L、W) | -3.456 | 0.001 |
| | 线性回归-支持向量机 II (L、W、L×W) | -3.951 | 0 |
| | 线性回归-支持向量机 III (L、W、S) | -3.863 | 0 |
| | 支持向量机 I (L、W)-支持向量机 II (L、W、L×W) | -0.311 | 0.755 |
| | 支持向量机 I (L、W)-支持向量机 III (L、W、S) | -0.016 | 0.987 |
| | 支持向量机 II (L、W、L×W)-支持向量机 III (L、W、S) | -2.143 | 0.032 |

表 7 中早 35 水稻叶面积的显著性检验结果

Table 7 Significant test results of rice (ZZ35) rice leaf area

| 方法 | 组别 | Z 值 | P 值 |
|-------------|--------------------------------------|--------|-------|
| Wilcoxon 检验 | 线性回归-支持向量机 I (L、W) | -1.521 | 0.128 |
| | 线性回归-支持向量机 II (L、W、L×W) | -1.971 | 0.049 |
| | 线性回归-支持向量机 III (L、W、S) | -1.596 | 0.110 |
| | 支持向量机 I (L、W)-支持向量机 II (L、W、L×W) | -0.801 | 0.423 |
| | 支持向量机 I (L、W)-支持向量机 III (L、W、S) | -0.934 | 0.350 |
| | 支持向量机 II (L、W、L×W)-支持向量机 III (L、W、S) | -5.413 | 0 |

从 Wilcoxon 检验结果来看,金优 458 水稻品种中除了支持向量机 I (L 、 W)-支持向量机 II (L 、 W 、 $L \times W$) 与支持向量机 I (L 、 W)-支持向量机 III (L 、 W 、 S),其他配对样本的 P 值均小于给定的显著性水平 0.05,拒绝原假设,判定其他模型之间存在显著性差异。而在中早 35 水稻品种中,除了线性回归-支持向量机 II (L 、 W 、 $L \times W$) 和支持向量机 II (L 、 W 、 $L \times W$)-支持向量机 III (L 、 W 、 S), P 值小于给定的显著性水平 0.05,模型间具有显著性差异,其他模型间显著性差异不明显。

3 讨论

本研究建立的 4 种叶面积模型均可以实现水稻叶面积的快速、无损测定。通过改进网格搜索算法参数优化后的支持向量机建立水稻叶面积模型并预测叶面积,模型的输入参数为水稻叶长 (L)、叶宽 (W) 以及回归拟合值 (S),试验样本为 3 种水稻品种的数据。结果表明,水稻叶长、叶宽、回归拟合值与叶面积之间存在显著的相关性,模型预测结果误差小,可以较好地应用于叶面积的测定,具有较好的使用价值和应用前景。

目前关于图像处理以及支持向量机方法应用于农产品的研究已有一些进展。程洪等^[23]使用图像处理技术以及传统网格搜索算法参数优化后的支持向量机建立树上苹果早期模型,模型的预测效果较好。鉴于前期研究成果得到利用图像处理技术对水稻叶片面积的计算误差小于 5%^[24],本研究通过图像处理技术求取目标物的最小外接矩形以获得水稻叶长、叶宽,叶长的平均相对误差为 0.40%,叶宽的平均相对误差为 3.10%。本研究还采用了改进的网格搜索算法对支持向量机参数进行寻优,虽然牺牲了均方根误差,但大大缩减了大量的样本空间下支持向量机参数寻优的时间,为模型的预测提供了一定的实时性。

水稻叶面积是水稻株型研究通常需要测量的形态指标^[25]。关于植物叶面积的测量方法较为常见的有:复印称质量法,长宽矫正法,叶面积仪测定法等。这些方法各有利弊,尤其对于野外测量,需要消耗大量的时间与精力,而且操作复杂影响测量精度^[26]。本研究通过图像处理技术对水稻叶片进行简单测量,采用多元线性回归和改进网格搜索算法参数优化的支持向量机建立模型预测水稻叶面积。

3 种水稻品种基于支持向量机预测水稻叶面积模型的均方根误差和平均相对误差均小于多元线性回归方法,且其中 2 种水稻品种(两优培 9、金优 458)基于支持向量机的模型与线性回归模型之间具有显著差异($P < 0.05$),表明支持向量机模型能更好地预测水稻叶面积,具有较强的预测精度。

模型变量的选择会影响模型的精度和复杂性。不少研究通过图像处理技术对植物叶片进行长、宽、面积的测算,发现叶面积与叶长、叶宽、长宽乘积显著相关^[27-29]。本研究不仅证明了上述结论,还利用支持向量机模型对叶面积进行预测,预测效果良好。本研究还提出了一种新的思路,通过多元线性回归建立数学模型,得到 3 种水稻品种的叶面积回归方程,将水稻叶片长宽的实测值代入回归方程中得到叶面积回归拟合值,并将拟合值与叶长、叶宽构成一个三维输入变量,建立支持向量机模型预测水稻叶面积,模型的预测效果良好,相比于其他模型,对于两优培 9、金优 458、中早 35 3 种水稻品种,本研究提出模型的均方根误差和平均相对误差均为最低。Wilcoxon 检验结果表明,在统计学意义上,本研究提出的支持向量机 III 模型与其他 2 种支持向量机模型的预测结果具有较高一致性,显示本研究提出的支持向量机 III 模型具有较强的适用性和有效性,与支持向量机 II 相比,显著性差异明显,说明本研究提出的支持向量机 III 模型预测水稻叶面积更为准确,为水稻叶面积的测定提供了一种新的方法。

参考文献:

- [1] 左欣,韩斌,程嘉林. 基于数字图像处理的植物叶面积测量方法[J]. 计算机工程与应用, 2006, 42(27):194-196.
- [2] 王忠芝,张金瑞. 基于图像处理的叶面积测量方法[J]. 网络新媒体技术, 2010, 31(5):68-72.
- [3] 赵燕,刘晶,王辉. 洛阳市 13 种绿化树种叶面积的回归测算[J]. 江苏农业科学, 2016, 44(10):254-257.
- [4] 陈超君,黄敏,尹小红. 广金钱草叶面积测量方法的研究[J]. 作物杂志, 2007(5):46-48.
- [5] 曹红梅,宋世鑫,谢兆森. 蓝莓叶面积测定回归方程的建立[J]. 北方园艺, 2014(24):1-6.
- [6] 薛义霞,栗东霞,李亚灵. 番茄叶面积测量方法的研究[J]. 西北农林科技大学学报(自然科学版), 2006, 34(8):116-120.
- [7] 刘洪波,张江辉,白云岗,等. 香梨叶面积测定方法对比研究[J]. 新疆农业科学, 2013, 50(3):453-459.
- [8] 谢德鸿,陈诗隆,杨伟中,等. 木菠萝叶面积无损测定方法研究[J]. 东南园艺, 2010(1):4-7.

- [9] 宋世威,刘厚诚,陈日远. 节瓜叶面积回归方程的比较[J]. 中国蔬菜, 2004, 1(6):18-20.
- [10] 杨劲峰,陈清,韩晓日,等. 数字图像处理技术在蔬菜叶面积测量中的应用[J]. 农业工程学报, 2002, 18(4):155-158.
- [11] 赵英,付海天,黄国弟,等. 基于图像处理技术的芒果叶面积非损伤估测技术初探[J]. 中国南方果树, 2013, 42(5):87-88.
- [12] 高君亮,郝玉光,张景波,等. 基于数字图像处理的防护林体系三种杨树叶面积测定[J]. 农机化研究, 2013(7):39-42.
- [13] 崔世钢,秦建华,图像处理法测定油菜叶面积的研究[J]. 湖北农业科学, 2017, 56(14):2756-2757.
- [14] 陆秀明,黄庆,孙雪晨,等. 图像处理技术估测水稻叶面积指数的研究[J]. 中国农学通报, 2011, 27(3):65-68.
- [15] 孙涛,刘振波,葛云健,等. 基于数码相片 Gamma 校正的水稻叶面积指数估算[J]. 生态学报, 2014, 34(13):3548-3557.
- [16] 顾婷,高晓雯,孙成明,等. 基于无人机图像的水稻拔节期叶面积指数估算[J]. 农业网络信息, 2015(12):63-65.
- [17] 吴大伟,李聪,徐丽,等. 基于图像处理技术的植物叶片相关参数测量研究[J]. 科技信息, 2010(4):106.
- [18] 常军,李祯,朱业玉,等. 基于支持向量机(SVM)方法的冬季温度预测[J]. 气象科技, 2005(S1):102-106.
- [19] 王磊. 支持向量机学习算法的若干问题研究[D]. 成都:电子科技大学,2007.
- [20] 丁世飞,齐丙,谭红艳. 支持向量机理论与算法研究综述[J]. 电子科技大学学报, 2011, 40(1):2-10.
- [21] 彭光金,司海涛,俞集辉,等. 改进的支持向量机算法及其应用[J]. 计算机工程与应用, 2011, 47(18):218-221.
- [22] 王健峰,张磊,陈国兴,等. 基于改进的网格搜索法的 SVM 参数优化[J]. 应用科技, 2012(3):28-31.
- [23] 程洪,LUTZ DAMEROW,MICHAEL BLANKE,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]. 农业机械学报, 2015, 46(3):9-14, 22.
- [24] 杨红云,孙爱珍,何火娇. 水稻叶片几何参数图像视觉测量方法研究[J]. 湖北农业科学, 2015(17):4317-4320.
- [25] 马敏,李绪孟. 水稻叶面积测量方法的研究[J]. 湖南农机, 2014(6):119-120.
- [26] 白云岗,董新光,张江辉,等. 无核白葡萄叶面积及一年生枝条干物质质量简易测定方法研究[J]. 新疆农业科学, 2010, 47(9):1744-1748.
- [27] 卜海东,顾蔚,齐永平,等. 基于图像处理华中五味子叶面积的回归测算[J]. 植物生理学报, 2008, 44(3):543-547.
- [28] 安佳佳,李茂富,黄绵佳,等. 香蕉叶面积回归模型的建立[J]. 南方农业学报, 2009, 40(6):724-727.
- [29] 曾燕蓉,朱方容,林强,等. 几种桑树叶面积测量方法及叶片大小与叶片质量的相关性分析[J]. 蚕业科学, 2013(5):868-876.

(责任编辑:陈海霞)