

基于高光谱图像的龙眼叶片叶绿素含量分布模型

岳学军^{1,2} 凌康杰^{1,2} 洪添胜^{1,3} 甘海明^{1,2} 刘永鑫^{1,2} 王林惠^{1,2}

(1. 华南农业大学南方农业机械与装备关键技术省部共建教育部重点实验室, 广州 510642;

2. 华南农业大学电子工程学院, 广州 510642; 3. 华南农业大学工程学院, 广州 510642)

摘要: 针对传统高光谱单点法检测叶绿素含量效率低、精度不足等问题, 提出一种基于高光谱图像和卷积神经网络(CNN)多特征融合的深度学习方法, 对龙眼叶片叶绿素含量分布预测模型。首先进行 Savitzky-Golay 光谱去噪, 然后通过奇异值分解(SVD)和独立成分分析(ICA)提取特征光谱, 再对特征光谱图像提取灰度共生矩阵(GLCM)和CNN纹理特征, 最后建立粒子群优化(PSO)支持向量回归(SVR)、深度神经网络(DNNs)分布模型。结果表明, 基于特征光谱建模的 PSO-SVR 预测效果最佳, 全期的校正集和验证集模型决定系数 R^2 为 0.822 0 和 0.815 2。对比多种主流模型, 基于特征光谱、GLCM 纹理、CNN 纹理特征的 ICA-DNNs 模型预测精度最高, 校正集和验证集 R^2 分别为 0.835 8 和 0.821 0。试验结果表明, 高光谱图像可快速无损地对龙眼叶片叶绿素含量分布进行检测, 可为龙眼树实时营养监测和病害早期防治提供理论依据。

关键词: 龙眼叶片; 叶绿素; 深度学习; 高光谱图像; 支持向量回归

中图分类号: S123; S667.2 **文献标识码:** A **文章编号:** 1000-1298(2018)08-0018-08

Distribution Model of Chlorophyll Content for Longan Leaves Based on Hyperspectral Imaging Technology

YUE Xuejun^{1,2} LING Kangjie^{1,2} HONG Tiansheng^{1,3} GAN Haiming^{1,2} LIU Yongxin^{1,2} WANG Linhui^{1,2}

(1. Key Laboratory of Key Technology on Agricultural Machine and Equipment, Ministry of Education, South China Agricultural University, Guangzhou 510642, China

2. College of Electronic Engineering, South China Agricultural University, Guangzhou 510642, China

3. College of Engineering, South China Agricultural University, Guangzhou 510642, China)

Abstract: Traditional hyperspectral single point detection methods of obtaining chlorophyll content of longan leaves are inefficiency, low accuracy and time consuming. Combined with the state-of-the-art deep learning technology, a distribution model of chlorophyll content for longan leaves based on convolution neural networks (CNN) and deep neural networks (DNNs) was proposed. Firstly, the spectral noise was reduced by Savitzky-Golay filter. The initial features extraction was carried out by using a principle component analysis (PCA) to identify a number of potential characteristic wavelengths (483 nm, 518 nm, 625 nm, 631 nm, 642 nm and 675 nm) according to the weight coefficient distribution curve of the first three principle component images (PC1, PC2 and PC3) under the full wavelengths. For the characteristic spectral images and the principal component images, the texture based on the gray level co-occurrence matrix was extracted from those images, and the structure information of those images was also extracted based on CNN simultaneously. Among the 300 samples, there were total of 1 800 spectral images and 900 principal component images, in which a sample corresponding to six characteristic spectrum images and three PCA images for a sample. Gray-level co-occurrence matrix (GLCM) was utilized to extract texture features. The hyperspectral wavelength feature data, texture data, images structure data and the combined data were utilized to develop particle swarm optimization-support vector regression (PSO-SVR) and independent component analysis-deep neural networks (ICA-DNNs),

收稿日期: 2018-02-01 修回日期: 2018-03-30

基金项目: 国家自然科学基金项目(30871450)、广东省科技计划项目(2015A020224036、2014A020208109)、广东省水利科技创新项目(2016-18)和广州市科技计划项目(201803020022)

作者简介: 岳学军(1971—),女,教授,博士生导师,主要从事农业工程、通信和信息技术应用研究,E-mail: yuexuejun@scau.edu.cn

通信作者: 洪添胜(1955—),男,教授,主要从事农业工程、机电一体化和信息技术应用研究,E-mail: tshong@scau.edu.cn

respectively. The particle swarm optimization (PSO) was introduced to intelligently optimize the parameters (γ and c) in the SVR model to find the optimum. Some main conclusions were obtained: the performance of PSO - SVR model based on characteristic spectrum was the best, and the coefficient of determination (R^2) of calibration set and validation set of the entire growth state were up to 0.822 0 and 0.815 2, respectively. Multi-source data fusion performance of ICA - DNNs was the best, and the precision of ICA - DNNs model was improved based on feature spectrum, texture feature and images structural feature, the R^2 of calibration set and validation set were 0.835 8 and 0.821 0, respectively. Compared with traditional methods of SVR, DNNs was more robust for larger data set. The longan leaf textures and CNN characteristics were less relevant to longan chlorophyll content distribution. Chlorophyll content distribution region for tender, pale and dark green longan leaves were: mesophyll of the leaf-root, part of mesophyll of lateral veins and the whole mesophyll. Finally, hyperspectral technology could obtain accurate chlorophyll content of longan leaves rapidly, quantitatively and non-destructively. The research result can provide a theoretical basis for nutrition surveillance of longan growth and longan disease such as leaf spot, brown spot and leaf blight.

Key words: longan leaves; chlorophyll; deep learning; hyperspectral imaging; support vector regression

0 引言

光合作用直接影响植株生长和叶绿素含量,进而影响其叶色。叶色是氮素营养诊断最常用指标^[1],也是诊断植株染病与抗病能力的辅助指标^[2],获得准确的叶色诊断指标是植株精确定量施氮施药的基础^[3]。利用计算机视觉技术识别叶色,以此鉴别农作物长势已成为研究热点与难点^[4]。龙眼具有药用价值^[5-6],是我国福建、广西、广东和海南等地支柱性经济树种。龙眼的叶斑病、褐斑病、叶枯病等能导致叶片叶绿素含量分布变化,降低光合效率,进而影响龙眼的品质与产量。龙眼急性病如急性炭疽病易引起幼苗部分叶片枯死,导致果树落花落果,该龙眼病初期病理特征不明显,但叶绿素含量却变化显著^[7]。快速无损检测龙眼叶片叶绿素含量及其分布对龙眼树的栽培管理、病害早期快速监测防治、实时精确控制水肥灌溉等具有重要意义。

传统的植物叶片叶绿素含量检测常用分光光度法和高效液相色谱法(High performance liquid chromatography, HPLC),存在耗时、费力、时效性差和损坏样本等弊端。近年来,随着高光谱技术的发展和应用的成熟,国内外研究者针对水稻、冬小麦、玉米等一体化大田植物做了大量研究,但对龙眼等个体化果树的叶绿素含量估测和分布研究甚少。岳学军等^[8]分别采用了主成分分析-支持向量机回归(Principle component analysis - support vector regression, PCA - SVR)和小波去噪-偏最小二乘回归(Wavelet denoising - partial least square regression, WD - PLSR)算法建立柑橘叶片叶绿素含量估测模型,其模型决定系数分别达到了0.871 3和0.867。KALACSKA等^[9]、ZHAO等^[10]、SHI等^[11]采用PCA、连续小波变换(Continuous wavelet transform, CWF)、多元

线性回归(Multiple linear regression, MLR)、神经网络等建模方法分析了GNDVI、mNDVI₇₀₅、MSAVI2、PSRI等10余种光谱指数与农作物叶绿素浓度的相关性,建立了高光谱图像数据对叶绿素含量估测及其分布模型。邹小波等^[12]研究了黄瓜叶片高光谱图像450~850 nm波段,建立以特征波长与对应的光谱反射值植被指数关系模型,结果表明基于695~709 nm和750~800 nm波段下的 $(R_{695-709})^{-1} - (R_{750-800})^{-1}$ 模型预测具有最优效果,其校正集和预测集相关系数分别为0.841和0.828 6。孙俊等^[13]对生菜叶选取了390~1 050 nm可见光-近红外波段的高光谱图像,分别采用特征波段和基于灰度共生矩阵纹理特征建立SVR模型,最优特征模型决定系数为0.996,均方根误差(RMSE)为0.034。

数据量大、波段多、混合像元间关系复杂是高光谱图像处理中的难点,且图像中精细光谱特征受目标光谱变异性、背景信息分布与模型假设差异性、图像亚像元小等问题的影响,传统机器学习模型无法适应这种特征,需要新的信息提取模型与方法^[14]。

本文采样地点为广州市郊华南农业大学红满堂龙眼园(23°09'N, 113°21'E),样本树为20株长势相当的龙眼树,采用HyperSIS高光谱成像仪分别采集嫩绿、浅绿、深绿共3种不同生长状态的龙眼叶高光谱数据,先进行Savitzky - Golay滤波,再获取PCA降维高光谱图像,使用卷积神经网络(CNN)提取特征,最后分别使用粒子群优化-支持向量回归(PSO - SVR)和独立成分分析-深度神经网络(ICA - DNNs)建立龙眼叶片叶绿素含量预测及其分布模型。

1 材料与方 法

1.1 样本采集与处理

选取华南农业大学红满堂龙眼园内20株长势

相当的处于采果期的龙眼树为样本,采集数据时间为2016年8月23日08:00。龙眼叶分嫩绿、浅绿和深绿3种,从每棵龙眼树冠的上、中、下3个部位和东、南、西、北4个方位分别均匀随机采集龙眼叶共15片,共300个龙眼叶样本,每种颜色100个。其中嫩绿、浅绿、深绿3种状态是指叶片SPAD均值分布区间为(9,25]、(25,41]、(41,57)的叶片,叶龄分别是7~27 d、28~48 d和49 d以上。叶片新鲜、无明显病斑和破损,将采集的样本保存于冰袋中,避免叶绿素高温见光分解。所有样本编号后立即带回实验室进行高光谱信息采集及叶绿素含量化学测定^[15]。

1.2 高光谱图像获取和标定

该实验所用的高光谱成像仪(HyperSIS, Zolix, 中国)主要由4个卤光灯、CCD相机(Gilden Photonics Ltd, UK)、样品移动台等部件组成。成像仪光谱波长范围369~988 nm,分辨率1.2 nm。高光谱成像仪控制及图像校正软件为SpectraSENS。按照文献[16]的方法采集龙眼叶片的高光谱图像。

为了防止基线漂移,需打开卤光灯,等待30 min后对300个样本进行扫描。由于光源强度分布不均匀以及暗电流噪声存在,需要在扫描生成高光谱图之前,对每个样本进行黑白标定。先将龙眼叶片平铺在移动台的白纸上,CCD相机曝光时间设置为15 ms,平台移动速度为4.5 mm/s。再次关闭相机快门,盖上CCD相机镜头扫描获取反射率为零的全黑标定图像 B ,对白板扫描获取反射率接近100%的全白标定图像 W ,设 R_{raw} 为未标定的样本图像, I_c 为标定后的高光谱图像,则 I_c 的计算公式为^[17]

$$I_c = \frac{R_{\text{raw}} - B}{W - B} \quad (1)$$

1.3 叶绿素含量测定

采集完龙眼叶片的光谱信息后,立刻使用分光光度法测定每片叶子的叶绿素感兴趣区(Region of interest, ROI)含量。在叶脉附近逆时针均匀选取6个ROI点,每个点选取0.5 g的ROI新鲜龙眼叶组织,去除叶脉,再剪成细丝,置于研钵中,加入少量石英砂、碳酸钙粉及体积分数为80%丙酮溶液2~3 mL,研磨至组织变白,于室温(25℃)下暗处静置10 min。经过滤、反复冲洗研钵和研棒,确保叶片色素全部进入容量瓶。用紫外-可见分光光度计分别测量波长为663、645、652 nm处的吸光度,最后依据叶绿素a和叶绿素b的浓度与吸光度关系计算出总的叶绿素含量^[18](表1),叶绿素含量用叶片叶绿素质量比表示。

1.4 高光谱图像预处理

1.4.1 ROI智能选取

由图1可知,在519 nm波段附近,构成龙眼叶

表1 不同叶片状态下龙眼叶片叶绿素含量统计数据
Tab.1 Descriptive statistic of chlorophyll content during different growth states mg/g

统计量	叶绿素质量比		
	嫩绿	浅绿	深绿
最小值	0.302 4	1.069 9	1.663 4
最大值	1.586 6	2.744 1	3.009 7
平均值	0.824 5	1.674 7	2.280 6
标准差	0.257 4	0.326 6	0.255 3

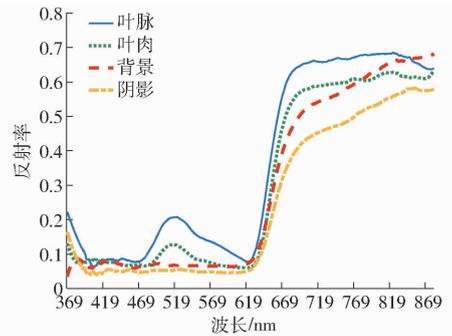


图1 叶片图像中不同区域的光谱特征图

Fig.1 Spectral characteristic of different regions in longan leaves

主体的叶脉和叶肉部分的反射率明显比背景和阴影的高。高光谱图像中,背景和阴影的反射率在波长419~619 nm波段范围内较为平稳,没有明显波峰起伏。阴影区域的反射强度低于背景区域反射强度。基于该特征,采用自动二值化去除背景和阴影部分,提取整片叶子。

1.4.2 图像去噪与特征波长选取

高光谱数据在获取过程中,往往受到仪器噪声等因素的影响,存在吸光度的非线性、基线变动和附加散射变动^[19],不可避免的产生误差,为了减少噪声对模型的干扰,采用基于时域局部多项式最小二乘法拟合的Savitzky-Golay滤波方法,具有运算速度快,保留极大值、极小值和宽度等分布特性的优点^[20]。由于高光谱图像波段多且连续,数据量大,相邻波段间存在较大的相关性,本文采用基于SVD的PCA降维后特征值所占能量比来确定主成分数目。在300个样本中,每个样本降维后选取前3个主成分图像PC1、PC2、PC3。每个样本的前3个主成分图像的总能量贡献率均大于其全部主成分图像总能量的95.8%。再根据所有样本下PC1、PC2、PC3中的各波长平均权重选取8个特征波长。由于403、821 nm处光谱噪声较大,模型中去掉这两处的光谱,对300个龙眼叶样本提取483、518、625、631、642、675 nm处共6个对应波长下的光谱图像纹理信息作为样本光谱特征,结果如图2所示。

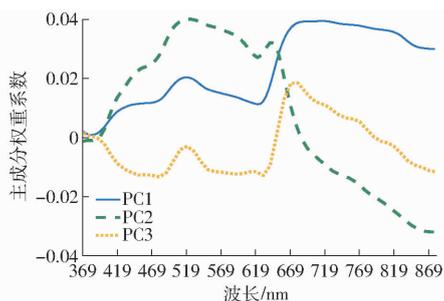


图 2 前 3 个主成分图像下各个波段平均权重系数

Fig. 2 Mean coefficient of all wavelengths for the first three principle component images

1.4.3 图像纹理信息提取

作物中叶绿素含量变化会导致生长趋势发生变化,进而影响叶片中的纹理特征^[21]。本文采用灰度共生矩阵 GLCM 提取纹理特征,利用卷积神经网络进行 CNN 纹理提取。基于 GLCM 纹理特征,提取距离参数设为 1 个像素,方向依次为 0° 、 45° 、 90° 和 135° ,分别对 ROI 区域进行包括能量、对比度、相关性、同质性^[22] 4 种纹理特征进行提取。基于预训练 AlexNet^[23] 神经网络提取主成分图像和特征波段下的高光谱图像的纹理特征。对 300 个样本分别提取主成分图像,共得 900 幅主成分图像,以此作为模型训练数据。

1.5 建模原理和方法

本文的 PSO-SVR 算法^[24-25] 是利用粒子群优化算法选取 SVR 中最优核参数 γ 和正则参数 c ,基本思想是将低维空间数据根据某种非线性变换(通过核函数实现)映射到高维特征空间,在高维特征空间内进行线性回归计算, γ 和 c 分别控制模型的偏差(Bias)和方差(Variance)^[26],对于总数为 n 的样本集,设回归模型 $f(\mathbf{x})$ 中核函数为 K ,偏置量为 d ,满足^[27]

$$f(\mathbf{x}) = \sum_{i=1}^n (\alpha_i - \alpha_i^*) K(\mathbf{x}_i, \mathbf{x}) + d \quad (0 \leq \alpha_i \leq c, 0 \leq \alpha_i^* \leq c) \quad (2)$$

其中 $K(\mathbf{x}_i, \mathbf{x}) = \exp\left(-\frac{1}{2\gamma^2} \|\mathbf{x}_i - \mathbf{x}\|^2\right) \quad (3)$

式中 α_i, α_i^* ——拉格朗日待定系数

\mathbf{x}_i ——第 i 个支持向量

\mathbf{x} ——未知样本的特征矢量

ICA-DNNs 是深度学习(Deep learning, DL)模型^[28],是近年来人工智能领域最热门的研究方向之一,其主要原理是逐层经过深度神经网络和反馈传播,自动提取数据中的特征^[29-31]。设网络权重矩阵为 \mathbf{V} ,偏置向量为 \mathbf{b} ,样本总数为 m ,网络总层数为 n_{l-1} ,假设函数 h ,网络层号为 $l(l=1, 2, \dots, n_{l-1})$,第 l 层总神经元数为 s_l ,权重衰减因子为 λ ,最小化损

失函数 J 为

$$\min J(\mathbf{V}, \mathbf{b}) = \frac{1}{m} \sum_{i=1}^m \left(\frac{1}{2} \|h_{\mathbf{V}, \mathbf{b}} \mathbf{p}_i - \mathbf{y}_i\|^2 \right) + \frac{\lambda}{2} \sum_{l=1}^{n_{l-1}} \sum_{i=1}^{s_l} \sum_{j=1}^{s_{l+1}} (\mathbf{V}_{ji}^{(l)})^2 \quad (4)$$

式中 $\mathbf{V}_{ji}^{(l)}$ ——第 $l+1$ 层网络中第 j 个神经元与第 l 层网络中第 i 个神经元间的权重矩阵

\mathbf{p}_i ——第 i 个输入样本向量

\mathbf{y}_i ——第 i 个输出向量

$h_{\mathbf{V}, \mathbf{b}}$ ——参数为 \mathbf{V} 和 \mathbf{b} 的假设函数

为了避免模型过拟合问题,本文均对目标函数进行 L1 正则化^[32] 处理。每个批次样本随机选取 80% 的数据作为校正集,用于建模,剩下 20% 作为验证集,用于评估模型的性能^[33]。建模评估采用模型决定系数 R^2 和均方根误差(RMSE)两个指标。每个模型方案独立运行 30 次,取性能指标结果的均值作为实验结果。

2 结果与分析

2.1 特征光谱建模

针对嫩绿、浅绿和深绿 3 种不同龙眼叶生长状态,分别采用 PSO-SVR 和 ICA-DNNs 进行建模实验,评估结果如表 2 所示。经过实验,PSO 种群数量设置为 20、权重衰减因子为 0.9 时,获得最佳参数性能。DNNs 层数设置为 6 层,隐藏神经元数为 20—10—10—20。从表 2 可知,由于仅在特征光谱建模下 ICA-DNNs 的验证集 R^2 均小于 0.8,故其性能较差,泛化能力较低,全期校正集和验证集的决定系数分别为 0.806 1 和 0.796 4,PSO-SVR 在 4 种不同叶片状态下均有良好表现,模型稳定性较高,嫩绿、浅绿、深绿、全期状态下的验证集的决定系数分别为 0.793 8、0.806 6、0.805 7、0.815 2。

2.2 GLCM 纹理与 CNN 纹理特征建模

对 DNNs 采用随机梯度下降(Stochastic gradient descent, SGD)法训练网络,每幅图像统一尺寸为 224 像素 \times 224 像素,经卷积(Convolution)和池化(Pooling)处理后得到特征数为 60。每个样本采集 6 幅特征光谱图像和 3 幅主成分图像,共得到特征光谱图像 1 800 幅,主成分图像 900 幅。采用 4 种纹理特征,每种特征采集方向依次为 0° 、 45° 、 90° 和 135° ,得到 16 维 GLCM 纹理, AlexNet 提取 160 维 CNN 纹理,总共有 176 维纹理特征。两种图像纹理特征分别导入 PSO-SVR 和 ICA-DNNs 模型,调节参数最优,分析结果如表 3 所示。

表2 龙眼叶在不同生长状态下的模型性能评估结果

Tab.2 Models evaluation result for longan leaves in different growth states

数据集	生长状态	PSO - SVR			ICA - DNNs		
		(γ, c)	R^2	RMSE	学习速率 α	R^2	RMSE
校正集	嫩绿	(4.980 4, 0.136 2)	0.806 1	0.181 8	0.010 0	0.789 7	0.187 1
	浅绿	(3.931 2, 0.381 4)	0.812 1	0.173 8	0.023 3	0.803 7	0.175 2
	深绿	(4.135 6, 0.503 7)	0.807 8	0.178 7	0.018 9	0.793 4	0.181 3
	全期	(4.019 2, 0.346 6)	0.822 0	0.170 9	0.032 2	0.806 1	0.171 2
验证集	嫩绿	(4.980 4, 0.136 2)	0.793 8	0.190 3	0.010 0	0.763 2	0.215 6
	浅绿	(3.931 2, 0.381 4)	0.806 6	0.180 5	0.023 3	0.783 5	0.198 3
	深绿	(4.135 6, 0.503 7)	0.805 7	0.183 3	0.018 9	0.764 0	0.209 6
	全期	(4.019 2, 0.346 6)	0.815 2	0.172 0	0.032 2	0.796 4	0.179 8

表3 基于纹理特征模型分析结果

Tab.3 Analysis results of models based on texture features

模型	校正集		验证集	
	R^2	RMSE	R^2	RMSE
PSO - SVR	0.640 5	0.231 1	0.574 2	0.288 5
ICA - DNNs	0.553 7	0.289 6	0.509 1	0.296 9

从表3可知, PSO - SVR 效果较优, 校正集和验证集模型决定系数分别为 0.640 5 和 0.574 2, 分别高于 ICA - DNNs 模型校正集和验证集模型的决定系数 0.553 7 和 0.509 1。从表3可知, 基于 GLCM 纹理与 CNN 纹理特征建模总体效果差, 说明纹理特征与叶绿素含量相关性不大。

2.3 多特征融合建模

将光谱特征数据与 GLCM 纹理特征、CNN 特征进行融合优化, 核参数与惩罚因子设置不变, 分别建立基于特征光谱与纹理特征数据融合模型, 特征光谱与 CNN 纹理特征数据融合模型, 特征光谱与 GLCM 纹理特征和 CNN 纹理信息融合模型。在训练过程中, 采用早停止方法选取最佳迭代次数, 方法如图3所示。

表4 3种融合模型对叶绿素含量的预测性能

Tab.4 Performance of three fusion models for prediction of chlorophyll content in longan leaves

模型	数据集	PSO - SVR		ICA - DNNs	
		R^2	RMSE	R^2	RMSE
特征光谱 + 特征波长图像与主成分图像 GLCM 纹理	校正集	0.816 8	0.172 4	0.819 0	0.171 2
	验证集	0.795 9	0.178 6	0.804 3	0.175 5
特征光谱 + 特征波长图像与主成分图像 CNN 纹理	校正集	0.778 5	0.225 1	0.812 5	0.174 1
	验证集	0.750 1	0.239 3	0.794 2	0.186 8
特征光谱 + 特征波长图像与主成分图像 GLCM 纹理 + 特征波长图像与主成分图像 CNN 纹理	校正集	0.807 6	0.176 2	0.835 8	0.162 1
	验证集	0.782 3	0.190 7	0.821 0	0.168 6

3种融合特征在 ICA - DNNs 模型中效果最优, 此时校正集和验证集 R^2 分别为 0.835 8 和 0.821 0, RMSE 分别为 0.162 1 和 0.168 6。在 PSO - SVR 中, 特征光谱与 GLCM 纹理融合模型效果最优, 与 2.1 节仅使用特征光谱建模相比, 基于特征光谱与

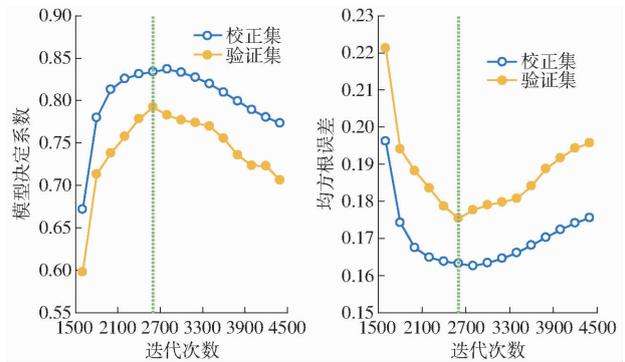


图3 基于早停止方法的 DNNs 网络训练迭代次数与模型性能评估

Fig.3 Model evaluation with number of epochs based on early stopping method

由图3可知, 采用早停止方法选取迭代次数为 2600 次时, 验证集的模型决定系数达到最大, 均方根误差达到最小, 模型性能综合最优。将 3 种融合特征分别导入 PSO - SVR 和 ICA - DNNs 模型中, 每次训练都采用图3中早停止方法选取最佳迭代次数, 将每种融合特征对应的每个模型的 30 次性能指标输出取平均, 得多特征融合的结果如表4所示。

GLCM 纹理、特征光谱与 CNN 纹理特征的建模性能有所下降, 可能是由于龙眼叶在不同叶绿素含量分布下, 叶片状态与结构发生改变性不明显, 图像纹理特征与叶绿素含量相关性较差, 相当于引入部分不相关信息, 加上 PSO - SVR 整体迁移融合能力不强,

融合模型变差。龙眼叶片叶绿素含量与特征光谱、GLCM 纹理特征、CNN 纹理信息特征相关性依次递减。ICA - DNNs 在融合包括特征光谱、图像纹理特征等多种数据信息下，模型稳定性得到了提高，泛化能力得到了一定的提升。

2.4 龙眼叶叶绿素含量验证与分布反演

根据上述结论，建立基于多源特征融合的 ICA - DNNs 模型预测龙眼叶叶绿素分布，将每个像素点对应的高光谱曲线全部导入 ICA - DNNs 模型，每个像素点和与它相关点的高光谱信息可预测该点的叶绿素含量，进而得到整块叶片的叶绿素分布，分布结果如图 4、5 所示。

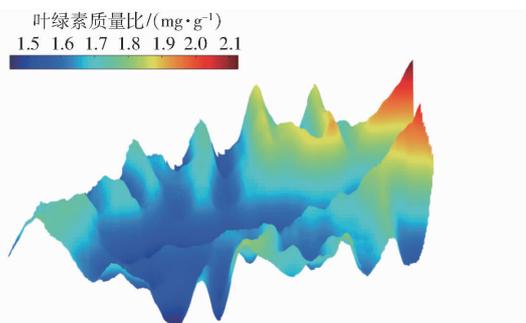


图 4 基于最优参数的 ICA - DNNs 龙眼叶绿素含量分布
Fig. 4 Image of chlorophyll content distribution for longan leaves based on ICA - DNNs with optimum parameters

由图 5 可知，嫩绿叶片叶绿素含量主要分布于靠近叶根叶肉部分。浅绿叶片叶绿素含量开始从叶根扩展至侧脉附近，叶子脉络清晰，叶绿素平均含量增大。叶片为深绿时，叶脉、叶肉均有比较大的叶绿素含量，分布均匀。主脉叶绿素含量在 3 种叶片状态下均比较低。

为了验证 ICA - DNNs 模型的可靠性，随机选取

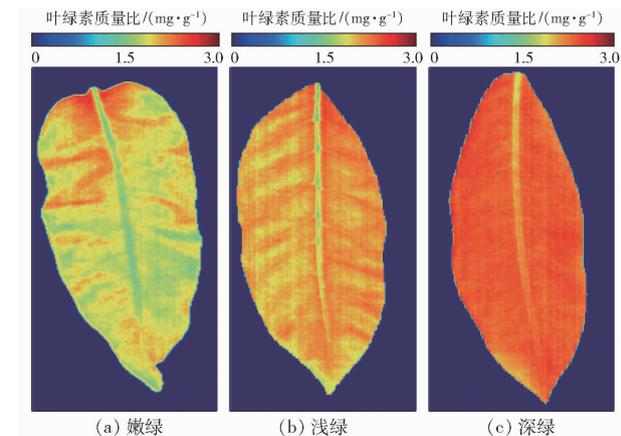
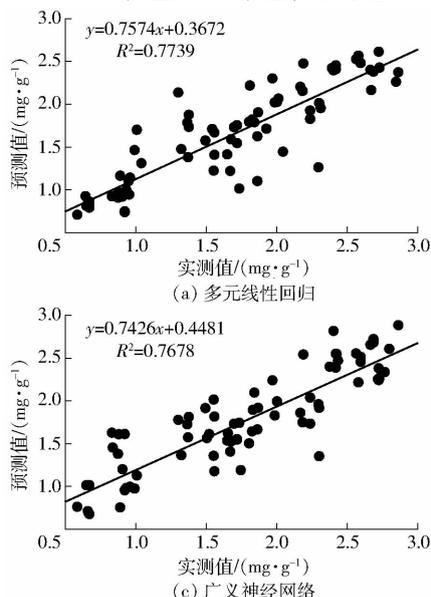


图 5 3 种不同生长状态下龙眼叶片叶绿素含量分布
Fig. 5 Distribution of chlorophyll content in three different growth states

每种状态叶片 30 个，共 90 个，每个样本随机选择 2 个点，用 TYS - B 型叶绿素测定仪测定叶片叶绿素相对含量 (SPAD)，共得到 180 组测量值，同时由 ICA - DNNs 叶绿素反演模型得到预测值，SPAD 值与预测值建立拟合关系，得回归方程 $y = 0.0534x - 0.4606$, $R^2 = 0.9172$ ，说明该模型具有可靠性和准确性。

2.5 其他模型建模对比分析

本实验将最优的基于多源特征融合的 ICA - DNNs 与多元线性回归 (Multiple linear regression, MLR)、偏最小二乘法回归 (Partial least square regression, PLSR)、广义神经网络 (General regression neural networks, GRNN)、最小邻域回归 (Nearest neighbors regression, NNR) 模型进行对比，将融合后的光谱特征导入这 4 个模型，模型参数分别调到最优，其训练结果如图 6 所示。

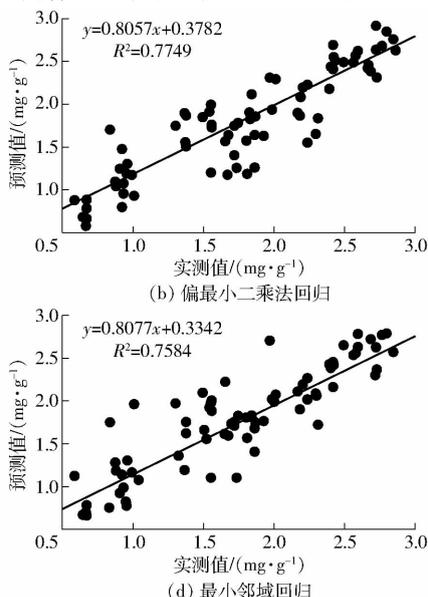


图 6 不同建模方法下叶绿素含量实测值和预测值的比较

Fig. 6 Comparison of real and predicted chlorophyll contents by different modeling methods

将这4种模型建模结果与表4相比较,得到相对ICA-DNNs模型性能百分比,结果如表5所示。这4种传统模型的实验结果表明,本文提出的ICA-DNNs模型效果最优,校正集和验证集 R^2 分别为0.8358和0.8210。其次取得较好效果的是偏最小二乘法回归模型,训练集上效果最差模型为最小邻域回归模型,验证集模型决定系数比ICA-DNNs性能下降8.15%。这4个模型均有不同程度的预测偏离数据点,主要是因为建模中涉及模型的偏差-方差(Bias-variance)矛盾,模型复杂度不够,提取到的有效特征不足或与原始数据并不完全相关,从而影响了模型的鲁棒性和泛化能力。ICA-DNNs模型能够较好进行多源特征融合,因此ICA-DNNs模型具有较高的预测精度和泛化能力。

表5 不同建模方法下的模型性能分析

Tab.5 Results of performance based on different modeling methods

模型	验证集		相对ICA-DNNs 模型性能百分比/%
	R^2	RMSE	
MLR	0.7546	0.1416	-8.09
PLSR	0.7667	0.1714	-6.61
GRNN	0.7354	0.1553	-10.42
NNR	0.7541	0.1509	-8.15

3 结论

(1)基于特征光谱的PSO-SVR模型预测效果最佳,嫩绿、浅绿、深绿和全期的验证集模型决定系

数分别为0.7938、0.8066、0.8057、0.8152。

(2)基于纹理特征的模型性能次之,纹理信息与叶绿素含量之间相关性较低,与ICA-DNNs模型相比,PSO-SVR模型性能较优,校正集和验证集模型决定系数分别为0.6405和0.5742。

(3)在多特征融合建模中,特征光谱与GLCM纹理融合在PSO-SVR模型中效果最优,校正集和验证集 R^2 分别为0.8168和0.7959, RMSE分别为0.1724和0.1786。对比MLR、PLSR、GRNN、NNR传统模型,融合3种特征在ICA-DNNs模型中效果最优,校正集和验证集 R^2 分别为0.8358和0.8210, RMSE分别为0.1621和0.1686。ICA-DNNs在融合包括特征光谱、图像纹理特征等多种数据信息下,模型稳定性得到提高,泛化能力得到了一定的提升。对比PSO-SVM,深度学习模型ICA-DNNs计算复杂度较高,但具有较强的特征提取和融合能力,适合大批量数据。在多特征融合中,该模型具有较强稳定性和泛化能力。深度模型的精度比单一传统简单模型更高。

(4)从嫩绿、浅绿到深绿的龙眼叶叶绿素含量分布规律是从局部到整体,从不均匀到均匀。嫩绿叶片的叶绿素主要分布于靠近叶根的叶肉部分,叶绿素平均含量较小。浅绿叶片叶绿素开始从叶根扩展至侧脉附近,叶绿素平均含量增大。深绿叶片的侧脉、叶肉分布均匀,叶绿素平均含量进一步增大。主脉叶绿素含量在嫩绿、浅绿和深绿这3种叶子状态下均较低。

参 考 文 献

- 张金恒,王珂,王人潮. 高光谱评价植被叶绿素含量的研究进展[J]. 上海交通大学学报(农业科学版), 2003, 21(1): 74-80. ZHANG Jinheng, WANG Ke, WANG Renchao. Study on hyperspectral remote sensing in estimate vegetation leaf chlorophyll content[J]. Journal of Shanghai Jiaotong University(Agricultural Science), 2003, 21(1): 74-80. (in Chinese)
- 王刚,侍瑞,努尔尼萨,等. 氮磷钾施肥对比对骏枣叶绿素荧光参数的影响[J]. 西北农林科技大学学报(自然科学版), 2012, 40(9): 187-194. WANG Gang, SHI Rui, NUER Ni-sa, et al. Effect of different ratios of N, P and K fertilizer on the fluorescence parameters of junjube[J]. Journal of Northwest A&F University(Natural Science Edition), 2012, 40(9): 187-194. (in Chinese)
- 弓建国,徐松鹤. 氮磷钾有机肥对马铃薯叶绿素含量的影响[J]. 安徽农业科学, 2009, 37(23): 10940-10942. GONG Jian'guo, XU Songhe. Effect of nitrogen phosphorus and potassium fertilizers and organic fertilizer on chlorophyll content in potato[J]. Journal of Anhui Agricultural Sciences, 2009, 37(23): 10940-10942. (in Chinese)
- CAVALCANTE I H L, DA SILVA G B, SANTOS E M, et al. Relationship between chlorophyll meter readings and leaf nitrogen concentration in papaya (*Carica papaya* L.)[J]. Philippine Journal of Crop Science, 2016, 41(1): 75-79.
- 黄美香,林河通,李辉,等. 龙眼果实加工技术研究进展[J]. 包装与食品机械, 2012, 30(2): 48-52. HUANG Meixiang, LIN Hetong, LI Hui, et al. Advances in the researches of processing technologies of longan fruits[J]. Packaging and Food Machinery, 2012, 30(2): 48-52. (in Chinese)
- 康由发. 氯酸钾对龙眼催花的效应及其机理研究综述[J]. 华南热带农业大学学报, 2006, 12(4): 23-26. KANG Youfa. Research review on the effects and mechanism of potassium chlorate on inducing flowering in longan[J]. Journal of South China University of Tropical Agriculture, 2006, 12(4): 23-26. (in Chinese)
- 伍南,刘君昂,周国英,等. 基于可见-近红外光谱技术的炭疽病侵染后油茶叶片叶绿素含量预测研究[J]. 光谱学与光谱分析, 2012, 32(5): 1221-1224. WU Nan, LIU Jun'ang, ZHOU Guoying, et al. Prediction of chlorophyll content of leaves of oil camlliae after being indected with anthracnose based on Vis/NIR spectroscopy[J]. Spectroscopy and Spectral Analysis, 2012, 32(5): 1221-1224. (in Chinese)
- 岳学军,全东平,洪添胜,等. 柑橘叶片叶绿素含量高光谱无损检测模型[J]. 农业工程学报, 2015, 31(1): 294-302. YUE Xuejun, QUAN Dongping, HONG Tiansheng, et al. Non-destructive hyperspectral measurement model of chlorophyll content for citrus leaves[J]. Transactions of the CSAE, 2015, 31(1): 294-302. (in Chinese)
- KALACSKA M, LALONDE M, MOORE T R. Estimation of foliar chlorophyll and nitrogen content in an ombrotrophic bog from

- hyperspectral data: scaling from leaf to image[J]. *Remote Sensing of Environment*, 2015, 169: 270 – 279.
- 10 ZHAO Jiewen, WANG Kailiang, QIN Ouyang, et al. Measurement of chlorophyll content and distribution in tea plant's leaf using hyperspectral imaging technique[J]. *Spectroscopy and Spectral Analysis*, 2011, 31(2): 512 – 515.
- 11 SHI Jiyong, ZOU Xiaobo, ZHAO Jiewen, et al. Nondestructive diagnostics of nitrogen deficiency by cucumber leaf chlorophyll distribution map based on near infrared hyperspectral imaging[J]. *Scientia Horticulturae*, 2012, 138:190 – 197.
- 12 邹小波, 张小磊, 石吉勇, 等. 基于高光谱图像的黄瓜叶片叶绿素含量分布检测[J]. *农业工程学报*, 2014, 30(13): 169 – 175.
- ZOU Xiaobo, ZHANG Xiaolei, SHI Jiyong, et al. Detection of chlorophyll content distribution in cucumber leaves based on hyperspectral imaging[J]. *Transactions of the CSAE*, 2014, 30(13): 169 – 175. (in Chinese)
- 13 孙俊, 金夏明, 毛罕平, 等. 基于高光谱图像光谱与纹理信息的生菜氮素含量检测[J]. *农业工程学报*, 2014, 30(10): 167 – 173.
- SUN Jun, JIN Xiaming, MAO Hanping, et al. Detection of nitrogen content in lettuce leaves based on spectroscopy and texture using hyperspectral imaging technology[J]. *Transactions of the CSAE*, 2014, 30(10): 167 – 173. (in Chinese)
- 14 张兵. 高光谱图像处理与信息提取前沿[J]. *遥感学报*, 2016, 20(5): 1062 – 1090.
- ZHANG Bing. Advancement of hyperspectral image processing and information extraction[J]. *Journal of Remote Sensing*, 2016, 20(5): 1062 – 1090. (in Chinese)
- 15 王伟, 彭彦昆, 马伟, 等. 冬小麦叶绿素含量高光谱检测技术[J]. *农业机械学报*, 2010, 41(5): 172 – 177.
- WANG Wei, PENG Yankun, MA Wei, et al. Prediction of chlorophyll content of winter wheat using leaf-level hyperspectral data [J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2010, 41(5): 172 – 177. (in Chinese)
- 16 邹小波, 陈正伟, 石吉勇, 等. 基于近红外高光谱图像的黄瓜叶片色素含量快速检测[J/OL]. *农业机械学报*, 2012, 43(5): 152 – 156. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=20120526&journal_id=jcsam. DOI: 10.6041/j.issn.1000-1298.2012.05.026.
- ZOU Xiaobo, CHEN Zhengwei, SHI Jiyong, et al. Rapid detection of cucumber leaves pigments based on near infrared hyperspectral image technology[J/OL]. *Transactions of the Chinese Society for Agricultural Machinery*, 2012, 43(5): 152 – 156. (in Chinese)
- 17 孙俊, 卫爱国, 毛罕平, 等. 基于高光谱图像及 ELM 的生菜叶片氮素水平定性分析[J/OL]. *农业机械学报*, 2014, 45(7): 272 – 277. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=20140742&journal_id=jcsam. DOI: 10.6041/j.issn.1000-1298.2014.07.042.
- SUN Jun, WEI Aiguo, MAO Hanping, et al. Discrimination of lettuce leaves' nitrogen status based on hyperspectral imaging technology and ELM[J/OL]. *Transactions of the Chinese Society for Agricultural Machinery*, 2014, 45(7): 272 – 276. (in Chinese)
- 18 李民赞. 光谱分析技术及其应用[M]. 北京: 科学出版社, 2006: 45 – 48.
- 19 DAUGHTRY C S T, WAHALL C L, KIM M S, et al. Estimating corn leaf chlorophyll concentration from leaf and canopy reflectance[J]. *Remote Sensing of Environment*, 2000, 74(2): 229 – 239.
- 20 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 Tiansheng, et al. Early detection and grading of citrus Huanglongbing using hyperspectral imaging technique[J]. *Transactions of the CSAE*, 2014, 30(9): 140 – 147. (in Chinese)
- 21 岳学军, 全东平, 洪添胜, 等. 基于流形学习算法的柑橘叶片氮含量光谱估测模型[J/OL]. *农业机械学报*, 2015, 46(6): 244 – 250. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=20150635&journal_id=jcsam. DOI: 10.6041/j.issn.1000-1298.2015.06.035.
- YUE Xuejun, QUAN Dongping, HONG Tiansheng et al. Estimation model of nitrogen content for citrus leaves by spectral technology based on manifold learning algorithm[J/OL]. *Transactions of the Chinese Society for Agricultural Machinery*, 2015, 46(6): 244 – 250. (in Chinese)
- 22 HUANG Xin, LIU Xiaobo, ZHANG Liangpei. A multichannel gray level co-occurrence matrix for multi/hyperspectral image texture representation[J]. *Remote Sensing*, 2014, 6(9): 8424 – 8445.
- 23 KRIZHEVSKY A, SUTSKEVER I, HINTON G E. Imagenet classification with deep convolutional neural networks [J]. *Communication of the ACM*, 2017, 60(6): 84 – 90.
- 24 黄双萍, 洪添胜, 岳学军, 等. 基于高光谱的柑橘叶片氮素含量多元回归分析[J]. *农业工程学报*, 2013, 29(5): 132 – 138.
- HUANG Shuangping, HONG Tiansheng, YUE Xuejun, et al. Multiple regression analysis of citrus leaf nitrogen content using hyperspectral technology[J]. *Transactions of the CSAE*, 2013, 29(5): 132 – 138. (in Chinese)
- 25 XUE Zhaohui, DU Peijun, SU Hongjun. Harmonic analysis for hyperspectral image classification integrated with PSO optimized SVM[J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2014, 7(6): 2131 – 2146.
- 26 梁亮, 杨敏华, 张连蓬, 等. 基于 SVR 算法的小麦冠层叶绿素含量高光谱反演[J]. *农业工程学报*, 2012, 28(20): 162 – 171.
- LIANG Liang, YANG Minhua, ZHANG Lianpeng, et al. Chlorophyll content inversion with hyperspectral technology for wheat canopy based on support vector regression algorithm[J]. *Transactions of the CSAE*, 2012, 28(20): 162 – 171. (in Chinese)
- 27 HINTON G E, SALAKHUTDINOV R R. Reducing the dimensionality of data with neural networks[J]. *Science*, 2006, 313(5786): 504 – 507.
- 28 LECUN Y, BENGIO Y, HINTON G. Deep learning[J]. *Nature*, 2015, 521(7553): 436 – 444.
- 29 HINTON G E, DENG L, YU D, et al. Deep neural networks for acoustic modeling in speech recognition[J]. *IEEE Signal Processing Magazine*, 2012, 29(6): 82 – 97.
- 30 HINTON G E, OSINDERO S, TEH Y W. A fast learning algorithm for deep belief nets[J]. *Neural Computation*, 2006, 18(7): 1527 – 1554.
- 31 YU D, HINTON G, MORGAN N, et al. Introduction to the special section on deep learning for speech and language processing [J]. *IEEE Transactions on Audio, Speech, and Language Processing*, 2012, 20(1): 4 – 6.
- 32 LEE H, GROSSE R, RANGANATH R, et al. Unsupervised learning of hierarchical representations with convolutional deep belief networks[J]. *Communications of the ACM*, 2011, 54(10): 95 – 103.
- 33 SHIN H C, ROTH H R, GAO M C, et al. Deep convolutional neural networks for computer-aided detection: cnn architectures, dataset characteristics and transfer learning[J]. *IEEE Transactions on Medical Imaging*, 2016, 35(5): 1285 – 1298.