

基于深度信念网络的多品种水稻生物量无损检测

段凌凤¹ 潘井旭¹ 郭子龙² 刘海北¹ 覃建祥¹ 柯希鹏¹

(1. 华中农业大学工学院, 武汉 430070; 2. 华中农业大学作物遗传改良国家重点实验室, 武汉 430070)

摘要: 提出了基于深度信念网络的多品种生殖生育期水稻生物量无损检测方法。对在正常生长及干旱胁迫两个不同环境下的 483 个水稻品种, 分别于胁迫前、胁迫后和复水后 3 个时间点进行图像采集。利用 HSL 颜色空间固定阈值分割法分割图像, 并对处理后的图像进行特征提取, 共提取 57 个特征值。对数据进行归一化处理后, 构建基于深度信念网络的水稻生物量模型, 根据决定系数 R^2 、平均相对误差 (MAPE) 及相对误差绝对值的标准差 (SAPE) 选择最优模型, 并与逐步线性回归模型进行比较。结果表明, 基于深度信念网络的生物量测量模型性能更优, R^2 为 0.929 9, MAPE 为 11.19%, SAPE 为 18.36%。本研究提供了一种精度高且适用于多品种、不同生殖生育期、不同生长环境的水稻生物量无损检测模型, 为水稻研究提供了新的测量工具。

关键词: 水稻; 生物量; 深度信念网络; 多品种; 生殖生育期; 无损检测

中图分类号: S126; TP391.41 **文献标识码:** A **文章编号:** 1000-1298(2019)11-0136-08

Non-destructive Measurement of Rice Biomass Based on Deep Belief Network

DUAN Lingfeng¹ PAN Jingxu¹ GUO Zilong² LIU Haibei¹ QIN Jianxiang¹ KE Xipeng¹

(1. College of Engineering, Huazhong Agricultural University, Wuhan 430070, China

2. National Key Laboratory of Crop Genetic Improvement, Huazhong Agricultural University, Wuhan 430070, China)

Abstract: Rice is one of the most significant food crops all over the world. Biomass is a key phenotypic trait in rice research. A new method for nondestructive detection of rice biomass for multiple varieties at reproductive stage based on deep belief network was proposed. RGB images of 483 different rice varieties under normal growth environment and drought stress environment were captured at three time points: before stress, after stress and after rehydration. After image acquisition, the images were segmented by using fixed threshold in HSL color space and 57 image-derived features related to rice biomass were extracted. After data normalization, a rice biomass model was built based on deep belief network. The influences of visible layer type, hidden layer number, hidden layer neuron number, learning rate, epoch number and momentum on the performance of deep belief network were tested. The best model was selected based on the coefficient of determination (R^2), mean absolute percent error (MAPE) and standard deviation of absolute percent error (SAPE). The deep belief network model was also compared with the stepwise linear regression model. The results showed that the biomass measurement model based on the deep belief network performed better (R^2 was 0.929 9, MAPE was 11.19% and SAPE was 18.36%). The research offered a new nondestructive method for accurately measuring rice biomass for multiple varieties under different growth environments, which would provide a new tool for rice research.

Key words: rice; biomass; deep belief network; multiple varieties; reproductive stage; nondestructive measurement

0 引言

水稻是世界上最主要的粮食作物之一, 高产一

直是水稻育种与栽培的重要目标^[1-2]。水稻生长发育中的生物量积累是经济产量(以下简称产量)的物质基础^[3-4]。不同品种生物量积累动态规律^[5-6]

收稿日期: 2019-03-04 修回日期: 2019-08-09

基金项目: 国家自然科学基金项目(31701317)、湖北省自然科学基金项目(2017CFB208)和国家级大学生创新创业训练计划项目(201810504075)

作者简介: 段凌凤(1986—), 女, 副教授, 博士, 主要从事植物表型组学研究, E-mail: duanlingfeng@mail.hzau.edu.cn

及各生育阶段生物量与产量的关系不一致^[7],连续测量不同品种的水稻生物量,对研究不同品种生物量积累动态和水稻产量形成规律、指导水稻生产及育种具有现实意义。

近年来,已有学者在基于图像特征的作物生物量测量模型构建上进行了探索研究。梁淑敏等^[7]以植株图像周长为表征因子,建立了玉米鲜生物量测量模型。更多研究则基于植株不同角度下投影面积构建生物量测量模型,如基于顶视投影面积^[8]、基于 2 幅互呈 90°的侧视投影面积^[9]、基于顶视投影面积与 2 幅互呈 90°侧视投影面积^[10-13],以及基于多幅侧视投影面积平均值与顶视投影面积^[14-15]等。研究表明,仅基于植株投影面积无法建立准确有效的多品种全生育期生物量测量模型^[13]。通过加入其他表征因子,如绿色比^[16]、株龄^[17]、株高及分蘖数^[18]等,可提高生物量测量模型的精度。然而,这些研究的对象大多为处于营养生长早期的少量品种,构建的模型品种及生育期适应性较差。另外,上述研究均基于传统的回归分析建模,不足以描述图像特征与生物量之间复杂的非线性关系。

自从 HINTON 等^[19]开创性地提出深度信念网络以来,基于大数据驱动的深度学习方法以其优越的性能及充足的生理学基础而广泛地应用于各个领域^[20-26]。深度学习的优势在于可以直接将原始数据输入到模型中,而不需要单独对特征进行选择与变换,由模型通过学习给出合适的特征表示^[27]。从理论上来说,将原始图像输入深度学习网络即可得到很好的检测结果,但本研究中的原始图像尺寸非常大(2 452 像素×2 056 像素),现有网络无法处理,若直接压缩则会造成很多图像细节丢失。另外,即使压缩到适合现有网络的图像尺寸(如 224 像素×224 像素),对于这么多的参数输入,需要更多样本(数十万甚至数百万)才能得到较好的效果。因此,本研究采用先提取图像特征、再建模的方式。深度信念网络(Deep belief network, DBN)首先对网络进行无监督预训练,然后通过反向传播(Back propagation, BP)算法对网络进行微调,尤其适用于有标签样本的数量有限情况下的学习问题。深度信念网络具备层次化特征学习与表达的能力,在探索分析输入、输出间复杂非线性关系上具有独特优势^[28]。

图像特征和水稻生物量间具有复杂的非线性关系,深度信念网络是解决这一问题的有效工具。本研究使用深度信念网络进行生物量建模的另一个重要原因在于,研究中有大量无标签样本,而深度信念网络能充分利用无标签样本提高模型性能,是研究

水稻生物量测量模型的有力工具。本研究首先通过图像分析从原始水稻图像中提取生物量相关图像特征,然后引入深度信念网络,构建适用于多品种水稻生殖生育期的水稻生物量模型。

1 材料与方法

1.1 成像系统

本研究中的图像采集装置如图 1 所示。水稻植株由工业输送线自动输送至检测区域,电动旋转台带动水稻旋转一周,由彩色工业相机(AVT Stingray F-504)采集水稻图像,图像尺寸为 2 452 像素×2 056 像素。每盆植株每次采集 14 个不同角度下的可见光图像。采集到的图像将自动传输至工作站。

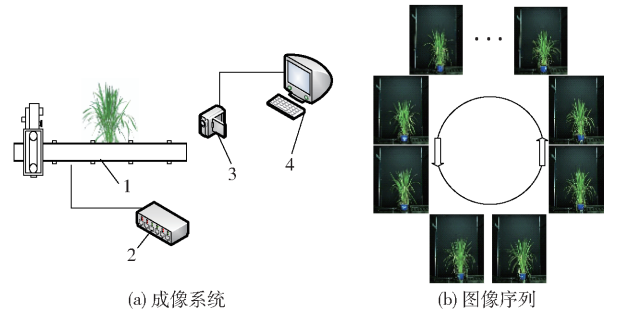


图 1 图像采集装置

Fig. 1 Image acquisition equipment

1. 旋转台 2. 控制器 3. 相机 4. 工作站

1.2 试验材料与图像采集

本研究的试验材料为存在广泛自然变异的 483 份水稻核心种质资源。样本种植于直径为 22 cm 的塑料盆中。试验设置对照组(N)和干旱胁迫组(D),每个组设置 4 个重复,共 483×2×4=3 864 个样本。对照组正常浇水,种植于露天环境下,而胁迫组种植在温室中,于孕穗期进行干旱胁迫。

当水稻样本长至孕穗期时,对胁迫组所有品种的 4 个重复采集图像。随后,胁迫组断水进行胁迫。采用 TRIME-PICO32 型土壤水分测量仪(IMKO Micromoduletechnik GmbH, Ettlingen, 德国)测量土壤含水率。当土壤含水率降至 15% 时,给胁迫组浇水,使土壤含水率保持在 15% 的水平 5 d。再次采集胁迫组和对照组所有样本的水稻图像,并对胁迫组和对照组所有品种的两个重复进行有损人工测量获取水稻地上部分干物质量。复水后对胁迫组和对照组正常浇水,于成熟期进行第 3 次水稻图像采集,并对胁迫组和对照组所有品种的两个重复进行有损人工测量获取水稻地上部分干物质量。由于部分样本在测量中出现人为失误,最终具有人工测量干物质量数据的有标签样本为 3 811 个,仅有图像而无人工测量干物质量数据的无标签样本为 3 829 个。

1.3 图像处理及性状提取

1.3.1 图像处理

图像采集完成后进行处理,主要步骤如下:①图像阈值分割。首先将图像转换到 HSL 颜色空间,利用 HSL 颜色空间固定阈值分割法分割图像,提取图像的 L 分量,设定阈值为 100。去除小区域后得到水稻植株二值图像。②提取颜色分量。利用二值图像和原始 RGB 图像进行掩模操作,获取植株的 RGB 图像,提取水稻植株图像的超绿分量(G_{Ex})和超红分量(R_{Ex}),若某像素的 G_{Ex} 值大于预定义 G_{Ex} 阈值, R_{Ex} 值小于预定义 R_{Ex} 阈值,则该像素被分割为植株绿色部分。提取 i_2 组分,测定水稻植株的黄色部分。各颜色分量的计算公式为

$$r = \frac{R}{R + G + B}$$
 (1)

$$g = \frac{G}{R + G + B}$$
 (2)

$$b = \frac{B}{R + G + B}$$
 (3)

$$G_{Ex} = 2g - r - b$$
 (4)

$$R_{Ex} = 1.4r - b$$
 (5)

$$i_2 = 0.5R - 0.5B$$
 (6)

式中 $R、G、B$ ——彩色 RGB 图像的 $R、G、B$ 分量
 $r、g、b$ ——归一化 $R、G、B$ 分量

③检测水稻植株边缘和最小外接矩形。利用 IMAQ EdgeDetection VI 提取水稻植株边缘,并利用 IMAQ Particle Analysis VI 检测水稻植株最小外接矩形(图 2)。

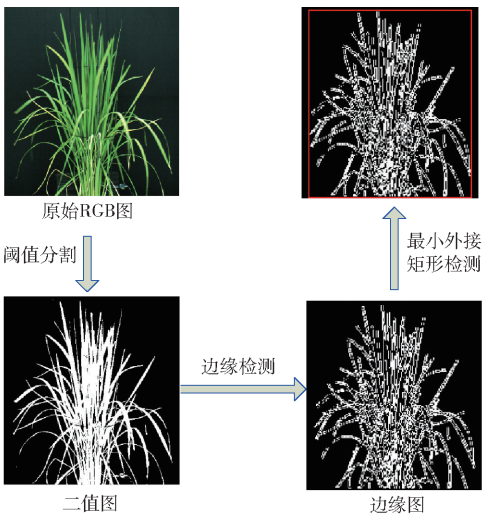


图 2 水稻植株边缘和最小外接矩形检测
Fig.2 Detection of edge and border of rice plants

1.3.2 图像特征提取

基于处理好的水稻图像,对每个植株共提取 57 个图像特征,分别为:

株高(H_p):将水稻拉直后测量得到的高度。

分蘖数(N_T):分蘖的数目。
绿色投影面积(A_G):植株绿色部分的像素数。
茎秆面积(A_{ST}):茎秆像素数。
植株投影面积(A):植株像素数。
绿色比(R_{GLA}):绿色投影面积与植株投影面积的比值。

植株密度($C_1 \sim C_6$):将图像分为若干个尺寸为 5 像素 $\times$ 5 像素的子图像,计算每个子图像内的前景像素比例。像素点的密度定义为该像素点所在子图像中的前景像素数占子图像总像素数的比例。对植株 180° 范围的图像(本研究中共 7 幅)进行同样处理,以这些图像得到的密度均值作为植株密度。根据子图像内的前景像素比例,将子图像划分为 6 个不同水平,分别为水平 1(0 ~ 10%)、水平 2(10% ~ 20%)、水平 3(20% ~ 40%)、水平 4(40% ~ 60%)、水平 5(60% ~ 80%)、水平 6(80% ~ 100%)。统计不同前景像素比例水平中所含子图像数占所有子图像数的比例,计算得到 6 个不同植株密度水平,分别记为 $C_1、C_2、C_3、C_4、C_5、C_6$ [29]。

周长面积比(R_{PA}):植株周长与植株投影面积之比,其中周长为植株边缘像素的数目。

分形维数 1(D_{F1}):由植株原始图像计算得到的分形维数。

植株外接矩形高(H):植株最小外接矩形的高度。

植株外接矩形宽(W):植株最小外接矩形的宽度。

植株占空比(R_{PB}):植株面积与其最小外接矩形面积的比值。

高宽比(R_{HW}):植株外接矩形高与植株外接矩形宽的比值。

分形维数 2(D_{F2}):以植株最小外接矩形为边界裁剪图像后计算的分形维数。

深绿色面积(A_{DG}):植株深绿色部分的前景像素数。

植株相对频数($F_1 \sim F_{14}$):本文共采集了植株 360° 旋转范围内的 14 幅图像,每幅侧视图对应不同的植株角度,侧视图中相同的像素点位置在不同的侧视角度下对应着不同的植株位置。统计侧视图像中每个像素点处出现前景像素(植株像素)的次数。本实验共 14 幅侧视图,即前景像素出现次数为 0 ~ 14。不考虑前景像素出现次数为 0 的像素数,统计图像中出现 i 次前景像素的像素点的个数,根据统计的个数,计算出出现 i 次前景像素的相对频率,得到 14 个分布频率特征,记为特征 $F_1、F_2、F_3、F_4、F_5、F_6、F_7、F_8、F_9、F_{10}、F_{11}、F_{12}、F_{13}、F_{14}$ [30]。

浅绿色面积(A_{LG}):植株浅绿色部分的前景像素数。

绿色等级(G_c):植株颜色等级分为深绿色和浅绿色,分别用 1 和 0 表示。当植株深绿色面积大于浅绿色面积时为 1,否则为 0。

直方图特征:均值 M 、标准差 S 、三阶矩 M_{U3} 、平滑度 R 、熵 E 、一致性 $U^{[29]}$ 。

灰度-梯度共生矩阵特征 $T_1 \sim T_{15}$: T_1 、 T_2 、 T_3 、 T_4 、 T_5 、 T_6 、 T_7 、 T_8 、 T_9 、 T_{10} 、 T_{11} 、 T_{12} 、 T_{13} 、 T_{14} 、 T_{15} 分别为相关、小梯度优势、大梯度优势、能量、灰度分布不均匀性、梯度分布不均匀性、灰度平均值、梯度平均值、灰度熵、梯度熵、混合熵、差分矩、逆差分矩、灰度均方差、梯度均方差^[30]。

表 1 为各时期 57 个特征的平均值及各特征与生物量之间的相关性。从表中可以看出,大多数特征与生物量之间具有极显著的相关性。

2 水稻生物量无损检测模型构建

2.1 数据预处理及性能评价

本研究中有标签样本 3 811 个,无标签样本 3 829 个,深度信念网络能充分利用无标签样本提高模型性能,是研究水稻生物量测量模型的有效工具。将有标签样本随机划分为训练集和测试集,其中训练样本 3 000 个,测试样本 811 个。为了消除不同水稻生物量特征数据量纲的差异,提高训练收敛速度,利用 Matlab 工具箱中的 mapminmax 函数将数据归一化至[0,1]区间。

本文采用决定系数 R^2 (Coefficient of determination)、平均相对误差 (Mean absolute percent error, MAPE)、相对误差绝对值的标准差 (Standard deviation of absolute percent error, SAPE) 来判断所构建模型的性能。计算公式为

$$V_{APE_i} = \left| \frac{y_{ai} - y_{mi}}{y_{mi}} \right| \times 100\% \tag{7}$$

$$V_{MAPE} = \frac{1}{n} \sum_{i=1}^n V_{APE_i} \tag{8}$$

$$V_{SAPE} = \sqrt{\frac{\sum_{i=1}^n (V_{APE_i} - V_{MAPE})^2}{n - 1}} \times 100\% \tag{9}$$

式中 V_{APE_i} ——第 i 个样本的相对误差绝对值
 V_{MAPE} ——平均相对误差
 V_{SAPE} ——相对误差绝对值的标准差
 y_{ai} ——第 i 个样本的生物量系统测量值
 y_{mi} ——第 i 个样本的生物量人工测量值
 n ——样本数量

2.2 基于多元线性回归的水稻生物量模型构建

通过逐步线性回归法对水稻生物量数据进行回归分析,并通过方差膨胀因子 (Variance inflation factor, VIF) 进行共线性检验,当某个回归变量的 VIF 小于 10 时,认为建立的回归模型无显著的共线性。最终从 57 个特征值中筛选出 4 个特征值,并利用 SPSS 对其进行共线性检验,确定无共线性。最终构建的逐步线性回归模型方程为

$$Y = 0.072 + 1.185X_1 - 0.224X_2 - 0.270X_3 - 0.224X_4 \tag{10}$$

式中 Y ——植株干物质量
 X_1 ——三阶矩 X_2 ——株高
 X_3 ——茎秆面积
 X_4 ——植株外接矩形高

该模型测试集的 R^2 为 0.807 6、MAPE 为 22.15%、SAPE 为 37.53%。

2.3 基于深度信念网络的水稻生物量模型构建

线性回归模型的性能相对较差,为了提高模型性能,构建基于深度信念网络的生物量模型。

DBN 模型训练中可见层数据类型、隐含层层数、隐含层节点数、学习率、迭代次数、动量都会对模型性能产生影响,本文分别测试其对 DBN 模型性能的影响。

(1) 可见层数据类型

分别测试可见层数据类型为高斯型和概率型时对 DBN 模型性能的影响,结果表明高斯型模型的 R^2 为 0.925 9,MAPE 为 11.60%,SAPE 为 18.00%;概率型模型的 R^2 为 0.906 9,MAPE 为 14.51%,SAPE 为 33.18%,高斯型数据类型性能优于概率型。

(2) 隐含层层数

DBN 隐含层层数很大程度上影响着 DBN 模型的性能。选择 5 种不同 DBN 隐含层层数,分别测得它们对 DBN 模型性能的影响,结果如表 2 所示。当隐含层层数为 3 时,模型性能最优。

(3) 隐含层节点数

分别选择 8 种不同的隐含层节点数测试其对 DBN 模型性能的影响,结果如表 3 所示。从结果可以看出,隐含层节点数为 70-70-70 时,模型性能最优。

(4) 学习率

DBN 训练中选择合适的学习率可以有效提高参数学习的收敛速度和学习性能。首先设置受限玻尔兹曼机 (Restricted Boltzmann machines, RBM) 预训练的学习率为 0.1,然后测试 BP 算法不同学习率对 DBN 模型性能的影响。表 4 和表 5 分别为学习

表 1 各时期特征均值与生物量之间的相关性分析

Tab.1 Correlation analysis of feature values for each period and biomass

特征	各时期特征均值					相关系数 r
	胁迫前	胁迫后		复水后		
		对照组	胁迫组	对照组	胁迫组	
H_P	106.96	105.15	115.79	108.51	100.44	0.33 **
N_T	19.50	20.61	15.02	18.37	15.79	0.44 **
A_G	260 921.8	250 039.2	75 231.0	59 959.7	37 695.1	0.26 **
R_{GLA}	0.912	0.812	0.393	0.170	0.133	-0.18 **
A_{ST}	96 417.8	116 926.6	134 947.4	131 617.3	144 418.9	0.19 **
C_1	0.116	0.108	0.165	0.103	0.107	-0.62 **
C_2	0.249	0.250	0.256	0.256	0.254	-0.30 **
C_3	0.232	0.244	0.259	0.262	0.256	-0.28 **
C_4	0.139	0.151	0.150	0.149	0.152	0.03
C_5	0.141	0.147	0.104	0.132	0.131	0.59 **
C_6	0.123	0.101	0.065	0.098	0.100	0.55 **
R_{PA}	0.208	0.194	0.266	0.183	0.190	-0.73 **
W	1 061.09	1 002.88	801.12	1 002.24	925.47	0.73 **
H	1 552.37	1 505.61	1 415.55	1 538.39	1 549.73	0.55 **
D_{F1}	1.517	1.530	1.488	1.544	1.534	0.82 **
R_{PB}	0.191	0.217	0.204	0.230	0.229	0.47 **
A	297 004.6	322 609.0	232 220.7	359 743.1	327 990.4	0.86 **
R_{HW}	1.477	1.530	1.818	1.574	1.719	-0.27 **
D_{F2}	1.621	1.655	1.631	1.671	1.661	0.66 **
F_1	0.257	0.250	0.266	0.240	0.247	-0.37 **
F_2	0.138	0.137	0.144	0.140	0.136	-0.37 **
F_3	0.091	0.092	0.101	0.099	0.094	-0.38 **
F_4	0.068	0.071	0.080	0.079	0.076	-0.37 **
F_5	0.056	0.060	0.069	0.067	0.065	-0.34 **
F_6	0.049	0.053	0.061	0.058	0.059	-0.26 **
F_7	0.046	0.050	0.055	0.053	0.054	-0.13 **
F_8	0.045	0.049	0.050	0.049	0.050	0.05
F_9	0.046	0.050	0.044	0.046	0.047	0.27 **
F_{10}	0.047	0.050	0.039	0.045	0.045	0.45 **
F_{11}	0.047	0.048	0.033	0.042	0.043	0.56 **
F_{12}	0.046	0.042	0.027	0.038	0.038	0.59 **
F_{13}	0.040	0.031	0.020	0.028	0.029	0.52 **
F_{14}	0.024	0.015	0.011	0.015	0.017	0.36 **
A_{DG}	109 389.9	85 325.6	43 181.0	42 955.3	39 966.3	0.36 **
A_{LG}	186 877.1	236 466.0	188 435.8	315 850.7	287 179.0	0.84 **
G_C	0.384	0.180	0.042	0.000	0.001	-0.09 *
M	1.605	1.924	1.607	2.583	2.489	0.81 **
S	4.878	5.602	5.592	7.140	7.266	0.72 **
R	0.814	0.839	0.837	0.873	0.875	0.64 **
M_{U3}	411.259	552.475	693.667	993.404	1 124.840	0.58 **
U	0.818	0.803	0.856	0.782	0.800	-0.86 **
E	0.850	0.911	0.703	1.006	0.928	0.86 **
T_1	0.173	0.224	0.290	0.758	0.691	0.48 **
T_2	0.478	0.394	0.321	0.231	0.244	-0.35 **
T_3	8.869	8.782	13.403	8.191	8.555	-0.69 **
T_4	0.022	0.017	0.013	0.006	0.007	-0.37 **
T_5	0.055	0.044	0.042	0.016	0.018	-0.50 **
T_6	0.231	0.161	0.102	0.051	0.057	-0.29 **
T_7	7.450	6.743	6.410	4.194	4.373	-0.50 **
T_8	1.757	1.599	1.966	1.178	1.231	-0.65 **
T_9	1.965	1.699	1.593	1.197	1.241	-0.48 **
T_{10}	1.293	1.243	1.395	1.006	1.033	-0.63 **
T_{11}	2.997	2.623	2.627	1.785	1.858	-0.55 **
T_{12}	52.346	49.024	40.143	28.148	29.179	-0.33 **
T_{13}	0.035	0.025	0.033	0.019	0.020	-0.64 **
T_{14}	3.272	3.440	3.278	3.913	3.857	0.54 **
T_{15}	2.194	2.182	2.633	2.195	2.236	-0.73 **

注: **表示极显著($P < 0.001$); *表示显著($0.001 < P < 0.05$)。

表 2 隐含层层数对 DBN 模型性能的影响

Tab.2 Influence of hidden layer number on performance of DBN model

隐含层 层数	各隐含层结构	R^2	MAPE/%	SAPE/%
1	70	0.921 4	12.75	28.68
2	70-70	0.927 8	11.79	24.63
3	70-70-70	0.929 9	11.19	18.36
4	70-70-70-70	0.917 4	12.93	25.65
5	70-70-70-70-70	0.914 5	12.69	28.26

表 3 隐含层节点数对 DBN 模型性能的影响

Tab.3 Influence of hidden layer neuron number on performance of DBN model

隐含层节点数	R^2	MAPE/%	SAPE/%
20-20-20	0.923 5	12.67	25.96
30-30-30	0.922 1	12.21	22.96
40-40-40	0.925 2	11.99	26.05
50-50-50	0.926 9	12.21	22.46
60-60-60	0.926 4	11.73	19.02
70-70-70	0.929 9	11.19	18.36
80-80-80	0.923 5	11.69	16.36
90-90-90	0.919 3	13.38	24.99

率随迭代次数改变(式(11))和不随迭代次数改变对模型性能的影响。结果显示,学习率随迭代次数改变且初始学习率为 0.2 时,DBN 训练的精度高,建立的模型更准确。

$$l_i = \frac{500l_0}{\max(i,500)}$$

(11)

式中 l_i ——迭代次数 i 时的学习率
 l_0 ——初始学习率
 i ——迭代次数

表 4 DBN 训练中学习率随迭代次数改变对

DBN 模型性能的影响

Tab.4 Influence of learning rate which varied with change of iterations on performance of DBN model

初始学习率	R^2	MAPE/%	SAPE/%
0.01	0.864 3	16.69	19.93
0.05	0.915 6	12.94	29.25
0.10	0.921 7	12.00	20.85
0.20	0.929 9	11.19	18.36
0.40	0.919 4	11.94	19.37
0.60	0.924 6	13.17	25.45
0.80	0.920 6	15.18	18.58
1.00	0.916 9	13.42	14.34
1.50	0.922 6	21.33	24.09
2.00	0.921 3	17.51	20.43

(5) 迭代次数

迭代次数对模型性能有较大的影响。从迭代次

表 5 DBN 训练中学习率不随迭代次数改变对

DBN 模型性能的影响

Tab.5 Influence of learning rate which did not vary with change of iterations on performance of DBN model

学习率	R^2	MAPE/%	SAPE/%
0.01	0.900 7	14.99	27.97
0.05	0.919 1	12.25	21.61
0.10	0.925 5	11.48	18.86
0.20	0.918 5	13.21	16.32
0.40	0.923 8	12.70	18.86
0.60	0.913 3	16.17	26.67
0.80	0.919 6	12.97	24.68
1.00	0.924 5	15.33	23.39
1.50	0.915 9	19.24	23.47
2.00	0.903 6	12.78	14.95

数 50~2 000 之间选择 9 种不同的迭代次数,分别测试其对 DBN 模型性能的影响,结果如表 6 所示。从表中可知,当 DBN 模型迭代次数为 200 时,模型性能最优。

表 6 迭代次数对 DBN 模型性能的影响

Tab.6 Influence of different numbers of iteration on performance of DBN model

迭代次数	R^2	MAPE/%	SAPE/%
50	0.924 3	11.46	22.52
100	0.923 4	11.62	18.15
200	0.925 9	11.60	17.99
400	0.915 7	13.48	21.76
600	0.913 1	12.58	14.57
1 000	0.908 2	13.48	24.19
1 400	0.910 2	13.96	27.83
1 800	0.914 9	13.23	29.88
2 000	0.917 9	13.41	34.74

(6) 动量

为了测试动量对性能的影响,分别测试了动量随迭代次数改变(式(12))和不随迭代次数改变对 DBN 模型性能的影响。结果表明,可变动量条件下 DBN 的 R^2 为 0.929 9,MAPE 为 11.19%,SAPE 为 18.36%;固定动量条件下 DBN 的 R^2 为 0.915 4,MAPE 为 12.70%,SAPE 为 23.34%。

$$m_i = 0.5 + \frac{0.4}{1 + e^{500-i}}$$

(12)

式中 m_i ——第 i 次迭代的动量

综上,最终确定了性能最优的 DBN 模型,即可见层数据类型为高斯型,隐含层层数为 3 层,隐含层节点数为 70-70-70,学习率为 0.2,迭代次数为 200,动量随迭代次数改变而改变。

2.4 DBN 模型与回归模型性能比较

图 3 为 DBN 模型与回归模型对测试集处理性能比较的结果。

分析图 3 可知, DBN 模型 R^2 较线性回归模型增加了 0.122 3, MAPE 和 SAPE 分别降低了 10.96 个

百分点和 19.17 个百分点, 因此 DBN 模型具有更优的拟合效果。

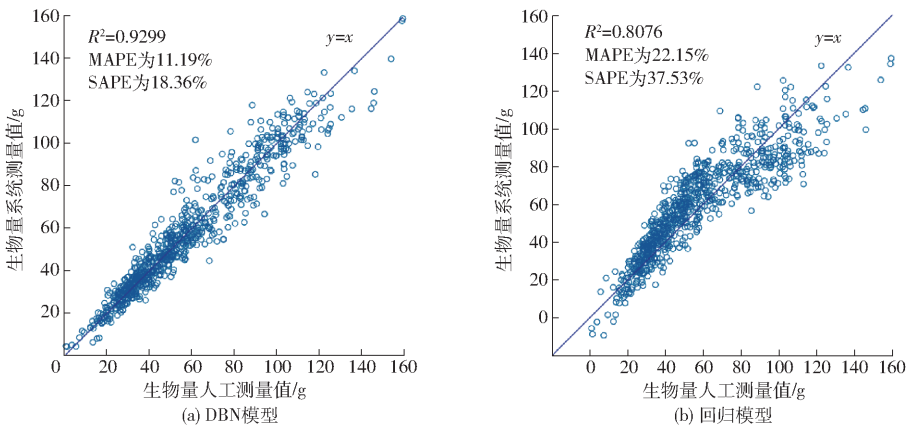


图 3 DBN 模型及回归模型性能比较

Fig.3 Comparison of performance between DBN model and regression model

3 结论

- (1) 对不同品种水稻核心种质资源进行图像采集, 并利用 HSL 颜色空间固定阈值分割法分割图像, 对处理后的图像进行特征提取, 每个植株共获得 57 个水稻特征值。
- (2) 通过逐步线性回归方法构建水稻生物量回归模型, 模型测试集的 R^2 为 0.807 6、MAPE 为

- 22.15%、SAPE 为 37.53%。
- (3) 通过反复试验构建一组可见层数据类型为高斯型、隐含层数为 3、隐含层节点数为 70-70-70、学习率为 0.2、迭代次数为 200 的多品种生物量无损检测 DBN 模型, 模型测试集 R^2 为 0.929 9, MAPE 为 11.19%, SAPE 为 18.36%。
- (4) 将回归模型与 DBN 模型的性能进行对比, 结果表明, DBN 模型具有更优的训练结果。

参 考 文 献

[1] SANG T, GE S. Understanding rice domestication and implications for cultivar improvement[J]. Current Opinion in Plant Biology, 2013, 16(2): 139-146.

[2] 朱德峰, 张玉屏, 陈惠哲, 等. 中国水稻高产栽培技术创新与实践[J]. 中国农业科学, 2015, 48(17): 3404-3414. ZHU Defeng, ZHANG Yuping, CHEN Huizhe, et al. Innovation and practice of high-yield rice cultivation technology in China [J]. Scientia Agricultura Sinica, 2015, 48(17): 3404-3414. (in Chinese)

[3] 张仕颖, 夏运生, 史静, 等. 超高产生态区水稻产量成因研究进展[J]. 中国农学通报, 2014, 30(18): 14-17. ZHANG Shiyong, XIA Yunsheng, SHI Jing, et al. Progress in research on determination factors for super-high-yield rice in special eco-site of Yunnan Province[J]. Chinese Agricultural Science Bulletin, 2014, 30(18): 14-17. (in Chinese)

[4] 范乐乐, 冯跃华, 何腾兵, 等. 黔中地区实地氮肥管理对水稻产量、干物质积累量及氮肥利用率的影响[J]. 中国农学通报, 2011, 27(27): 184-190. FAN Lele, FENG Yuehua, HE Tengbing, et al. Effects of site-specific nitrogen management on yield and dry matter accumulation and fertilizer-N use efficiency of rice in the central areas of Guizhou Province[J]. Chinese Agricultural Science Bulletin, 2011, 27(27): 184-190. (in Chinese)

[5] 闫平, 张书利, 于艳敏, 等. 不同水稻品种干物质积累与产量性状的相关研究[J]. 中国农学通报, 2015, 31(18): 1-6. YAN Ping, ZHANG Shuli, YU Yanmin, et al. Correlation research on dry matter accumulation and yield characters of different rice varieties[J]. Chinese Agricultural Science Bulletin, 2015, 31(18): 1-6. (in Chinese)

[6] 吴桂成, 张洪程, 戴其根, 等. 南方粳型超级稻物质生产积累及超高产特征的研究[J]. 作物学报, 2010, 36(11): 1921-1930. WU Guicheng, ZHANG Hongcheng, DAI Qigen, et al. Characteristics of dry matter production and accumulation and super-high yield of japonica super rice in South China[J]. Acta Agronomica Sinica, 2010, 36(11): 1921-1930. (in Chinese)

[7] 梁淑敏, 杨锦忠. 图像处理技术在玉米株型上的应用研究[J]. 玉米科学, 2007, 15(4): 146-148. LIANG Shumin, YANG Jinzhong. Study on image process application in maize plant type [J]. Journal of Maize Sciences, 2007, 15(4): 146-148. (in Chinese)

[8] 武聪玲, 滕光辉, 李长缨. 黄瓜幼苗生长信息的无损监测系统的应用与验证[J]. 农业工程学报, 2005, 21(4): 109-112. WU Congling, TENG Guanghui, LI Changying. Application and validation of computer vision based nondestructive measurement system for cucumber seedling growth conditions[J]. Transactions of the CSAE, 2005, 21(4): 109-112. (in Chinese)

[9] NAGEL K, PUTZ A, GILMER F, et al. GROWSCREEN-Rhizo is a novel phenotyping robot enabling simultaneous measurements of root and shoot growth for plants grown in soil-filled rhizotrons[J]. Functional Plant Biology, 2012, 39(11): 891-904.

[10] PARENT B, SHAHINNIA F, MAPHOSA L, et al. Combining field performance with controlled environment plant imaging to identify the genetic control of growth and transpiration underlying yield response to water-deficit stress in wheat[J]. Journal of Experimental Botany, 2015, 66(18): 5481-5492.

- [11] HONSDORF N, MARCH T, BERGER B, et al. High-throughput phenotyping to detect drought tolerance QTL in wild barley introgression lines[J]. PLoS ONE, 2014, 9(5): e97047.
- [12] HAIRMANSIS A, BERGER B, TESTER M, et al. Image-based phenotyping for non-destructive screening of different salinity tolerance traits in rice[J]. Rice, 2014, 7:16.
- [13] POORTER H, POT C, LAMBERS H. The effect of an elevated atmospheric CO₂ concentration on growth, photosynthesis and respiration of *Plantago major*[J]. Physiologia Plantarum, 1988, 73(4):553 – 559.
- [14] CHEN D, NEUMANN K, FRIEDEL S, et al. Dissecting the phenotypic components of crop plant growth and drought responses based on high-throughput image analysis[J]. Plant Cell, 2014, 26(12): 4636 – 4655.
- [15] 方伟, 冯慧, 杨万能, 等. 基于可见光成像的单株水稻植株地上部分生物量无损预测方法研究[J]. 中国农业科技导报, 2015, 17(3):63 – 69.
FANG Wei, FENG Hui, YANG Wanneng, et al. Studies on non-destructive optical method for predicting above-ground biomass of individual rice plant based on visible light imaging[J]. Journal of Agricultural Science & Technology, 2015, 17(3): 63 – 69. (in Chinese)
- [16] TACKENBERG O. A new method for non-destructive measurement of biomass, growth rates, vertical biomass distribution and dry matter content based on digital image analysis[J]. Annals of Botany, 2007, 99(4):777 – 783.
- [17] GOLZARIAN M, FRICK R, RAJENDRAN K, et al. Accurate inference of shoot biomass from high-throughput images of cereal plants[J]. Plant Methods, 2011, 7(2): 1 – 11.
- [18] 刘翠红. 水稻盆栽试验水分消耗检测系统研制及其试验研究[D]. 沈阳: 沈阳农业大学, 2012.
- [19] HINTON G, OSINDERO S, TEH Y. A fast learning algorithm for deep belief nets[J]. Neural Computation, 2006, 18(7): 1527 – 1554.
- [20] 李小昱, 库静, 颜伊芸, 等. 基于高光谱成像的绿皮马铃薯检测方法[J/OL]. 农业机械学报, 2016, 47(3):228 – 233.
LI Xiaoyu, KU Jing, YAN Yiyun, et al. Detection method of green potato based on hyperspectral imaging[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2016, 47(3):228 – 233. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?file_no=20160332&flag=1&journal_id=jcsam. DOI: 10. 6041/j. issn. 1000-1298. 2016. 03. 032. (in Chinese)
- [21] 谭文学, 赵春江, 吴华瑞, 等. 基于弹性动量深度学习神经网络的果体病理图像识别[J/OL]. 农业机械学报, 2015, 46(1):20 – 25.
TAN Wenxue, ZHAO Chunjiang, WU Huarui, et al. A deep learning network for recognizing fruit pathologic images based on flexible momentum[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2015, 46(1):20 – 25. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?file_no=20150104&flag=1&journal_id=jcsam. DOI: 10. 6041/j. issn. 1000-1298. 2015. 01. 004. (in Chinese)
- [22] 张善文, 张传雷, 丁军. 基于改进深度置信网络的大棚冬枣病虫害预测模型[J]. 农业工程学报, 2017, 33(19):202 – 208.
ZHANG Shanwen, ZHANG Chuanlei, DING Jun. Disease and insect pest forecasting model of greenhouse winter jujube based on modified deep belief network[J]. Transactions of the CSAE, 2017, 33(19):202 – 208. (in Chinese)
- [23] 黎焯, 赵建, 高云, 等. 基于深度信念网络的猪咳嗽声识别[J/OL]. 农业机械学报, 2018, 49(3):179 – 186.
LI Xuan, ZHAO Jian, GAO Yun, et al. Recognition of pig cough sound based on deep belief nets[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2018, 49(3):179 – 186. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?file_no=20180322&flag=1&journal_id=jcsam. DOI: 10. 6041/j. issn. 1000-1298. 2018. 03. 022. (in Chinese)
- [24] 余永维, 杜柳青, 易小波, 等. 基于时序深度学习的数控机床运动精度预测方法[J/OL]. 农业机械学报, 2019, 50(1):421 – 426.
YU Yongwei, DU Liuqing, YI Xiaobo, et al. Prediction method of NC machine tools' motion precision based on sequential deep learning[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2019, 50(1):421 – 426. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?file_no=20190149&flag=1&journal_id=jcsam. DOI: 10. 6041/j. issn. 1000-1298. 2019. 01. 049. (in Chinese)
- [25] 杨红艳, 杜健民, 王圆, 等. 基于无人机遥感与卷积神经网络的草原物种分类方法[J/OL]. 农业机械学报, 2019, 50(4):188 – 195.
YANG Hongyan, DU Jianmin, WANG Yuan, et al. Classification method of grassland species based on unmanned aerial vehicle remote sensing and convolutional neural network[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2019, 50(4):188 – 195. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?file_no=20190421&flag=1&journal_id=jcsam. DOI: 10. 6041/j. issn. 1000-1298. 2019. 04. 021. (in Chinese)
- [26] 张领先, 陈运强, 李云霞, 等. 基于卷积神经网络的冬小麦麦穗检测计数系统[J/OL]. 农业机械学报, 2019, 50(3):144 – 150.
ZHANG Lingxian, CHEN Yunqiang, LI Yunxia, et al. Detection and counting system for winter wheat ears based on convolutional neural network[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2019, 50(3):144 – 150. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?file_no=20190315&flag=1&journal_id=jcsam. DOI: 10. 6041/j. issn. 1000-1298. 2019. 03. 015. (in Chinese)
- [27] 胡越, 罗东阳, 花奎, 等. 关于深度学习的综述与讨论[J]. 智能系统学报, 2019, 14(1):1 – 19.
HU Yue, LUO Dongyang, HUA Kui, et al. Overview on deep learning[J]. Transactions on Intelligent Systems, 2019, 14(1):1 – 19. (in Chinese)
- [28] LECUN Y, BENGIO Y, HINTON G. Deep learning[J]. Nature, 2015, 521(7553):436 – 444.
- [29] GUO Z, YANG W, CHANG Y, et al. Genome-wide association studies of image traits reveal the genetic architecture of drought resistance in rice[J]. Molecular Plant, 2018, 11(6):789 – 805.
- [30] 蒋霓. 单株水稻绿叶面积无损测量方法研究[D]. 武汉: 华中科技大学, 2014.