

基于栈式自编码神经网络的包衣单籽粒玉米品种识别

李浩光^{1,2} 李卫军^{1,3} 覃 鸿^{1,3} 于丽娜^{1,3} 于云华² 逢 燕²

(1. 中国科学院半导体研究所, 北京 100083; 2. 中国石油大学胜利学院, 东营 257061;

3. 中国科学院大学微电子学院, 北京 100049)

摘要: 常规近红外定性识别研究中,玉米籽粒为表皮裸露状态,未经种衣剂覆盖处理,但是在实际农业生产中,为抵御病虫害侵袭,提高玉米种子发芽率,达到保产增产的功效,玉米种子常需经种衣剂包裹处理。玉米种衣剂的类型多样,对近红外光谱具有一定的吸收,因此种衣剂对近红外定性识别具有干扰作用。本文针对种衣剂对玉米品种识别准确性影响的问题,提出了一种基于栈式自编码神经网络(SAE)的近红外光谱定性建模方法。首先采用无种衣剂玉米籽粒光谱作为训练集,通过栈式自编码无监督学习算法与 softmax 分类器构建栈式自编码网络定性分析模型,再利用所建模型对有种衣剂玉米籽粒进行品种识别。实验结果表明,基于 SAE 的建模方法能够将种衣剂对玉米籽粒识别率的影响降低至 3% 以内。

关键词: 玉米; 品种识别; 栈式自编码神经网络; 种衣剂

中图分类号: O657.3;S123 **文献标识码:** A **文章编号:** 1000-1298(2017)S0-0422-07

Varietal Identification for Single Maize Seed Based on Stacked Auto Encoder Neural Network

LI Haoguang^{1,2} LI Weijun^{1,3} QIN Hong^{1,3} YU Li'na^{1,3} YU Yunhua² PANG Yan²

(1. Institute of Semiconductors, Chinese Academy of Sciences, Beijing 100083, China

2. Shengli College, China University of Petroleum, Dongying 257061, China

3. School of Microelectronics, University of Chinese Academy of Sciences, Beijing 100049, China)

Abstract: In the conventional near infrared qualitative identification, the maize seed kernel epidermis was not treated with seed coating agent. However, in the actual agricultural production, in order to resist the invasion of diseases and insect pests, improve the germination rate, and achieve the effect of maintaining and increasing yield, maize seeds often need to be coated with seed coating agents. In reality, on the market, it is usually necessary to model maize seed kernels without seed coating to identify the ones with seed coating, so as to achieve the purpose of cracking down fake and shoddy products. The maize seeds coating usually consist of a mixture of insecticides, fungicides, fertilizer, plant growth regulators and other ingredients. Their types are diverse and the components are different. These components contain hydrogen group organic compounds, which have certain absorption to near infrared spectrum. Therefore, the seed coating agent had an interference effect on near infrared spectroscopy qualitative identification, which reduced the performance of some conventional shallow learning model. According to the effects of seed coating on maize variety authenticity identification accuracy, a method of near infrared spectroscopy qualitative modeling based on stacked autoencoder (SAE) neural networks has been proposed. Firstly, taking maize seed spectrum without seed coating agent as the training set, a qualitative analysis model was constructed through SAE unsupervised learning algorithm and Softmax classifier. Then, based on this model, the authenticity of maize seeds with seed coating agents was identified. The experimental results showed that, by using the method based on SAE, the effect of seed coating on maize varietal authenticity recognition rate was controlled within 3%.

Key words: maize; varietal authenticity identification; stack auto encoder neural network; seed coating

收稿日期: 2017-07-17 修回日期: 2017-11-20

基金项目: 国家重大科学仪器设备开发专项(2014YQ470377)和中国石油大学胜利学院科技计划项目(KY2017006、KY2015011)

作者简介: 李浩光(1981—),男,博士生,中国石油大学讲师,主要从事智能信息处理和模式识别研究,E-mail: lihaoguang@semi.ac.cn

通信作者: 李卫军(1975—),男,研究员,博士生导师,主要从事智能信息处理、机器视觉、近红外光谱研究,E-mail: wjli@semi.ac.cn

引言

我国假玉米种子事件屡有发生,影响农业生产,损害农民利益。因此,玉米种子品种真实性识别是目前农业生产、作物育种和种子检验的重要问题之一。国内外种子品种识别的常用方法主要有形态学方法、荧光扫描鉴定法、化学鉴定法和电泳鉴定法。形态学方法识别所需时间长,而且精度不高。荧光扫描鉴定法、化学鉴定法和电泳鉴定法等识别精度高,但所需的时间较长,且识别过程复杂,非专业人员难以完成,不适宜对样品进行批量快速分析与无损在线监测^[1-2]。

近红外光谱分析技术由于其无损、便捷、准确等优点在农业领域取得了广泛的应用,该技术目前已被用于玉米品种真实性识别,并取得了较好的效果^[3-4]。常规近红外定性识别研究中,玉米籽粒为表皮裸露状态,未经种衣剂覆盖处理,但是在实际农业生产中,为抵御病虫害侵袭,提高玉米种子发芽率,达到保产增产的功效,玉米种子常需经种衣剂包裹处理^[5]。目前仍需使用无种衣剂玉米籽粒的模型,以识别市场上购得的含种衣剂的玉米籽粒。玉米种衣剂的类型多样,成分不尽相同,通常由杀虫剂、杀菌剂、微肥、植物生长调节剂等成分混合而成,组成成分一般都含有氢基团有机物,对近红外光谱具有一定的吸收。因此,种衣剂对近红外定性识别具有干扰作用,常规浅层学习模型^[6-12]性能易受种衣剂影响。

当前,深度学习是机器学习领域的一大热点,被应用至各个领域的识别与回归问题中,尤其是在图像识别领域取得巨大成功^[13]。深度神经网络采用近似人类大脑的多层次构造,通过学习提取输入信息的多抽象层表示,对输入信息依次从底层到高层进行特征提取,形成适合分类或回归的高阶抽象特征。深度学习的发展也为近红外定性分析提供了新思路。栈式自编码神经网络(Stacked auto encoder

neural network,SAE)具有强大的表达能力,能够获取到输入的“层次型分组”或者“部分-整体分解”结构^[14-20]。近红外光谱数据是高维数据,其中包含了很多无关信息,对数据分类起到干扰作用。通过SAE进行深度学习近红外光谱中的高阶特征,剔除光谱中的无关信息,提取得到能够反映品种间差别的有效信息。

本文针对种衣剂对玉米品种真实性识别准确性影响的问题,提出一种基于栈式自编码神经网络(SAE)的近红外光谱定性建模方法,为有农种衣剂玉米种子品种识别提供新的方法和工具。

1 方法和算法

1.1 栈式自编码神经网络

基于深度学习思想的自编码神经网络是一种无监督学习神经网络^[19],自编码器(Auto encoder, AE)利用编码与解码操作来实现原始输入信息的重构。自编码器信息重构过程,看似毫无意义,实则可通过隐藏神经元环节稀疏性限制,获得一组基向量,并通过该组基向量反映输入向量本征结构。

使用稀疏自编码方法获取输入样本在稀疏约束下的低维特征后,再结合有监督学习方法如 logistic 回归或 softmax 回归,即可实现输入样本的二分类或者多分类任务。

图1是栈式自编码神经网络结构示意图。对于分类问题,将多层稀疏自编码器首尾相接,并与有监督分类器相结合,即可构成能够完成多分类任务的栈式自编码神经网络。在图1的每一级自编码器中,都执行编码和解码操作,在编码和解码过程中实现对上一层自编码器输出的特征抽取。自编码器的输出只是对其输入的重构或近似,并不能直接将其用于对输入样本的直接分类,需结合有监督分类器才能进行对输入信息的分类识别。图1中的3个自编码器经过特征提取即可获得3个隐藏层,在最末级加上有监督分类器后可实现分类识别。

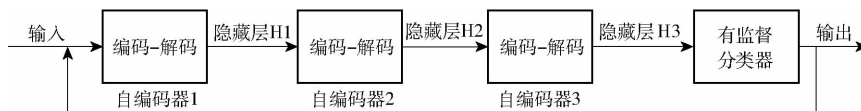


图1 栈式自编码神经网络结构示意图

Fig. 1 Schematic map of stack self-encoding neural network

1.2 自编码器原理

图2是单个自编码器内部网络结构图,由图2可知该网络由3层结构组成,分别是输入层L1、隐藏层L2、输出层L3,其目标是使自编码器输入与输出近似相等,即 $h_{w,b}(x) \approx x$,隐层激活值为 $a^2 = (a_1^2, a_2^2, \dots, a_n^2)$,即输入向量在自编码方式下得到新特征值。

将 sigmoid 函数作为单个神经元激活函数,训练数据中包括 m 个样本,代价函数为

$$J(W, b) = \frac{1}{m} \sum_{i=1}^m \left(\frac{1}{2} \| h_{w,b}(x^{(i)}) - y^{(i)} \|^2 \right) + \frac{\lambda}{2} \sum_{l=1}^{n_l-1} \sum_{i=1}^{s_i} \sum_{j=1}^{s_{i+1}} (W_{ji}^{(l)})^2 \quad (1)$$

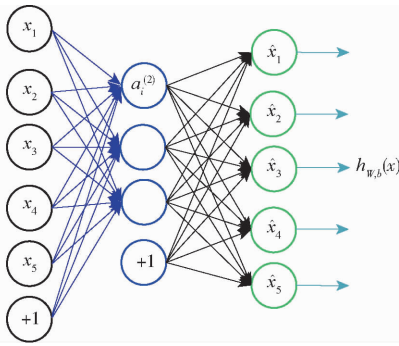


图2 自编码器内部网络结构图

Fig.2 Map of self-encoder internal network structure

式中 $x^{(i)}$ ——输入 $y^{(i)}$ ——输出
 $W_{ji}^{(l)}$ ——前层与后层的连接权重
 b ——偏置值

为叙述方便,假设 x 和 y 分别代表所有的输入与输出。在前向传播过程中,后层与前层的递推关系表示为

$$z^{(l+1)} = \omega^{(l)} a^{(l)} + b^{(l)} \quad (2)$$

$$a^{(l+1)} = f(z^{(l+1)}) \quad (3)$$

对隐藏层加入稀疏性约束条件,即可构成稀疏自编码器,稀疏表达通常比一般方式更为有效。

$a_j^{(2)}(x)$ 为 AE 中隐藏神经元 j 的激活度,其均值计算式为

$$\hat{\rho}_j = \frac{1}{m} \sum_{i=1}^m a_j^{(2)} x^{(i)} \quad (4)$$

式中 ρ ——稀疏性系数,通常设为接近于 0 的数

通过正则化参数 c 可以使激活度均值逼近设定的稀疏性参数,正则化参数为

$$c = \sum_{j=1}^n L(\rho \parallel \hat{\rho}_j) \quad (5)$$

$$L(\rho \parallel \hat{\rho}_j) = \rho \ln \frac{\rho}{\hat{\rho}_j} + (1 - \rho) \ln \frac{1 - \rho}{1 - \hat{\rho}_j} \quad (6)$$

式中, n 为隐藏层中隐藏神经元数量,两个伯努利随机变量分别以 ρ 以及 $\hat{\rho}_j$ 为均值,可通过 $L(\rho \parallel \hat{\rho}_j)$ 度量两个分布之间差异,又称作相对熵。

最小化相对熵可使 $\hat{\rho}_j$ 接近 ρ ,因此总代价函数为

$$J_{\text{sparse}}(W, b) = J(W, b) + \beta \sum_{j=1}^n L(\rho \parallel \hat{\rho}_j) \quad (7)$$

式中, β 是控制稀疏性正则化参数的权重。计算得到代价函数后,即可使用反向传播算法计算代价函数的偏导数。

AE 的输出及误差计算式为

$$\delta_i^{(3)} = -(y_i - a_i^{(3)}) f'(z_i^{(3)}) \quad (8)$$

$$\delta_i^{(2)} = \left[\sum_{j=1}^2 W_{ji}^{(2)} \delta_j^{(3)} + \beta \left(-\frac{\rho}{\hat{\rho}_j} + \frac{1 - \rho}{1 - \hat{\rho}_j} \right) \right] f'(z_i^{(2)}) \quad (9)$$

计算偏导数

$$\nabla_{W^{(l)}} J(W, b; x, y) = \delta^{(l+1)} a^{(l)T} \quad (10)$$

$$\nabla_{b^{(l)}} J(W, b; x, y) = \delta^{(l+1)} \quad (11)$$

再利用批量梯度下降法迭代更新参数 W 和 b , 并求得该网络最优解。

将各级自编码器经过执行编码和解码操作,并在编码和解码过程中通过稀疏性限制得到的多个隐藏层级联即构成了如图 3 所示的多隐藏层结构。

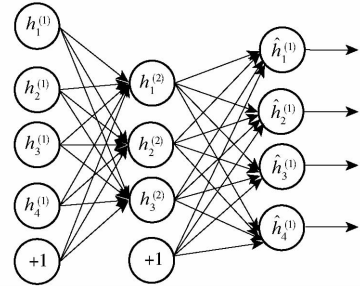


图3 多隐藏层结构示意图

Fig.3 Diagram of multi-hidden layer structure

网络结构末端的 softmax 模型可以表示为

$$h_{\theta}(x^{(i)}) = \begin{bmatrix} p(y^{(i)} = 1 | x^{(i)}; \theta) \\ p(y^{(i)} = 2 | x^{(i)}; \theta) \\ \vdots \\ p(y^{(i)} = k | x^{(i)}; \theta) \end{bmatrix} = \frac{1}{\sum_{j=1}^k e^{\theta_j^T x^{(i)}}} \begin{bmatrix} e^{\theta_1^T x^{(i)}} \\ e^{\theta_2^T x^{(i)}} \\ \vdots \\ e^{\theta_k^T x^{(i)}} \end{bmatrix} \quad (12)$$

向量 $h_{\theta}(x^{(i)})$ 的每一个元素 $p(y^{(i)} = j | x^{(i)}; \theta)$ 表示输入 $x^{(i)}$ 属于第 j 类的概率。 $\theta_1^T, \theta_2^T, \dots, \theta_k^T$ 为模型参数向量,其中 $\theta_k^T \in \mathbf{R}^{n+1}$,将这些向量写成矩阵形式

$$J'(\theta) = -\frac{1}{m'} \left[\sum_{i=1}^{m'} \sum_{j=1}^k 1 \{y^{(i)} = j\} \ln \frac{e^{\theta_j^T x^{(i)}}}{\sum_{i=1}^k e^{\theta_j^T x^{(i)}}} \right] + \frac{\lambda'}{2} \sum_{i=1}^k \sum_{j=0}^n \theta_{ij}^2 \quad (13)$$

式中 $1 \{ \cdot \} = 1$,为示性函数,该函数取值规则为 $\{ \cdot \}$ 值为真的表达式 $\{ \cdot \} = 1$ 。加入权重衰减项 ($\lambda' > 0$) 后,代价函数则是严格的凸函数,使用梯度下降法或牛顿法可使 $J'(\theta)$ 收敛至全局最优解,得到最优解 θ 后即完成了网络的训练过程,后期可将训练所得网络用于对未知类别光谱样本进行有监督分类识别。

1.3 算法设计

基于 SAE 的近红外光谱定性分析模型算法流程如下:

(1) 根据设定的 SAE 初始参数,计算首级自编码器参数。训练参数时,以学习到最能反映原始输入近红外光谱数据的新特征向量为目标,将所有原始特征代入第一级自编码器计算并得到其输出。

(2)将前级自编码器输出作为下级自编码器输入,并使用步骤(1)中的方法训练该级自编码器参数。重复步骤(2),直至训练完成所有的隐藏层。

(3)将最后一级自编码器输出作为有监督分类器 softmax 分类输入,并利用网络输入近红外光谱数据所对应类别标签来训练分类器参数。

(4)通过梯度下降算法优化网络所有节点参数,获得最终用于近红外光谱识别的栈式自编码神经网络模型。

基于上述分析,设计的栈式自编码神经网络隐藏层结构为3层,分别是:第1层自编码器结构:125-50-125,输入输出层为125维,隐藏层神经元为50维。第2层自编码器结构:50-25-50,输入输出层为50维,隐藏层神经元为25维。第3层自编码器结构:25-5-25,输入输出层为25维,隐藏层神经元为5维,构建的网络隐藏层结构为50-25-5。

综上,本节中自编码器稀疏参数 ρ 设为0.09,学习率为0.001,神经元激活函数为sigmoid。训练网络时使用随机梯度下降法,每一层迭代100次,分别得到自编码网络权重 W_1 、 W_2 和 W_3 ,预训练后,构造栈式自编码神经网络。

2 试验结果与分析

试验所用材料为农华032、农大108、京玉16、洛单248、屯玉2102、屯玉2104、屯玉2105、屯玉2106、屯玉2107、屯玉2109、郑单958、浚单20等12个品种足量含种衣剂以及不含种衣剂的玉米籽粒。在连续8天分别采集上述12个品种玉米单籽粒的光谱,并将每日采集的数据定义为T1,T2,...,T8数据集。采集时使用课题组自制的单籽粒光谱采集装置^[11],以漫透射的方式逐粒交替采集,即按照农华032(无种衣剂)—农华032(有种衣剂)—农大108(无种衣剂)—农大108(有种衣剂)—京玉16(无种衣剂)—京玉16(有种衣剂)—洛单248(无种衣剂)—洛单248(有种衣剂)……的方式采集单粒光谱,中间无校验。每个品种有种衣剂玉米籽粒以及无种衣剂玉米籽粒各采集100粒,采集时随机排放每个玉米籽粒的胚面朝向,以模拟实际自动化分选时的情形,每个籽粒玉米采集10条光谱取平均。采用这种方式采集光谱的原因是:在自动化测试过程中,一堆未知属性的玉米种子并不知道具体每一粒是何品种,因此在采集光谱时,要按照间隔交替的方式,以适应在自动化测试时的情况。

试验中为测试SAE方法所建模型的性能,同时采用几种分类方法对单日无种衣剂光谱数据建立模型,再使用多个测试集中的无种衣剂籽粒光谱与有

种衣剂籽粒光谱进行对比试验。

分类方法设置如下:

SVM:试验中使用LIBSVM工具箱,设置SVM分类器类型为Nu-SVC,多分类类型,参数设置通过工具箱中的网格寻参函数交叉验证来确定,选择径向基核函数,参数Nu设为0.5,核函数中的gamma值设为2.8。

BPANN:采用平滑(11)+一阶导(9)+归一化预处理方法,将预处理后得到的输出直接作为BP神经网络的输入,通过5折交叉验证的方法确定最优的隐藏层数目为8。

BPR:使用KS方法选择构网样本点,超香肠作为基本覆盖单元,通过交叉验证确定超香肠神经元的半径。

利用T1数据集的无种衣剂玉米籽粒光谱作为建模集,使用后7日时段内测量的玉米籽粒近红外光谱分别作为测试集进行测试,每一个测试集中又分为无种衣剂玉米籽粒光谱及有种衣剂玉米籽粒光谱。实验中使用4种机器学习方法对近红外光谱进行分类识别。

图4是不同测试集中无种衣剂玉米籽粒的正确识别率变化曲线。

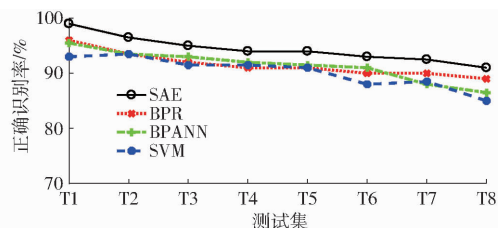


图4 不同测试集无种衣剂籽粒识别率变化曲线

Fig. 4 Diagram of grain recognition rate without seed coat in different test sets

由图4中曲线可以看出,若以T1数据集中所有的无种衣剂籽粒的光谱数据作为建模集,使用其他几个单日数据集中的无种衣剂光谱数据进行测试时,随着测试集与建模集时间间隔的增大,识别准确率出现下降的趋势。

上述识别率下降趋势说明近红外光谱定性分析模型的预测能力与时间推移密切相关。种子作为活体,内部存在新陈代谢作用,随着时间推移,其内部成分也会相应发生变化,此外温度、湿度、仪器状态等周围环境也会发生细微的变化,导致识别准确率的下降。

随着时间推移,几种方法的识别率都出现了下降,但是下降程度不同,由图4可知:

SAE方法:在时间推移过程中,其识别率趋势曲线下降最为缓慢,至T8数据集时,识别准确率仍在90%以上。

BPR 方法:在时间推移过程中,其识别率趋势曲线下下降相对也较为缓慢,至 T5 数据集后,识别准确率开始低于 90%。

BPANN、SVM 方法:所对应的识别准确率趋势曲线在 T6 测试集后即开始低于 90%。

图 5 是以 T1 数据集中的无种衣剂籽粒光谱数据为建模集,对 T1 ~ T8 不同测试集中有种衣剂玉米籽粒光谱数据进行识别时的识别率变化趋势曲线。

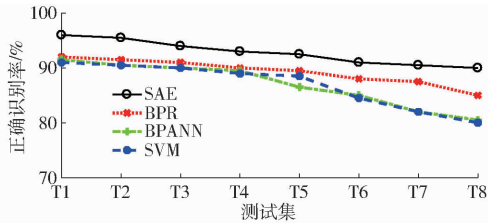


图 5 不同测试集有种衣剂籽粒识别率变化曲线

Fig. 5 Diagram of grain recognition rate with seed coat in different test sets

图 5 中,以 T1 中无种衣剂数据作为建模集,使用多个数据集中有种衣剂籽粒光谱数据进行测试时,随着测试集与建模集时间间隔的增大,识别率也呈现出下降的趋势。

由图 5 中 4 条曲线对比可知:

SAE 方法在时间推移过程中,识别率下降最为缓慢,在 T7 时正确识别率仍然高于 90%。

BPR 方法在 T5 数据集时识别准确率低于 90%,模型的泛化性能略弱于 SAE 方法。

BPANN、SVM 方法在 T3 测试集后正确识别率开始低于 90%。

图 6 是分别使用 4 种建模方法建立定性分析模型所得识别结果。

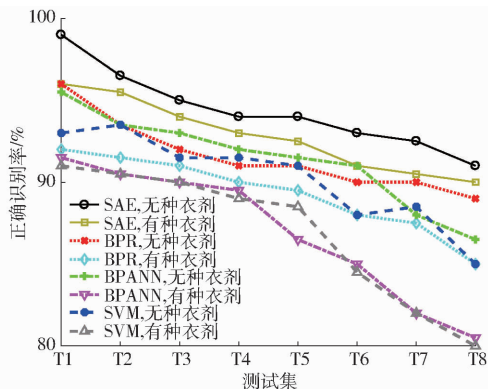


图 6 不同测试集有、无种衣剂籽粒正确识别率比较

Fig. 6 Comparison of the accuracy of grain identification in different test sets

建立模型时,采用 T1 数据集中无种衣剂玉米籽粒光谱数据作为建模集,对不同测试集中有种衣剂玉米籽粒与无种衣剂玉米籽粒光谱数据进行识别得

到准确率变化曲线。

由图 6 可知,有种衣剂籽粒光谱数据的趋势曲线均明显低于无种衣剂光谱数据的趋势曲线,说明种衣剂对玉米品种的正确识别率有非常明显影响,有种衣剂籽粒识别率相对于无种衣剂籽粒识别率均出现不同程度下降。

在时间推移过程中,有如下现象:

SAE 方法所得的有种衣剂数据识别率与无种衣剂数据识别率差异最小。

BPR 方法所得的种衣剂数据识别率与无种衣剂数据识别率差异比 SAE 方法略高。

BPANN、SVM 方法所得的有种衣剂数据识别率与无种衣剂数据识别率下降最为明显。

综上,几种建模方法中,SAE 方法对种衣剂噪声具有最强抗干扰作用。

图 7 中 4 条曲线对应 4 种方法,分别为不同测试集中 12 个品种无种衣剂与有种衣剂玉米籽粒的正确识别率差异曲线。

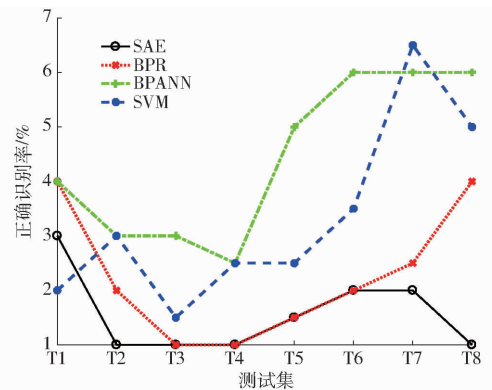


图 7 有种衣剂、无种衣剂籽粒正确识别率差异

Fig. 7 Differences in the accuracy of seed identification with or without seed coat

由图 7 可知,对于 T2 ~ T7 等 7 个测试集,有如下结果:

(1)使用 SAE 方法时,有种衣剂籽粒数据相对于无种衣剂籽粒数据的正确识别率下降最小,正确识别率下降值范围为 1% ~ 3%,随时间推移,正确识别率下降值基本不变。

(2)使用 BPR 方法时,有种衣剂籽粒数据相对于无种衣剂籽粒数据的识别率下降值比 SAE 方法略高,正确识别率下降值范围为 1% ~ 4%。

(3)使用 SVM 方法时,有种衣剂籽粒数据相对于无种衣剂籽粒数据的正确识别率下降较大,正确识别率下降值范围为 1.5% ~ 6.5%,正确识别率下降值随着时间推移有增大的趋势。

(4)使用 BPANN 方法时,有种衣剂籽粒数据相对于无种衣剂籽粒数据的识别率下降最为明显,正确识别率下降值变化范围为 2.5% ~ 6%,且随着时间

推移也有增大的趋势。由图可以直观看出 SAE 方法所建的定性分析模型受玉米种衣剂的影响最小。

综上,基于栈式自编码神经网络的定性分析模型对种衣剂光谱噪声具有较强的抗干扰能力。

3 结论

(1) 以无种衣剂数据作为建模集,使用多个数据集中有种衣剂籽粒光谱数据进行测试,随着测试集与建模集时间间隔的增大,使用 4 种模式识别方法识别率均呈现出下降趋势,表明玉米籽粒外表覆

盖种衣剂对近红外正确识别率造成影响。

(2) 针对种衣剂对玉米品种真实性识别准确性影响的问题,提出了一种基于栈式自编码神经网络(SAE)的近红外光谱定性建模方法。首先采用无种衣剂玉米籽粒光谱作为训练集,通过栈式自编码无监督学习算法与 softmax 分类器构建栈式自编码网络定性分析模型,再利用所建模型对有种衣剂玉米籽粒进行品种真实性识别。实验结果表明,基于 SAE 的建模方法能够将种衣剂对玉米籽粒识别率的影响降低至 3% 以内。

参 考 文 献

- 1 严衍禄,陈斌,朱大洲. 近红外光谱分析的原理、技术与应用[M]. 北京:中国轻工业出版社,2007:60-114.
- 2 陈绍江. 作物育种工程化与工程化育种思考[J]. 作物杂志,2013:1-4.
CHEN Shaojiang. Crop breeding engineering and engineered breeding[J]. Crops, 2013(6):1-4. (in Chinese)
- 3 覃鸿,王徽蓉,李卫军,等. 基于 DPLS 特征提取的 LDA 方法在玉米近红外光谱定性分析中的应用[J]. 光谱学与光谱分析,2011,31(7):1777-1781.
QIN Hong, WANG Huirong, LI Weijun, et al. Application of DPLS-based LDA in corn qualitative near infrared spectroscopy analysis[J]. Spectroscopy and Spectral Analysis, 2011, 31(7):1777-1781. (in Chinese)
- 4 李浩光,李卫军,覃鸿,等. 一种添加历史数据的近红外光谱特征提取方法研究[J]. 光谱学与光谱分析,2016,36(10):3148-3153.
LI Haoguang, LI Weijun, QIN Hong, et al. A new feature extraction method of the near-infrared spectra based on the addition of the historical data[J]. Spectroscopy and Spectral Analysis, 2016, 36(10):3148-3153. (in Chinese)
- 5 姜军,赵霞,黄璐,等. 玉米种衣剂研究进展[J]. 河北农业科学,2008,12(9):49-50.
JIANG Jun, ZHAO Xia, HUANG Lu, et al. Research progress of the seed coating formulations of maize[J]. Journal of Hebei Agricultural Sciences, 2008, 12(9):49-50. (in Chinese)
- 6 张丽萍,李卫军,王平,等. 玉米品种近红外光谱鉴别技术中的参数漂移问题研究[J]. 光谱学与光谱分析,2012,32(10):2785-2788.
ZHANG Liping, LI Weijun, WANG Ping, et al. Research on the parameter drift problem of near infrared spectra based corn variety discrimination technology[J]. Spectroscopy and Spectral Analysis, 2012, 32(10):2785-2788. (in Chinese)
- 7 曹吾,李卫军,王平,等. 近红外定性分析模型的稳健性与适应性分析[J]. 光谱学与光谱分析,2014,34(6):1506-1511.
CAO Wu, LI Weijun, WANG Ping, et al. Analysis of stability and adaptability of near infrared spectra qualitative analysis model[J]. Spectroscopy and Spectral Analysis, 2014, 34(6):1506-1511. (in Chinese)
- 8 柳培忠,张丽萍,李卫军,等. 基于 DS 算法的玉米近红外定性分析光谱校正方法研究[J]. 光谱学与光谱分析,2014,34(6):1533-1537.
LIU Peizhong, ZHANG Liping, LI Weijun, et al. Study on spectral calibration of discrimination of corn variety using near infrared spectra based on DS algorithm[J]. Spectroscopy and Spectral Analysis, 2014, 34(6):1533-1537. (in Chinese)
- 9 李浩光,李卫军,覃鸿,等. 基于最小平方误差的单倍体和二倍体分类方法研究[J]. 农业机械学报,2016,47(6):259-264.
LI Haoguang, LI Weijun, QIN Hong, et al. Research on haploid and diploid classifying method based on least square error[J]. Transactions of the Chinese Society for Agricultural Machinery, 2016, 47(6):259-264. (in Chinese)
- 10 李天昕,贾仕强,刘旭,等. 应用近红外反射和透射光谱法鉴定玉米杂交种的纯度[J]. 光谱学与光谱分析,2015,35(12):3388-3392.
LI Tianxin, JIA Shiqiang, LIU Xu, et al. Maize hybrid seed purity identification based on near infrared reflectance and transmittance spectra[J]. Spectroscopy and Spectral Analysis, 2015, 35(12):3388-3392. (in Chinese)
- 11 覃鸿,马竞一,陈绍江,等. 基于近红外漫反射与漫透射光谱的玉米单倍体鉴别比较研究[J]. 光谱学与光谱分析,2016,36(1):292-297.
QIN Hong, MA Jingyi, CHEN Shaojiang, et al. Identification of haploid maize kernel using NIR spectroscopy in reflectance and transmittance modes: a comparative study[J]. Spectroscopy and Spectral Analysis, 2016, 36(1):292-297. (in Chinese)
- 12 覃鸿,马竞一,陈绍江,等. 光照强度和光阑孔径对近红外漫透射光谱鉴别单倍体玉米影响研究[J]. 红外技术,2015,37(1):78-81.
QIN Hong, MA Jingyi, CHEN Shaojiang, et al. Research on the effect of the intensity and amount of light source to maize haploid identification based on near infrared diffuse transmission[J]. Infrared Technology, 2015, 37(1):78-81. (in Chinese)
- 13 郭丽丽,丁世飞. 深度学习研究进展[J]. 计算机科学,2015,42(5):28-33.

- GUO Lili, DING Shifei. Research progress on deep learning[J]. Computer Science, 2015, 42(5):28-33. (in Chinese)
- 14 VINCENT P, LAROCHELLE H, BENGIO Y, et al. Extracting and composing robust features with denoising autoencoders [C] // Proceedings of the 25th International Conference on Machine Learning, 2008: 1096-1103.
- 15 VINCENT P, LAROCHELLE H, LAJOIE I, et al. Stacked denoising autoencoders: learning useful representations in a deep network with a local denoising criterion [J]. Journal of Machine Learning Research, 2010, 11(6): 3371-3408.
- 16 FARABET C, COUPRIE C, NAJMAN L, et al. Learning hierarchical features for scene labeling [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2013, 35(8):1915-1929.
- 17 HINTON G, DENG L, YU D, et al. Deep neural networks for acoustic modeling in speech recognition: the shared views of four research groups [J]. IEEE Signal Processing Magazine, 2012, 29(6):82-97.
- 18 STELLAN O. Deep learning [M]. Cambridge University Press, 2011:112-156.
- 19 LI Weijia, FU Haohuan, LE Yu, et al. Stacked autoencoder-based deep learning for remote-sensing image classification: a case study of African land-cover mapping [J]. International Journal of Remote Sensing, 2016, 37(23): 5632-5646.
- 20 宋鹏, 吴科斌, 张俊雄, 等. 基于计算机视觉的玉米单倍体自动分选系统[J]. 农业机械学报, 2010, 41(增刊):249-252.
SONG Peng, WU Kebin, ZHANG Junxiong, et al. Sorting system of maize haploid kernels based on computer vision [J]. Transactions of the Chinese Society for Agricultural Machinery, 2010, 41(Supp.): 249-252. (in Chinese)

(上接第 421 页)

- 13 刘燕德, 靳昙昙, 王海阳. 基于拉曼光谱的三组分食用调和油快速定量检测[J]. 光学精密工程, 2015, 23(9):2490-2496.
LIU Yande, JIN Tantan, WANG Haiyang. Rapid quantitative determination of components in ternary blended edible oil based on Raman spectroscopy [J]. Optics and Precision Engineering, 2015, 23(9): 2490-2496. (in Chinese)
- 14 杨坤, 褚昆昆, 朱祥, 等. 几种典型食用油脂拉曼光谱分析研究[J]. 粮食与油脂, 2015(11):56-58.
YANG Kun, CHU Kunkun, ZHU Xiang, et al. Analysis on Raman spectroscopy of several typical edible oils [J]. Cereals & Oils, 2015(11):56-58. (in Chinese)
- 15 BARTHUS R C, POPPI R J. Determination of the total unsaturation in vegetable oils by Fourier transform Raman spectroscopy and multivariate calibration[J]. Vibrational Spectroscopy, 2001, 26(1):99-105.
- 16 赵薇, 刘翠玲, 孙晓荣, 等. 应用拉曼光谱技术识别食用油的种类[J]. 食品科技, 2015(3):274-278.
ZHAO Wei, LIU Cuiling, SUN Xiaorong, et al. Application of Raman spectroscopy for classification of edible oils [J]. Food Science and Technology, 2015(3): 274-278. (in Chinese)
- 17 邓之银, 张冰, 董伟, 等. 拉曼光谱和 MLS-SVR 的食用油脂脂肪酸含量预测研究[J]. 光谱学与光谱分析, 2013, 33(11): 2997-3001.
DENG Zhiyin, ZHANG Bing, DONG Wei, et al. Research on prediction method of fatty acid content in edible oil based on Raman spectroscopy and multi-output least squares support vector regression machine [J]. Spectroscopy and Spectral Analysis, 2013, 33(11): 2997-3001. (in Chinese)
- 18 刘善梅, 李小昱, 钟雄斌, 等. 基于高光谱成像技术的生鲜猪肉含水率无损检测[J/OL]. 农业机械学报, 2013, 44(增刊 1):165-170. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=2013s130&journal_id=jcsam. DOI:10.6041/j.issn.1000-1298.2013.S1.030.
LIU Shanmei, LI Xiaoyu, ZHONG Xiongbn, et al. Non-destructive detection of water content in fresh pork based on Hyperspectral imaging technology[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2013, 44(Supp.1): 165-170. (in Chinese)
- 19 杨国强, 张淑娟, 赵艳茹. 基于近红外透射光谱的汾阳王酒快速鉴别[J/OL]. 农业机械学报, 2013, 44(增刊 1):189-193. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=2013s134&journal_id=jcsam. DOI:10.6041/j.issn.1000-1298.2013.S1.034.
YANG Guoqiang, ZHANG Shujuan, ZHAO Yanru. Fast discrimination of adulterated Fenyangwang wine based on near infrared spectroscopy[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2013, 44(Supp.1): 189-193. (in Chinese)
- 20 ASHTON L, JOHANNESSEN C, GOODACRE R, et al. The importance of protonation in the investigation of protein phosphorylation using Raman spectroscopy and Raman optical activity[J]. Analytical Chemistry, 2011, 83(20):7978-7980.