

近红外结合极限学习机快速识别牛肉中掺假猪肉

韩方凯¹, 刘璨¹, 黄煜¹, 马梅², 冯凡¹, 段腾飞¹, 张东京¹ (1. 宿州学院, 安徽宿州 234000; 2. 江苏大学, 江苏镇江 212013)

摘要 [目的]建立牛肉中掺假猪肉的快速鉴定方法。[方法]采用傅立叶变换近红外结合极限学习机(extreme learning machine, ELM)构建纯牛肉、牛肉中掺假猪肉、纯猪肉的快速识别模型,考察标准正态变换、多元散射校正(multiplicative scatter correction, MSC)、一阶导数及二阶导数结合核主成分分析(kernel principal component analysis, KPCA)等光谱预处理方法对 ELM 模型预测性能的影响。[结果]MSC+KPCA 预处理下,ELM 模型的预测效果最优,训练集及测试集的正确识别率分别为 86.67%和 83.33%。[结论]近红外光谱技术结合 ELM 在牛肉中掺假猪肉的快速鉴定方面具有较大的潜力。

关键词 牛肉掺假;近红外;极限学习机;核主成分分析

中图分类号 TS 207.3 **文献标识码** A

文章编号 0517-6611(2019)13-0185-03

doi: 10.3969/j.issn.0517-6611.2019.13.057



开放科学(资源服务)标识码(OSID):

Rapid Identification of Beef Adulteration with Pork by Near-infrared Combined Limit Learning Machine

HAN Fang-kai, LIU Can, HUANG Yu et al (Suzhou University, Suzhou, Anhui 234000)

Abstract [Objective] The research aimed to develop a rapid technique for identification of beef adulteration with pork. [Method] The Fourier transform near-infrared combined with extreme learning machine (ELM) was used to build prediction models for identification of pure beef, beef adulteration with pork, and pure pork. The influence of different spectral pretreatment methods on the performance of ELM models were studied, such as standard normal variate transformation, multiplicative scatter correction (MSC), first derivative and second derivative combined separately with kernel principal component analysis (KPCA). [Result] The best ELM model was obtained under MSC+KPCA with the correct recognition rate in train set and prediction set was 86.67% and 83.33% respectively. [Conclusion] The Fourier transform near-infrared in coupled with ELM has a great potential in rapid identification of beef adulteration with pork.

Key words Beef adulteration; Near-infrared; Extreme learning machine; Kernel principal component analysis

我国牛肉市场前景广阔。牛肉营养极为丰富,尤其是氨基酸的分布较其他大宗肉类食品更易满足机体需要,深受消费者欢迎;随着经济水平的提高,人们对牛肉的购买力亦在增强。据经济合作与发展组织统计资料显示^[1],2016年我国牛肉的消费量为771.3万t,较2015年增加了46.26万t,增长率为6.4%,增速较快。市场快速发展的同时,牛肉掺假现象也日趋严重。其中,以价格较低廉的猪肉等肉制品掺入牛肉中冒充正常牛肉,赚取高额利润,是牛肉掺假的常见方式。牛肉掺假严重损害消费者利益,阻碍行业发展,且有可能牵涉到宗教信仰问题^[2]。因此,保障生鲜牛肉的良好品质非常重要。在牛肉品质保障体系的众多环节中,对掺假牛肉进行客观、准确、快速的识别,是极其重要的一环。目前,PCR法、色谱法、质谱法等是掺假肉鉴定最被广泛采用的技术,虽然检测结果客观、可靠,但存在耗时长、样品预处理复杂,且难以实现在线检测的缺陷^[3]。掺假肉与正常肉风味组分的差异可被电子鼻有效识别^[3],然而基于物理吸附原理的传统金属氧化物气体传感器存在灵敏度低,难以区分相似物质,且受环境湿度影响大的缺陷;高光谱能够同时获取检测对象的图像及光谱信息,可有效识别掺假肉与正常肉理化组分及形态学的差异^[4]。然而,巨大的数据量给高光谱信号的储存及处理带来极大的困难。

近红外光谱是样品中化学组分含氢基团基频振动的倍

频和合频吸收,分析过程快速,且无需复杂的样品预处理,在食品原料的品质分级、食品加工过程监测、成品食品品质鉴定及溯源等领域应用颇为广泛^[5],非常适合开发掺假牛肉的快速识别方法。2013年,Alamprese等^[6]采用近红外光谱技术对牛肉中掺假火鸡肉进行识别研究,针对掺假比例分别为0~10%、15%~20%、30%~40%、50%及100%的5类样本构建了线性判别模型(linear discernment analysis, LDA),结果显示,LDA模型训练集及测试集的预测准确率分别为88.3%和71.2%。同年,Morsy等^[7]采用近红外光谱技术实现了纯牛肉及掺假牛肉(掺入猪肉等)2类样本的准确判别。2014年,孟一等^[8]采用近红外光谱技术对纯猪肉、牛肉和羊肉进行了鉴别研究,LDA模型结果显示,测试集对猪肉、牛肉和羊肉的识别准确率分别为99.28%、97.42%和100%。2015年,张丽华等^[9]采用近红外光谱技术对牛肉中掺假鸭肉进行了检测分析,并构建了纯牛肉、牛肉中掺假鸭肉、纯鸭肉3类样本的支持向量机(support vector machine, SVM)识别模型,结果显示,SVM模型训练集及测试集的识别准确率分别为97.09%和94.00%,效果良好。

极限学习机(extreme learning machine, ELM)是一种新型单隐含层前馈神经网络算法。ELM可避免采用梯度下降法的传统神经网络算法训练速度慢、容易陷入局部极小点及对网络学习率敏感的缺陷。ELM训练过程中无需调整输入层与隐含层的连接权值及隐含层神经元阈值,较传统的神经网络算法可显著提高运算速度,且泛化能力好,在食品品质的分析中应用广泛^[10-12]。然而,采用近红外光谱技术结合ELM用于牛肉中掺假猪肉的识别研究,迄今鲜见报道。笔者尝试采用近红外光谱技术结合ELM建立纯牛肉、猪肉及牛

基金项目 安徽省自然科学基金项目(1908085QC146);安徽高校自然科学研究重点项目(KJ2018A0447);宿州学院教授(博士)科研启动基金项目(2016j03)。

作者简介 韩方凯(1989—),男,安徽宿州人,讲师,博士,从事食品、农产品品质快速检测新方法研究。

收稿日期 2019-04-12

肉中掺假猪肉3类样本的快速鉴别方法。

1 材料与方法

1.1 试验装备 傅立叶变换近红外光谱仪(WQF-600N,北京北分瑞利分析仪器有限公司);小型绞肉机(AUX-J20,奥克斯集团有限公司);电子天平(BSA224S,赛多利斯科学仪器有限公司)。

1.2 样本制备 牛肋条肉、猪肋条肉(五花肉)购自宿州当地农贸市场,源于当日屠宰的生牛、生猪,以减小因腐败造成的鲜肉品质差异。将牛肋条肉、猪肋条肉分别置于绞肉机中,打浆2 min,备用。按照掺假比例为0%(纯牛肉)、20%、40%、60%、80%、100%配制试验样本。每种掺假比例下采样14次,单个样品重量为40.0 g,共得84个样本供近红外光谱数据采集。

1.3 近红外光谱数据采集 采用傅立叶变换近红外仪采集肉样漫反射光谱。仪器参数设置如下:3 300~10 000波数(wave number, WN),扫描步长0.316 495 746 WN,采样分辨率4 WN,扫描次数32次。试验时,将样品直接放置于样品室载物台进行漫反射光谱信息采集,扣除本底后,作为原始数据输出。

1.4 数据处理方法 近红外漫反射光谱信噪比低,且包含部分物理信息。对比采用标准正态变换(standard normal variate, SNV)、多元散射校正(multiplicative scatter correction, MSC)、一阶导数(first derivative, FD)、二阶导数(second derivative, SD)等算法对原始光谱信息进行处理,以削弱背景噪音,提高谱图与样品化学成分之间的相关性。

鲜肉中含氢化学组分为丰富,且同一含氢化合物会吸收多个频率的近红外光,不同含氢化合物会对同一频率的近红外光产生吸收,导致近红外光谱数据具有严重的共线性;近红外原始数据维度较高,这增大了后续模式建模的难度,同时会对模型的稳定性及预测性能产生不利影响。为了深入挖掘肉样近红外漫反射光谱信息中的非线性信息,采用核主成分分析(kernel principal component analysis, KPCA)对光谱数据进行降维和去相关处理。

采用ELM构建基于近红外的牛肉中掺假猪肉快速识别模型。以肉样近红外光谱数据经KPCA后,累计贡献率大于85%的最少主成分得分作为输入,构建ELM模型,对纯牛肉、牛肉中掺假猪肉及纯猪肉3类样本进行识别研究。最终以识别准确率作为ELM模型预测性能的评价指标。

2 结果与分析

2.1 核主成分分析 采用KPCA对肉样近红外光谱进行降维和去相关处理,并考察不同去噪方法对结果的影响。统计累计贡献率大于85%所需的最小主成分个数及累计贡献率,结果如图1所示。由图1可知,MSC法所得主成分个数最少,为18,此时累计贡献率为85.15%;FD法所得累计贡献率最大,为85.41%,此时主成分个数为30。

2.2 极限学习机识别结果 研究随机选取每组样本的2/7作为独立样本,其余60个样本作为训练集,构建ELM识别模型。经过多次试验,选取“hardlim”函数作为ELM隐含层

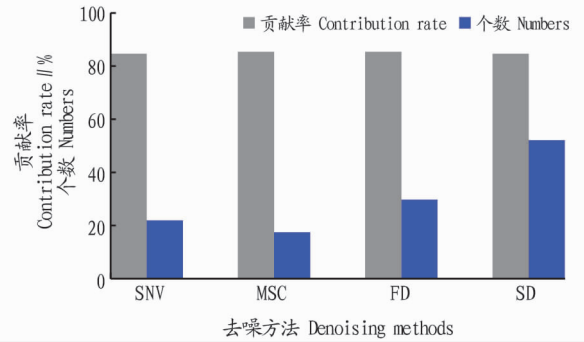


图1 不同去噪方法下KPCA累计贡献率大于85%时的最小主成分数及累计贡献率

Fig. 1 The minimum principal component number and accumulated contribution rate (ACR) of KPCA under different denoising methods, while ACR > 85%

神经元激活函数。作为典型ELM算法中唯一可变的参数,隐含层神经元个数对ELM预测性能有一定的影响。然而,ELM隐含层神经元个数的选取尚缺乏理论指导,往往需要依据具体情况而定。研究对ELM隐含层神经元个数进行筛选,范围设置为1~50,以独立样本识别准确率为评价指标,选择准确率最高的最少神经元个数作为最终ELM模型的最佳隐含层神经元个数。所建ELM模型识别结果如表1所示。

表1 ELM模型识别结果

Table 1 Recognition results of ELM models

预处理方法 Pretreatment method	输入层节点数 Node numbers of input layer	最佳隐含层 节点数 Node numbers of optimal hidden layer	正确识别率 Correct recognition rate / %	
			训练集 Training set	测试集 Testing set
SNV+KPCA	22	19	80	75
MSC+KPCA	18	43	86.67	83.33
FD+KPCA	30	18	83.33	75
SD+KPCA	52	15	75	70.83

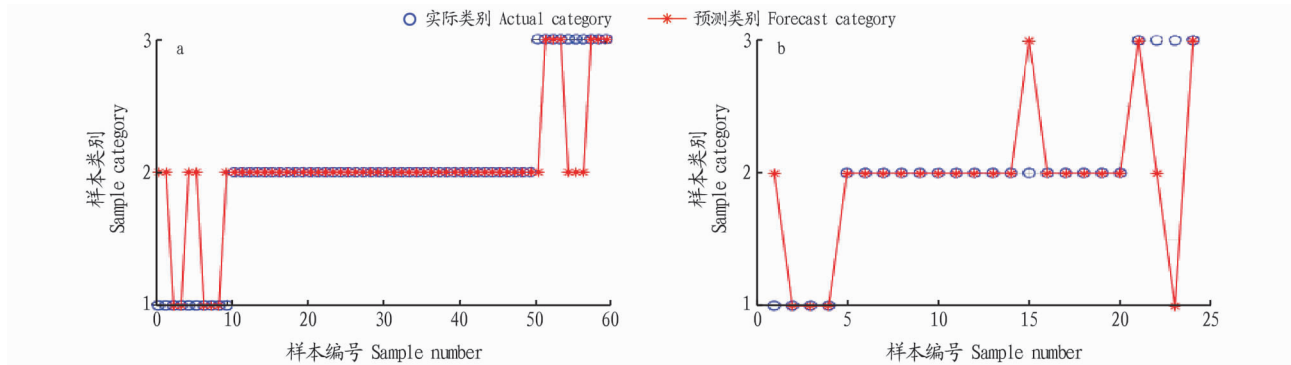
由表1可知,光谱预处理方法为MSC+KPCA的ELM模型获得最优的识别效果,训练集正确识别率为86.67%,测试集正确识别率为83.33%。这主要是因为相较于导数处理, MSC和SNV可消除或减弱光散射引入的基线漂移,更适用于固态样品近红外漫反射光谱的去噪处理;MSC+KPCA较SNV+KPCA处理下ELM模型识别率高的主要原因是MSC+KPCA的累计贡献率较大,ELM输入变量中包含更多的原始变量信息。MSC+KPCA处理下ELM模型的具体识别结果如图2所示。从图2可以看出,训练集有8个样品判错,其中有5个属于纯牛肉的错判给了掺假牛肉,3个属于纯猪肉的错判给了掺假牛肉;测试集有4个样品判错,其中有1个属于纯牛肉的错判给了掺假牛肉,有1个属于掺假牛肉的错判给了纯猪肉,有2个属于纯猪肉的分别错判给了掺假牛肉和纯牛肉。

3 结论

该研究采用傅立叶变换近红外结合极限学习机(ELM)构建纯牛肉、牛肉中掺假猪肉、纯猪肉的快速识别模型,考察

标准正态变换(SNV)、多元散射校正(MSC)、一阶导数(FD)及二阶导数(SD)结合核主成分分析(KPCA)等光谱预处理方法对 ELM 模型预测性能的影响。结果显示, MSC+KPCA

预处理下, ELM 模型的预测效果最优, 训练集及测试集的正确识别率分别为 86.67% 和 83.33%。可见近红外光谱技术结合 ELM 在牛肉中猪肉掺假的快速鉴定方面具有很大潜力。



注:纵坐标上的“1”表示纯牛肉;“2”表示牛肉掺假猪肉;“3”表示纯猪肉

Note:“1” on the ordinate indicates pure beef;“2” indicates beef adulteration with pork;“3” indicates pure pork.

图2 MSC+KPCA 预处理下 ELM 模型训练集(a)和测试集(b)识别结果

Fig. 2 Identification results of the training set(a) and testing set(b) of ELM model under MSC+KPCA preprocessing

参考文献

[1] OECD, Food, Agriculture Organization of the United Nations. OECD-FAO agricultural outlook 2018–2027[M]. Paris: OECD, 2018.

[2] 李婷婷, 张桂兰, 赵杰, 等. 肉及肉制品掺假鉴别技术研究进展[J]. 食品安全质量检测学报, 2018, 9(2): 409–415.

[3] 张娟, 张申, 张力, 等. 电子鼻结合统计学分析对牛肉中猪肉掺假的识别[J]. 食品科学, 2018, 39(4): 296–300.

[4] ROPODI A I, PAVLIDIS D E, MOHAREB F, et al. Multispectral image analysis approach to detect adulteration of beef and pork in raw meats[J]. Food research international, 2015, 67: 12–18.

[5] GRASSI S, ALAMPRESE C. Advances in NIR spectroscopy applied to process analytical technology in food industries[J]. Current opinion in food science, 2018, 22: 17–21.

[6] ALAMPRESE C, CASALE M, SINELLI N, et al. Detection of minced beef

adulteration with turkey meat by UV-vis, NIR and MIR spectroscopy[J]. LWT-Food Science and Technology, 2013, 53(1): 225–232.

[7] MORSY N, SUN D W. Robust linear and non-linear models of NIR spectroscopy for detection and quantification of adulterants in fresh and frozen-thawed minced beef[J]. Meat science, 2013, 93(2): 292–302.

[8] 孟一, 张玉华, 王家敏, 等. 基于近红外光谱技术快速识别不同动物源肉品[J]. 食品科学, 2014, 35(6): 156–158.

[9] 张丽华, 郝雨花, 李顺峰, 等. 基于支持向量机的近红外光谱技术快速鉴别掺假羊肉[J]. 食品工业科技, 2015, 36(23): 289–293.

[10] 韩方凯, 张雪柯, 吕日琴, 等. 基于极限学习机的高品质葡萄酒等级识别模型[J]. 宿州学院学报, 2018, 33(2): 101–104.

[11] 梁晓燕, 郭中华, 钱文瑶, 等. 基于高光谱和极限学习机的冷鲜羊肉颜色无损检测[J]. 食品工业科技, 2016, 37(24): 69–73.

[12] 黎源鸿, 王红军, 邓建猛, 等. 基于 PCA-ELM 和光谱技术预测香蕉成熟度[J]. 现代食品科技, 2017, 33(10): 268–274.

(上接第 148 页)

不适用于砖红壤上, 或是由于其水解吸附特性, 可能在苗期很难表现出其肥效优势, 需要在不同土壤上或是进行作物整个生长期的肥效试验来进一步验证聚磷酸铵的肥效, 为农业生产提供更多参考与指导。

参考文献

[1] 黄昌勇. 土壤学[M]. 北京: 中国农业出版社, 2000: 196–202.

[2] 邱燕, 张鼎华. 南方酸性土壤磷素化学研究进展[J]. 福建稻麦科技, 2003, 21(3): 14–17.

[3] DOMINGO S. Improvement of polyphosphates as soil fertilizers[D]. Manhattan: Kansas State University, 1985.

[4] BLANCHARD R W, HOSSNER L R. Hydrolysis and sorption reactions of orthophosphate, pyrophosphate, tripolyphosphate, and trimetaphosphate anions added to an Elliot soil[J]. Soil science society of America journal, 1969, 33(1): 141–144.

[5] STROECHLEIN J I, SABET S A, CLEMENTIZ D M. Response of plants to polyphosphate on calcareous soils[J]. Agronomy journal, 1968, 60(5): 576–577.

[6] SUBBARAO Y V, ROSCOE E. Reaction products of polyphosphates and orthophosphates with soils and influence on uptake of phosphorus by plants[J]. Soil science society of America journal, 1975, 39(6): 1085–1088.

[7] SUTTON C D, GUNARY D, LARSEN S. Pyrophosphate as a source of phosphorus for plants. II. Hydrolysis and initial uptake by a barley crop[J]. Soil science, 1966, 101(3): 199–209.

[8] RHUE R D, HENSEL D R, YUAN T L, et al. Ammonium orthophosphate and ammonium polyphosphate as sources of phosphorus for potatoes[J]. Soil science society of America journal, 1981, 45: 1229–1233.

[9] GHOSH G K, MOHAN K S, SARKAR A K. Characterization of soil-fertilizer P products and their evaluation as source of P for grams(*Cicer arietinum* L.)[J]. Nutrient cycling in agroecosystems, 1996, 46(1): 71–79.

[10] 李杰, 石元亮, 陈智文. 我国南方红壤磷素研究概况[J]. 土壤通报, 2011, 42(3): 763–768.

[11] 鲁如坤. 土壤农业化学分析方法[M]. 北京: 中国农业科技出版社, 1999: 296–336.

[12] 邢月华, 汪仁, 包红静, 等. 不同磷肥用量对玉米产量·效益及养分吸收的影响[J]. 安徽农业科学, 2011, 39(32): 19834–19835, 19923.

[13] 赵靓, 侯振安, 李水仙, 等. 磷肥用量对土壤有效磷及玉米产量和养分吸收的影响[J]. 玉米科学, 2014, 22(2): 123–128.

[14] DOBSON J W, WELLS K L, FISHER C D. Agronomic effectiveness of ammonium orthophosphate and ammonium polyphosphate as measured by ion uptake and yield of corn[J]. Soil science society of America journal, 1970, 34(2): 323–326.

[15] 陈日远, 代明, 侯文通, 等. 聚磷酸铵对玉米幼苗吸收磷、锌养分及生长的影响[J]. 江苏农业科学, 2014, 42(5): 104–106.