

基于可见-近红外光谱的鸡只粪便分类判别

周 敏，崔志航，单 蕾

(中国农业大学工学院，北京 100083)

摘 要：鸡只粪便的性状是反映鸡只健康状况的重要特征之一，不同性状的鸡便往往与特定的疾病相关联。针对鸡只粪便性状主要依靠人工监测，存在速度慢且易发生交叉感染等问题，该研究提出一种基于可见-近红外光谱技术的鸡只粪便分类判别模型。首先，通过扫描 4 类典型的鸡只粪便样本（正常粪便、红血丝粪便、绿色粪便和饲料粪便）在 400~900 nm 波段范围的光谱数据，对每一类别的鸡便样本按随机性原则以 3:1 划分为校正集和测试集。其次，分别采用多元散射校正、SG 卷积平滑和标准差标准化进行数据预处理，并建立偏最小二乘判别分析模型，根据模型评价指标确定最优预处理方法。然后，使用主成分分析、竞争性自适应重加权采样、改进的混合蛙跳 3 种方法对预处理后的样本进行数据降维，并最终建立分类判别模型。结果表明：基于模型评价指标确定最优数据预处理方法后，再采用改进后的混合蛙跳降维方法建立的判别模型区分正常粪便、红血丝粪便、绿色粪便表现最优，测试集判别准确率分别为 92.27%、92.59%、100%；而对于饲料粪便，所选 3 种降维方法建立的判别模型，其测试集准确率均可达 100%。因此，通过可见-近红外光谱检测手段，结合特征波长优选与偏最小二乘判别分析，可以有效判别不同类型的鸡只粪便，为实现鸡病智能化监测提供技术支持。

关键词：光谱；判别分析；鸡只粪便；偏最小二乘判别分析；特征波长优选；分类判别；可见-近红外光谱

doi: 10.11975/j.issn.1002-6819.202304145

中图分类号: S858.31

文献标志码: A

文章编号: 1002-6819(2023)-14-0200-07

周敏, 崔志航, 单蕾. 基于可见-近红外光谱的鸡只粪便分类判别[J]. 农业工程学报, 2023, 39(14): 200-206.

doi: 10.11975/j.issn.1002-6819.202304145 <http://www.tcsae.org>

ZHOU Min, CUI Zhihang, SHAN Lei. Classification and discrimination of chicken feces using visible-near infrared spectroscopy[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2023, 39(14): 200-206. (in Chinese with English abstract) doi: 10.11975/j.issn.1002-6819.202304145 <http://www.tcsae.org>

0 引 言

鸡只粪便与鸡只健康状况有着紧密联系，不同颜色、性状的粪便与不同的鸡病相关联^[1-2]，因此鸡便是畜禽养殖人员最为关注的对象。目前，鸡便的监测主要依靠有经验的工作人员或兽医人工巡检进行判断，主观性强，准确性低，且易引发人与动物的交叉感染。因此，十分有必要采用先进的检测技术对鸡的粪便进行分析判断，以提高鸡便判别效率和准确性，为智能化、无人化养殖^[3-7]提供技术支持。

可见-近红外光谱检测技术因其简单、快速、准确等优点^[8-10]，近年来已在动物疾病诊断、动物生产信息和动物组分检测^[11-13]方面取得了一定的研究进展。梁亮等^[14]获取粪便的可见-近红外光谱数据并建立判别模型对高山麝慢性肠炎进行诊断，诊断准确率达 93.3%，疾病检出率达 100%。TOLLESON 等^[15]对牛和马在蜱虫轻微和严重感染状态下所产生的粪便进行分析，提出一种通过功能性近红外光谱成像技术手段分析蜱虫感染严重

程度的判定方法。WALKER 等^[16]对山羊粪便的近红外光谱研究发现，山羊的品种、年龄、性别都会对粪便光谱产生影响。ATANASSOVA 等^[17]用近红外光谱技术（波长为 400~1 100 nm）检测牛奶样品中的体细胞数量并通过偏最小二乘法建立变量数为 10 的判别模型，判定奶牛是否患有乳腺炎。在动物生产信息方面，LI 等^[18]采集绵羊粪便光谱数据并建立偏最小二乘模型和逐步回归模型检测绵羊日粮的质量，试验表明绵羊个体差异对预测模型的判别结果影响很小。JANCEWICZ 等^[19]指出使用近红外光谱检测牛粪便中的蛋白质浓度，可用于监测牛的生长状况。LANDAU 等^[20]利用山羊粪便的近红外光谱构建模型，实现了通过粪便检测对山羊饲料的化学成分百分比和成分组成的预测。JOHNSON 等^[21]应用近红外反射光谱采集粪便样本信息，分析并预测肉牛日粮营养特征及自愿采食量。而在动物粪便组分检测方面，梁浩等^[22]设计了近红外实时在线检测装置用以检测畜禽粪污降解过程中产生的挥发性脂肪酸。GALVEZ-SOLA 等^[23]指出可利用近红外光谱测定粪肥中的电导率、pH 值、总有机质等一系列指标和宏量营养素含量（N、P、K）等，实现对粪便堆肥质量的判别。孙迪等^[24]对处理后的奶牛粪水样本采集其近红外光谱数据，建立粪水氮磷季节内预测模型，表明季节因素对粪水中氮磷含量的模型预测结果存在不同程度的影响。SAEYS 等^[25]应

收稿日期: 2023-04-18 修订日期: 2023-06-19

基金项目: 国家重点研发计划（国际合作）: 规模化鸡舍环境和集群健康智能感知机器人技术研发与应用（2017YFE2122200）资助

作者简介: 周敏, 博士, 副教授, 研究方向为智能农机装备。

Email: zhoumin2016@cau.edu.cn

用可见-近红外光谱技术建立了猪粪中各组成成分预测模型，对 K、Ca 含量实现了定性分析，对有机质、干物质、NH₃、N、P、Mg 等实现了定量测量。

上述研究表明，通过可见-近红外光谱检测动物粪便可定性或定量地获取动物的生理信息，但目前未见针对鸡只粪便进行检测从而为鸡病诊断提供依据的研究。为此，本文提出采用可见-近红外光谱技术解决鸡只粪便的分类判别问题，优选数据预处理方法、数据降维方法，对 4 类典型的鸡只粪便分别建立最佳分类判别模型，实现 4 类典型鸡便的快速、准确判别，以期为智能化、无人化的智慧养殖提供技术基础。

1 材料与方法

1.1 仪器设备

试验采用美国 Ocean Optics Inc.S 公司的可见-近红外光纤光谱仪（USB2000+），光谱数据采集范围为 339~1 024 nm，光源分辨率为 1.3 nm，光谱数据点数为 2 048 个。光源采用 PHILIPS 35W12 V 卤素光源，光纤采用优等抗紫外曝光光纤，光纤前端加装 74-UV 光纤准直透镜。使用时将样品和光纤置于暗箱中，并将光源打开预热 30 min 以上，待光谱数据稳定后进行黑板校正。数据处理使用 MATLAB 2018b，整套采集装置如图 1 所示。

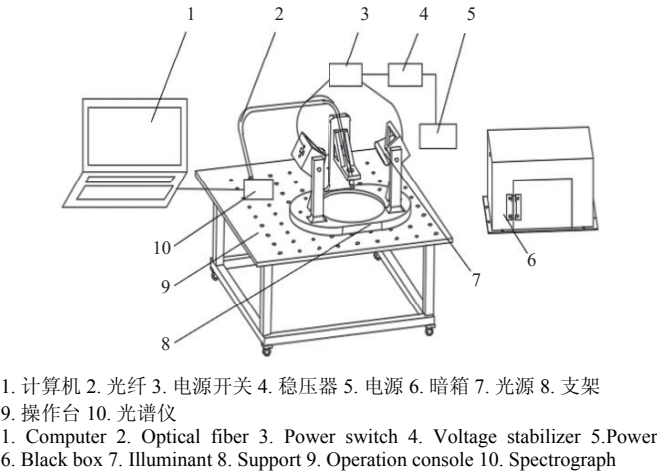


图 1 光谱采集试验平台
Fig.1 Spectral acquisition experiment platform

1.2 数据采集与标定

本试验所需样本均来源于河北省唐山市一蛋鸡养殖场，场内鸡只采用叠层笼养方式养殖，鸡粪利用鸡笼下方的传送带自动清理，随清粪带的运转可采集到一层鸡笼内不同鸡只的粪便。采样开始时鸡只平均日龄为 250 d 左右，采样共持续 14 d。每日清粪带首次运转时，采集传送带上的正常和异常的粪便样本，将其放入玻璃培养皿中并标注每份样本的取样日期，获取到的粪便样本在 2 h 内完成光谱数据采集。为保证原始样本的一致性，采集光谱数据前用干净镊子将样本中的明显干扰物（如羽毛）去除，每个样本每次采集 2 048 个光谱数据点，采集 3 次，取 3 次的平均值作为最终数据，去除首

尾光谱噪声后保留 1 732 个数据点。

根据粪便的形状、颜色、水分等表面特征，在兽医的协助下，对每份样本进行标定。最终共获取正常鸡只粪便样本 105 份，异常粪便样本 110 份。异常粪便样本进一步分为 3 类：红血丝粪便 48 份、绿色粪便 30 份、饲料粪便 32 份。其中，红血丝鸡便是由于肠道受细菌感染，肠粘膜增厚、脱落、出血，造成粪便呈现红色水样稀便；绿色鸡便则是因为病毒和细菌感染导致肝脏解毒功能加快，肝脏回收胆囊胆汁功能受损，引起胆汁随粪便排出并将粪便染成绿色；饲料鸡便的出现存在多种原因，鸡群热应激、肠道菌群紊乱、鸡只感染小肠球虫都会引起饲料在肠道中未完全消化就排出体外，形成“饲料颗粒粪便”的现象。4 类典型粪便样本如图 2 所示，其具体表征如表 1 所示。

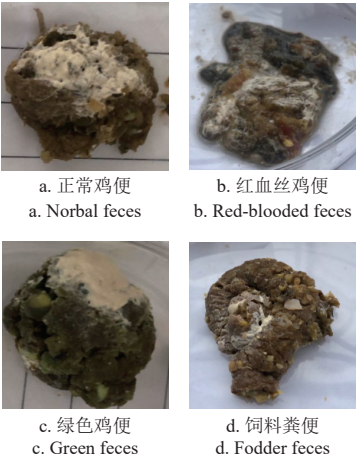


图 2 4 类粪便样本图
Fig.2 Four type of feces samples

表 1 鸡只粪便性状对比
Table 1 Comparison of chicken feces traits

样品 Samples	硬度 Hardness	形状 Shape	颜色 Color
正常粪便 Normal feces	软硬适中	堆状或粗条状	整体为灰色或灰褐色，粪便的顶端有少量的白色尿酸盐
红血丝鸡便 Red-blooded feces	偏稀状	滩状	灰褐色为主体，有红色血丝，部分可观测到混有肠粘膜
绿色鸡便 Green feces	软硬适中	堆状或粗条状	整体色彩呈现绿色、黄绿色
饲料鸡便 Fodder feces	软硬适中	堆状或粗条状	整体呈黄褐色，粪便内部掺杂有明显的未经消化的饲料

为保证校正集与测试集中两类样本分布的均匀性，对每一类粪便样本的光谱数据，按照随机性原则，以 3：1 的比例，划分校正集和测试集，即校正集 162 个样本，测试集 53 个样本，且校正集和测试集均包含 4 类典型鸡便。校正集样本用于建立判别模型，测试集样本用于检验判别模型的准确性。粪便样本的具体情况如表 2 所示。

表 2 鸡只粪便样本数据统计
Table 2 Statistics of chicken feces samples

序号 Sequence	类别 Categories	校正集 Calibration set	测试集 Test set	总计 Total
1	正常	79	26	105
2	异常	83	27	110
总计 Total		162	53	215

1.3 数据处理方法

将4种典型粪便样本的光谱数据分别采用无处理、多元散射校正 (multivariate scattering correction, MSC)、SG 卷积平滑 (Savitzky-Golay convolution smoothing) 和标准差标准化 (Z-score) 4 种方法进行数据预处理。偏最小二乘判别分析 (partial least squares discriminant analysis, PLS-DA) 方法^[26]是有监督的分类机器学习方法中的一种, 可以很好的区分组间差异性较小的样本。其通过在原始样本数据和所属类别间建立响应函数, 来预测未知样本类别, 适用于样本数量少, 变量数多的样本集。因此, 采用 PLS-DA 方法建立分类判别模型, 根据评价标准选择各类型粪便最适合的预处理方法。将预处理完成的数据分别通过主成分分析 (principal component analysis, PCA)、竞争性自适应重加权采样 (competitive adaptive reweighted sampling, CARS) 和改进的混合蛙跳算法 (improved shuffled frog leaping algorithm, ISFLA) 3 种变量优选方法^[27-29]进行数据降维, 以对模型进行轻量化改进, 根据模型评价指标确定各类型粪便最优的数据降维方法, 最终建立基于 PLS-DA 的鸡只粪便分类判别轻量化模型。

1.4 模型评价

所建立的分类判别模型采用受试者工作特征 (receiver operating characteristic, ROC) 曲线^[30]结合判别准确率的方法进行综合评判。其中 ROC 曲线评价法包含3个指标, 分别为敏感性、特异性、曲线下面积 (area under curve, AUC) 值。敏感性代表被正确预测的正实例样本占有所有正实例样本的比例, 特异性代表被正确预测的负实例样本占有所有负实例样本的比例。AUC 值用来表征模型分类性能的好坏。对于所有指标来说, 其值均分布在0和1之间, 且值越趋近于1, 代表模型分类效果越好。

2 结果与分析

2.1 鸡只粪便反射光谱特征

图3所示为3种异常鸡便在400~900 nm波段的反射光谱曲线, 红色、绿色和蓝色分别为红血丝粪便、绿色粪便和饲料粪便的曲线。从图3可知, 从反射强度来看, 饲料粪便的光强明显高于其他2种类型。由于采集数据过程中, 样本为非定量采集, 所以反射强度的差异是由样本体积差异造成的, 故不能将反射强度作为区分粪便类型的依据。从3种类别粪便的光谱曲线走势看, 波峰波谷出现位置基本一致, 无法直接进行类别区分。因此, 需要对光谱数据进行进一步的处理。

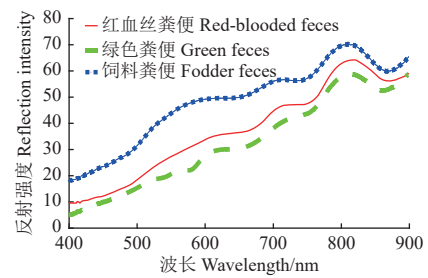


图3 3种类别粪便反射光谱曲线

Fig.3 Three types of stool reflectance spectrum curve

2.2 光谱数据预处理

对采集到的215份鸡只粪便样本, 分别采用无处理、SG 卷积平滑、MSC、SG 卷积平滑-MSC-Z-Score 标准化4种方法进行数据预处理。杜树新等^[31]研究中提出SG平滑多项式的阶次通常为奇数, 对平滑效果影响并不显著, 但平滑的窗宽大小对应平滑效果有一定的影响, 窗宽过小起不到平滑效果, 窗宽过大增加了计算复杂度。综合考虑模型性能影响, SG平滑中多项式拟合的阶数为3, 平滑窗口宽度设置为5。对于4种不同类型的粪便建立PLS-DA模型进行判别分析, 综合多项指标评价判别模型的优劣, 结果如表3所示。

表3 不同预处理方法模型判别结果

Table 3 Discriminant results of models by different preprocessing methods

粪便类型 Categories	光谱预处理方法 Spectral preprocessing methods	校正集 Calibration set				测试集 Test set			
		敏感性 Sensitivity	特异性 Specificity	AUC	准确率 Accuracy/%	敏感性 Sensitivity	特异性 Specificity	AUC	准确率 Accuracy/%
正常粪便 Normal feces	无处理	0.81	1.00	0.92	88.79	0.54	0.71	0.84	75.93
	SG 卷积平滑	0.85	0.89	0.94	86.79	0.84	0.65	0.84	74.07
	MSC	0.85	0.78	0.91	81.13	0.75	0.74	0.85	74.07
	SG-MSC-Z-scores	0.88	0.89	0.95	88.68	0.85	0.84	0.90	84.57
红血丝粪便 Red-blooded feces	无处理	0.50	0.87	0.88	97.56	0.44	0.97	0.90	85.18
	SG 卷积平滑	0.50	0.98	0.87	86.79	0.39	0.95	0.90	82.72
	MSC	0.83	0.95	0.96	92.45	0.81	0.95	0.98	91.98
	SG-MSC-Z-scores	0.50	0.93	0.88	83.02	0.50	0.93	0.88	83.02
绿色粪便 Green feces	无处理	0.43	1.00	0.92	92.45	0.26	0.97	0.74	87.04
	SG 卷积平滑	0.14	1.00	0.86	88.68	0.13	0.99	0.78	86.42
	MSC	0.86	0.98	1.00	98.23	0.91	0.99	1.00	96.15
	SG-MSC-Z-scores	0.29	0.98	0.92	88.68	0.30	0.99	0.90	88.89
饲料粪便 Fodder feces	无处理	0.13	1.00	0.79	86.79	0.29	1.00	0.73	89.51
	SG 卷积平滑	0.25	1.00	0.94	88.68	0.38	0.99	0.81	88.12
	MSC	1.00	1.00	1.00	100.00	1.00	1.00	1.00	100.00
	SG-MSC-Z-scores	0.38	1.00	0.89	90.57	0.29	1.00	0.85	89.51

注: AUC 为受试者工作特征曲线下的面积; MSC, 多元散射校正。下同。

Note: AUC is the area under receiver operating characteristic (ROC) curve; MSC, multivariate scattering correction. Same below.

由表 3 可知，对于正常粪便而言，经 SG 卷积平滑和 SG-MS-C-Z-scores 处理后模型的测试集敏感性相近，均高于其他 2 种方法；无处理、SG 卷积平滑和 MSC 的 AUC 值相当，均低于 SG-MS-C-Z-scores；而特异性和准确率方面，SG-MS-C-Z-scores 均远远高于其他 3 种方法，因此，SG-MS-C-Z-scores 预处理得到的光谱数据建立的判别模型效果最佳。对于红血丝粪便，经 MSC 处理后，模型测试集的敏感性、AUC 值和准确率均高于其他 3 种方法，尤其敏感性和准确率高出较多；其特异性值 0.95 虽低于无处理的 0.97，但相差不大，故综合考虑，MSC 为最佳预处理方法。对绿色粪便和饲料粪便，经 MSC 处理后，模型测试集的敏感性、特异性、AUC 值和准确率均高于其他 3 种方法，且敏感性和准确率高出较多，故 MSC 为最佳预处理方法。综上，正常粪便选择 SG-MS-C-Z-scores 作为预处理方法；异常粪便，选用 MSC 进行光谱数据的预处理。

2.3 模型优化

为减少冗余波长，改善模型计算性能，在保证判别准确率的前提下，提取用于区分粪便类型的特征波长，减少特征波长数量，实现模型轻量化。对经过预处理的鸡便样本数据分别采用 PCA、CARS 和改进的 ISFLA 方法进行数据降维处理，使原本的 1 732 个特征波长变量大大减少，然后利用降维后的数据建立 PLS-DA 分类判别模型，实现模型的轻量化。其中，相比原 ISFLA 算法，改进的 ISFLA 算法选用五折交叉验证结合 PLS 模型回归的方法，确定了更为精确的主因子数 l ，具体步骤为：将校正集所有样本按照每 5 个一组，分为 32 组，每次取一组作为测试集，其余为训练集，建立 PLS-DA 判别模型并输出判别结果，计算判别误差，最终选取判别误差最小的一组，将此时的 l 作为最终的主因子数；迭代次数由原本根据经验值选定，改为以模型交叉验证的均方根误差值最小为目标确定，正常粪便、红血丝粪便、绿色粪便、饲料粪便的迭代次数由原来的 1 000 分别降为 500、240、300 和 30。

对各个模型的评价参数进行统计，结果见表 4 所示。由表 4 可知，降维后特征波长的数量由 1 732 最低可降

到 3（红血丝粪便和饲料粪便经 PCA 降维），最高也只有 48 个（正常粪便经 CARS 降维），大大降低了特征波长冗余，减少了模型变量。此外，对正常粪便、红血丝和绿色粪便，采用改进的 ISFLA 方法进行数据降维后建立的 PLS-DA 模型在各方面的指标几乎都优于采用其他方法所对应的指标（除正常粪便测试集的敏感性 ISFLA 方法为 0.91，略低于 CARS 方法 0.92）。其中，正常粪便 ISFLA 方法测试集的准确率为 92.27%，比 PCA 和 CARS 方法准确率分别高出 16.80 和 7.36 个百分点；红血丝粪便 ISFLA 方法测试集的准确率为 92.59%，比 PCA 和 CARS 方法准确率分别高出 3.70 和 11.11 个百分点；绿色粪便 ISFLA 方法测试集的准确率为 100%，比 PCA 和 CARS 方法准确率分别高出 11.11 和 3.70 个百分点。可见，ISFLA 方法在敏感性、特异性和 AUC 值均高于其他方法的同时，其准确率也高于其他 2 种算法超过 3 个百分点。尽管 ISFLA 方法的优选特征波长数不是三种方法中最低的，但该方法针对 3 种粪便的特征波长数仍下降了 98.27%、99.19% 和 99.19%。因此，综合来看，改进的 ISFLA 方法为最佳降维方法。对饲料粪便而言，不同降维方法得到的分类模型的 ROC 曲线评价指标均为 1，准确率均为 100%，因此不能通过这些指标筛选模型，但经 PCA 降维后的特征波长数由 1 732 减少到 3，优于其他 2 种模型，故针对饲料粪便，PCA 方法略优。对比表 3 的数据可知，优化后的模型特征波长变量减少的同时，各项性能指标均有所提高或与优化前持平。其中，针对正常粪便、红血丝粪便、绿色粪便，优化后的模型校正集的准确率相比优化前，分别提高了 7.70、0.61 和 3.85 个百分点，而饲料粪便的各项指标均与优化前相同且为最大值。综上，正常鸡便、红血丝粪便和绿色粪便的最优判别模型为 ISFLA-PLS-DA，而饲料粪便的最佳判别模型为 PCA-PLS-DA。由于饲料粪便各模型的差别仅在于降维后的特征波长数不同，且该变量最大值与最小值差别较小，对模型计算性能影响不大，故仍可选择 ISFLA-PLS-DA 作为分类判别模型。因此，为提高判别算法效率，可确定 ISFLA-PLS-DA 为鸡只粪便分类判别最佳模型。

表 4 不同数据降维方法建立 PLS-DA 模型判别结果
Table 4 Discriminant results of PLS-DA models established by different data dimensionality reduction methods

粪便类型 Categories	降维方法 Dimensionality reduction methods	N	l	校正集 Calibration set				测试集 Test set			
				敏感性 Sensitivity	特异性 Specificity	AUC	准确率 Accuracy/%	敏感性 Sensitivity	特异性 Specificity	AUC	准确率 Accuracy/%
正常粪便 Normal feces	PCA	4	3	0.81	0.70	0.90	75.47	0.81	0.70	0.90	75.47
	CARS	48	6	0.94	0.82	0.93	87.65	0.92	0.78	0.92	84.91
	ISFLA	30	14	1	0.96	0.99	98.11	0.91	0.86	0.97	92.27
红血丝粪便 Red-blooded feces	PCA	3	1	1	0.83	0.97	90.36	0.75	1	0.98	88.89
	CARS	36	6	1	0.81	0.97	89.16	0.67	0.93	0.97	81.48
	ISFLA	14	6	1	1	1	100.00	0.83	1	1	92.59
绿色粪便 Green feces	PCA	3	1	0.57	1	0.98	87.95	0.57	1	1	88.89
	CARS	32	5	0.91	0.98	0.99	96.39	0.86	1	0.99	96.30
	ISFLA	14	6	0.96	0.98	1	97.59	1	1	1	100.00
饲料粪便 Fodder feces	PCA	3	1	1	1	1	100.00	1	1	1	100.00
	CARS	5	3	1	1	1	100.00	1	1	1	100.00
	ISFLA	7	6	1	1	1	100.00	1	1	1	100.00

注：N 为特征波长数量，l 为主因子数。PCA，主成分分析；CARS，竞争性自适应重加权采样；ISFLA，改进的混合蛙跳算法。
Note: N is the number of characteristic wavelength and l is the number of primary factor. PCA, principal component analysis; CARS, competitive adaptive reweighted sampling; ISFLA, improved shuffled frog leaping algorithm.

3 结 论

本研究基于可见-近红外光谱技术构建了一套针对 4 类典型鸡只粪便的偏最小二乘判别分析 (partial least squares discriminant analysis, PLS-DA) 模型, 该模型对于不同类型的样本均实现了良好的判别效果。主要结论如下:

1) 针对不同类型的鸡只粪便, 利用无处理、多元散射校正 (multivariate scattering correction, MSC)、SG 卷积平滑 (Savitzky-Golay convolution smoothing) 和标准差标准化 (Z-scores) 4 种数据预处理方法, PLS-DA 有着不同的表现: 正常粪便样本数据经过 SG 卷积平滑-MSC-Z-Scores 方法处理后建立的 PLS-DA 模型表现最佳, 3 类异常粪便样本数据则是通过 MSC 方法处理建立的 PLS-DA 模型表现更佳。

2) 应用改进后的混合蛙跳算法 (improved shuffled frog leaping algorithm, ISFLA) 进行数据降维处理后构建的 ISFLA-PLS-DA 模型在判别 4 类典型鸡只粪便样本时均有良好表现。从测试集准确率的角度看, ISFLA-PLS-DA 在判别正常粪便、红血丝粪便、绿色粪便样本时均有良好表现, 准确率高于其他 2 种算法超过 3 个百分点; 而对于饲料粪便, 该模型仍能与使用主成分分析 (principal component analysis, PCA) 最优模型 PCA-PLS-DA 的准确率持平。

3) 通过不同方法对数据进行降维, 可大大减少特征波长变量数, 同时提升各典型粪便样本校正集的准确率: 正常粪便、红血丝粪便、绿色粪便校正集的准确率相比模型优化前分别提高了 7.70, 0.61, 3.85 个百分点。

【参 考 文 献】

- [1] THORNTON, P K. Livestock production: Recent trends, future prospects [J]. Philosophical Transactions of the Royal Society B: Biological Sciences, 2010, 365: 2853-2867.
- [2] 彭欠欠, 苏佳文, 崔子鹤, 等. 鸡组织滴虫病的病理学诊断[J]. 中国动物传染病学报, 2017, 25(1): 72-74.
PENG Qianqian, SU Jiawen, CUI Zihe, et al. Pathological diagnosis of histomoniasis in chicken[J]. Chinese Journal of Animal Infectious Diseases, 2017, 25(1): 72-74. (in Chinese with English abstract)
- [3] 史兵, 赵德安, 刘星桥, 等. 基于无线传感网络的规模化水产养殖智能监控系统[J]. 农业工程学报, 2011, 27(9): 136-140.
SHI Bing, ZHAO Dean, LIU Xingqiao, et al. Intelligent monitoring system for industrialized aquaculture based on wireless sensor network[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2011, 27(9): 136-140. (in Chinese with English abstract)
- [4] 杨柳, 李保明. 蛋鸡福利化养殖模式及技术装备研究进展[J]. 农业工程学报, 2015, 31(23): 214-221.
YANG Liu, LI Baoming. Research progress of welfare-oriented breeding mode and technical equipments for laying hen[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2015, 31(23): 214-221. (in Chinese with English abstract)
- [5] 李保明, 王阳, 郑炜超, 等. 中国规模化养鸡环境控制关键技术与设施设备研究进展[J]. 农业工程学报, 2020, 36(16): 212-221.
LI Baoming, WANG Yang, ZHENG Weichao, et al. Research progress in environmental control key technologies, facilities and equipment for laying hen production in China[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2020, 36(16): 212-221. (in Chinese with English abstract)
- [6] 杨飞云, 曾雅琼, 冯泽猛, 等. 畜禽养殖环境调控与智能养殖装备技术研究进展 [J]. 中国科学院院刊, 2019, 34(2): 163-173.
YANG Feiyun, ZENG Yaqiong, FENG Zemeng, et al. Research status on environmental control technologies and intelligent equipment for livestock and poultry production[J]. Bulletin of Chinese Academy of Sciences, 2019, 34(2): 163-173 (in Chinese with English abstract)
- [7] 张铁民, 黄俊端. 基于音频特征和模糊神经网络的禽流感病鸡检测[J]. 农业工程学报, 2019, 35(2): 168-174.
ZHANG Tiemin, HUANG Junduan. Detection of chicken infected with avian influenza based on audio features and fuzzy neural network[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2019, 35(2): 168-174. (in Chinese with English abstract)
- [8] 褚小立, 陆婉珍. 近五年我国近红外光谱分析技术研究与应用进展[J]. 光谱学与光谱分析, 2014, 34(10): 2595-2605.
CHU Xiaoli, LU Wanzhen. Research and application progress of near infrared spectroscopy analytical technology in china in the past five years[J]. Spectroscopy and Spectral Analysis, 2014, 34(10): 2595-2605. (in Chinese with English abstract)
- [9] 田燕龙, 王毅, 王箫, 等. 近红外光谱技术在微生物检测中的应用进展[J]. 光谱学与光谱分析, 2022, 42(1): 9-14.
TIAN Yanlong, WANG Yi, WANG Xiao, et al. Advances in detection of microorganisms using near-infrared spectroscopy[J]. Spectroscopy and Spectral Analysis, 2022, 42(1): 9-14. (in Chinese with English abstract)
- [10] 李明, 韩东海, 鲁丁强, 等. 近红外光谱通用模型在农产品及食品检测中的研究进展[J]. 光谱学与光谱分析, 2022, 42(11): 3355-3360.
LI Ming, HAN Donghai, LU Dingqiang, et al. Research progress of universal model of near-infrared spectroscopy in agricultural products and foods detection[J]. Spectroscopy and Spectral Analysis, 2022, 42(11): 3355-3360. (in Chinese with English abstract)
- [11] 严旭, 杜周和, 白史且, 等. 动物粪便近红外光谱的应用潜力[J]. 光谱学与光谱分析, 2015, 35(12): 3382-3387.
YAN Xu, DU Zhouhe, BAI Shiqie, et al. Potential applicability of fecal nirs: A review [J]. Spectroscopy and Spectral Analysis, 2015, 35(12): 3382-3387. (in Chinese with English abstract)
- [12] 黄珏, 王正亮, 李慕雨, 等. 基于电子舌和近红外光谱技术的进口牛肉产地溯源[J]. 中国食品学报, 2021, 21(12): 254-260.
HUANG Yu, WANG Zhengliang, Li Muyu, et al. Origin traceability of imported beef based on electronic tongue and nir spectra[J]. Journal of Chinese Institute of Food Science and

- Technology, 2021, 21(12): 254-260. (in Chinese with English abstract)
- [15] 王巧华, 马逸霄, 付丹丹. 基于光谱技术的禽蛋内部品质无损检测研究进展[J]. 华中农业大学学报, 2021, 40(6): 220-230.
WANG Qiaohua, MA Yixiao, Fu Dandan. Progress of non-destructive detection of poultry egg internal quality based on spectroscopy[J]. Journal of Huazhong Agricultural University, 2021, 40(6): 220-230. (in Chinese with English abstract)
- [14] 梁亮, 刘志霄, 潘世成, 等. 基于粪便可见-近红外反射光谱的高山麝慢性肠炎诊断[J]. 光谱学与光谱分析, 2009, 29(7): 1772-1776.
LIANG Liang, LIU Zhixiao, PAN Shicheng, et al. Diagnosis of chronic enteritis of alpine musk deer (*moschus chrysogaster*) based on visible-near infrared reflectance spectra of feces[J]. Spectroscopy and Spectral Analysis, 2009, 29(7): 1772-1776. (in Chinese with English abstract)
- [15] TOLLESON D R, TEEL P D, STUTH J W, et al. Effects of a lone star tick (*amblyomma americanum*) burden on performance and metabolic indicators in growing beef steers[J]. Veterinary Parasitology, 2010, 173(1/2): 99-106.
- [16] WALKER J, CAMPBELL E, TAYLOR C, et al. Effects of breed, sex, and age on the variation and ability of fecal near-infrared reflectance spectra to predict the composition of goat diets[J]. *Journal of Animal Science*, 2007, 85(2): 518-526.
- [17] ATANASSOVA S, TSENKOVA R, VASU R M, et al. Identification of mastitis pathogens in raw milk by near infrared spectroscopy and SIMCA classification method[J]. Scientific Works of the University of Food Technologies - Plovdiv, 2009, 56(1): 567-572.
- [18] LI H, TOLLESON D, Stuth J, et al. Faecal near infrared reflectance spectroscopy to predict diet quality for sheep[J]. *Small Ruminant Research*, 2007, 68(3): 263-268.
- [19] JANCEWICZ L J, PENNER G, SWIFT M L, et al. Predictability of growth performance in feedlot cattle using fecal near infrared spectroscopy[J]. Canadian Journal of Animal Science, 2017, 95(1): 455-474.
- [20] LANDAU S, GIGER-REVERDIN S, RAPETTI L, et al. Data mining old digestibility trials for nutritional monitoring in confined goats with aids of fecal near infra-red spectrometry[J]. *Small Ruminant Research*, 2008, 77(2/3): 146-158.
- [21] JOHNSON J R, CARSTENS G E, PRINCE S D, et al. Application of fecal near-infrared reflectance spectroscopy profiling for the prediction of diet nutritional characteristics and voluntary intake in beef cattle[J]. *Journal of Animal Science*, 2017, 95(1): 447-454.
- [22] 梁浩, 黄圆萍, 沈广辉, 等. 在线旁路近红外实时监测粪污厌氧发酵挥发性脂肪酸含量[J]. 农业工程学报, 2020, 36(10): 220-226.
LIANG Hao, HUANG Yuanping, SHEN Guanghui, et al. Near-infrared real-time online bypass detection of volatile fatty acids in anaerobic fermentation of manure[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2020, 36(10): 220-226. (in Chinese with English abstract)
- [23] GALVEZ-SOLA L, MORAL R, PEREZ-MURCIA M D, et al. The potential of near infrared reflectance spectroscopy for the estimation of agroindustrial compost quality[J]. *Science of the Total Environment*, 2010, 408(6): 1414-1421.
- [24] 孙迪, 杨仁杰, 李梦婷, 等. 春秋季节对近红外光谱模型预测奶牛场粪水氮磷含量结果的影响[J]. 农业工程学报, 2020, 36(10): 197-205.
SUN Di, YANG Renjie, LI Mengting, et al. Influences of spring and autumn on the nitrogen and phosphorus contents of the slurry predicted by near-infrared spectrum model on dairy farms[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2020, 36(10): 197-205. (in Chinese with English abstract)
- [25] SAEYS W, MOUAZEN A M, RAMON H, et al. Potential for onsite and online analysis of pig manure using visible and near infrared reflectance spectroscopy[J]. *Biosystems Engineering*, 2005, 91(4): 393-402.
- [26] 韩娅红, 蹇林, 杜瑞铭, 等. 基于骨蛋白拉曼光谱特异性的肉骨粉种属鉴别方法[J]. 农业工程学报, 2020, 36(16): 267-273.
HAN Yahong, JIAN Lin, DU Ruiming, et al. Species-specific identification of meat and bone meal based on Raman spectral analysis of bone protein[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2020, 36(16): 267-273. (in Chinese with English abstract)
- [27] 孔德明, 董瑞, 崔耀耀, 等. 基于三维荧光光谱结合二维线性判别分析的油类识别方法的研究[J]. 光谱学与光谱分析, 2021, 41(8): 2505-2510.
KONG Deming, DONG Rui, CUI Yaoyao, et al. Study on oil identification method based on three-dimensional fluorescence spectrum combined with two-dimensional linear discriminant analysis[J]. Spectroscopy and Spectral Analysis, 2021, 41(8): 2505-2510. (in Chinese with English abstract)
- [28] 陈笑, 宦克为, 赵环, 等. 基于变量频次加权自助采样法的近红外光谱变量选择方法研究[J]. 分析化学, 2021, 49(10): 1743-1749.
CHEN Xiao, HUAN Kewei, ZHAO Huan, et al. Variable selection of near infrared spectroscopy based on variable frequency weighted bootstrap sampling[J]. Chinese Journal of Analytical Chemistry, 2021, 49(10): 1743-1749. (in Chinese with English abstract)
- [29] 雷德明, 王甜. 基于改进蛙跳算法的分布式两阶段混合流水车间调度[J]. 控制与决策, 2021, 36(1): 241-248.
LEI Deming, WANG Tian. An improve shuffled frog leaping algorithm for the distributed two-stage hybrid flow shop scheduling[J]. Control and Decision, 2021, 36(1): 241-248. (in Chinese with English abstract)
- [30] VERBAKEL J Y, STEYERBERG E W, UNO H, et al. ROC curves for clinical prediction models part 1. ROC plots showed no added value above the AUC when evaluating the performance of clinical prediction models[J]. *Journal of Clinical Epidemiology*, 2020, 126: 207-216.
- [31] 杜树新, 杜阳锋, 武晓莉. 基于 Savitzky-Golay 多项式的三维荧光光谱的曲面平滑办法[J]. 光谱学与光谱分析, 2011, 31(2): 440-443.
DU Shuxin, DU Yangfeng, WU Xiaoli. The surface smoothing methods for three-dimensional fluorescence spectrometry based on Savitzky-Golay polynomial smoothing[J]. Spectroscopy and Spectral Analysis, 2011, 31(2): 440-443. (in Chinese with English abstract)

Classification and discrimination of chicken feces using visible-near infrared spectroscopy

ZHOU Min, CUI Zhihang, SHAN Lei

(College of Engineering, China Agricultural University, Beijing 100083, China)

Abstract: Chicken disease identification with manual monitoring cannot fully meet the large-scale production in livestock and poultry breeding industry, due to slow speed and chickens prone to cross-infection. The characteristics of chicken feces can be one of the important indicators to reflect the health status of chickens. The different colors and traits of feces are associated with different chicken diseases. In this study, a classification and discrimination model was proposed for the chicken feces using visible-near-infrared spectroscopy. A research foundation was laid for the final realization of early warning of chicken diseases. Four types of typical chicken feces sample were selected, including the normal, red blood streak, green and fodder feces. The spectral data of samples was scanned in the 400-900 nm band. Each type of chicken feces samples was divided into the correction sets (162 samples) and test sets (53 samples) at 3:1, according to the principle of randomness. Multivariate scattering correction, Savitzky-Golay convolution smoothing, and Z-scores normalization were used to preprocess the data. A PLS-DA (partial least squares discriminant analysis) model was then established to select the optimal pretreatment, according to the evaluation index. An improved shuffled frog leaping algorithm was proposed to further optimize the existing partial least squares discriminant analysis model. The calculation rate of the model was effectively improved to obtain a more accurate number of principal factors. The number of iterations was determined to minimize the root mean square error of cross-validation. The PCA (principal component analysis), CARS (competitive adaptive reweighted sampling), and ISFLA (improved shuffled frog leaping algorithm) were used to reduce the data dimensionality of the processed samples using the preferred pretreatment. Finally, a categorical discriminant model was established after optimization. The results showed that the PLS-DA model established by the SG convolution smoothing-MS-C-Z-Scores shared the better performance of normal stool sample data, and for the three types of abnormal samples, the PLS-DA model established by MSC performed better. The accuracies of the test sets were achieved at 74.07%, 91.98%, 96.15%, and 100%, respectively, for four types of feces samples. The optimal data preprocessing was determined using model evaluation. The ISFLA-PLS-DA was used to distinguish between the normal, red-blooded feces and green feces, where the accuracies of the test set were 92.27%, 92.59% and 100%, respectively. More importantly, the PCA-PLS-DA should be used for the discrimination in the fodder feces, where the accuracy of the test set was 100%. The dimensionality of the data was reduced by feature extraction to construct the lightweight PLS-DA model. The accuracy of the correction set was significantly improved in each typical stool sample. Specifically, the accuracies of the correction set were significantly improved before optimization, which increased by 7.70, 0.61, and 3.85 percentage point, respectively, of normal, red-blooded feces and green feces. Therefore, rapid, accurate and non-contact detection was achieved in the different types of typical chicken feces using visible-near-infrared spectroscopy detection combined with the characteristic wavelength selection and PLS-DA discrimination. The finding can also provide the strong reference for the intelligent detection technology of chicken disease.

Keywords: spectroscopy; discriminant analysis; chicken feces; partial least squares-discriminant analysis; characteristic wavelength selection; classification discrimination; visible-near infrared spectroscopy