

红外光谱结合统计学方法对纸质包装盒的检验研究

汪子拓¹, 姜红^{2,3*}, 谢皓东⁴, 杨棋驭¹, 徐沐暄⁴

(1.中国人民公安大学 侦查学院, 北京 102600; 2.万子健检测技术(北京)有限公司司法鉴定中心, 北京 100141; 3.食品药品安全防控山西省重点实验室, 太原 030006; 4.中国人民公安大学 信息网络安全学院, 北京 102600)

摘要: **目的** 建立一种高效准确的红外光谱检验纸质包装盒的研究方法。**方法** 利用傅里叶变换红外光谱仪测得了 56 个不同类型不同来源的纸质包装盒的红外光谱数据; 根据纸盒主要填料的不同, 将样品初步分为三大类; 利用主成分分析对初步分类后的数据降维, 提取出 4 个主成分, 再利用系统聚类将样品最终分为 6 组, 使用 K-means 聚类算法结合 Pearson 卡方检验进行验证, 与最终分类的结果基本吻合。基于该分组, 训练朴素贝叶斯分类、随机森林模型、XGBoost 分类 3 种判别模型, 实现对新样品组别的分类预测。**结果** 56 个快递包装纸盒样品被分为 3 类, 而后进一步细分为 6 组, 3 种判别预测模型均有较高的准确率, 其中随机森林模型的准确率最高。**结论** 该方法快速方便地实现了对样品的区分, 并且可以实现无损检验, 为犯罪现场纸质包装盒的鉴别提供依据, 从而为公安侦查工作的开展提供帮助。

关键词: 纸质包装盒; 傅里叶变换红外光谱法; Pearson 卡方检验; 朴素贝叶斯分类; 随机森林模型; XGBoost 分类

中图分类号: TB484.1; O657.33; D918.2 文献标志码: A 文章编号: 1001-3563(2024)09-0178-07
DOI: 10.19554/j.cnki.1001-3563.2024.09.023

Infrared Spectroscopy Combined with Statistical Methods on Inspection of Paper Packaging Box

WANG Zituo¹, JIANG Hong^{2,3*}, XIE Haodong⁴, YANG Qiyu¹, XU Muxuan⁴

(1. School of Investigation, People's Public Security University of China, Beijing 102600, China;

2. Judicial Appraisal Center of Wanzijian Testing Technology Co., Ltd., Beijing 100141, China;

3. Shanxi Key Laboratory of Food and Drug Safety Prevention and Control, Taiyuan 030006, China;

4. School of Information Network Security, People's Public Security University of China, Beijing 102600, China)

ABSTRACT: The work aims to establish an efficient and accurate research method for infrared spectroscopy inspection of paper packaging boxes. Infrared spectral data of 56 paper packaging boxes of different types and sources were measured with a Fourier transform infrared spectrometer. According to the different main fillers of the paper box, the samples were preliminarily divided into three categories. Principal component analysis was conducted to reduce the dimensionality of the preliminarily classified data. Four principal components were extracted, and the samples were finally divided into six groups through systematic clustering. The K-means clustering algorithm combined with Pearson's chi-squared test was used for validation. The results were basically consistent with the final classification. Based on this grouping, three discriminative models, namely Naive Bayesian classification, random forest model, and XGBoost classification, were trained to achieve classification prediction for new sample groups. The 56 samples of express paper packaging boxes were divided into 3 categories and further subdivided into 6 groups. All

收稿日期: 2023-09-08

基金项目: 食品药品安全防控山西省重点实验室基金

*通信作者

three discriminant prediction models had high accuracy, with the random forest model having the highest accuracy. This method quickly and conveniently distinguishes samples, and can achieve non-destructive testing, providing a basis for the identification of paper packaging boxes at crime scenes, thus providing assistance for the development of public security investigation work.

KEY WORDS: paper packaging box; Fourier transform infrared spectroscopy; Pearson's chi-squared test; Naive Bayesian classification; random forest model; XGBoost classification

纸质包装盒是犯罪现场常见的物证,构建鉴别纸质包装盒的新方法可以为公安工作侦破案件提供新思路。近年来随着网络购物的迅速发展,各式各样的新型纸质包装盒在案发现场层出不穷。目前,研究纸质物的主要方法有 X 射线荧光光谱法^[1]、高光谱法^[2]、拉曼光谱法^[3]、红外光谱法^[4]、差分拉曼光谱法^[5]等。傅里叶红外光谱检验具有信噪比高、重现性好、扫描速度快、分辨率高等优势,且样品不需要预处理,方便快捷。徐冰冰等^[6]利用一维红外光谱、二维红外光谱等方法对普通和添加防水防油剂的食品包装牛皮纸进行了检验。本实验使用傅里叶变换红外光谱检验了普遍

意义上的纸质包装盒,并结合多元统计学方法,如系统聚类法、K-means 聚类算法、Pearson 卡方检验对实验数据进行处理,再使用朴素贝叶斯分类、随机森林模型、XGBoost 分类构建多元判别预测模型。

1 实验装置

1.1 主要来源

不同来源、不同种类的纸质包装盒 56 个,样品信息见表 1。

表 1 样品信息
Tab.1 Sample information

样品 编号	名称	来源	颜色	样品 编号	名称	来源	颜色
1	快递盒 1	北京市	深褐色	29	谷粒多燕麦牛奶包装盒	内蒙古自治区呼伦贝尔市	褐色
2	德生堂大药房快递盒	北京市	深褐色	30	山东杂粮煎饼包装纸	北京市	浅褐色
3	伊利高钙奶包装盒	内蒙古自治区呼和浩特市	白色	31	奥利奥包装盒	广东省深圳市	白色
4	晨光美工刀包装纸	上海市	白色	32	阿迪达斯鞋盒	广东省广州市	褐色
5	硬纸片 1	北京市	白色	33	快递盒 3	北京市	褐色
6	态度奶茶包装袋	北京市	白色	34	乳酸菌面包包装盒	浙江省宁波市	褐色
7	枣立方食品包装袋侧面	北京市	黄色	35	医用外科口罩包装盒	北京市	白色
8	枣立方食品包装袋底面	北京市	白色	36	必胜客包装袋	上海市	白色
9	姚记扑克牌包装盒	上海市	白色	37	彩虹真牛堡包装袋	北京市	白色
10	痤疮颗粒包装盒	上海市	白色	38	瑞幸咖啡包装袋	福建省厦门市	褐色
11	得力中性笔包装盒	浙江省宁波市	白色	39	星巴克包装袋	上海市	褐色
12	三金西瓜霜包装盒	广西壮族自治区桂林市	白色	40	喜茶包装袋	广东省深圳市	淡灰色
13	联合利华力士香皂包装盒	安徽省合肥市	白色	41	肯德基包装袋	北京市	浅褐色
14	托马斯保温杯包装盒	北京市	白色	42	吉列剃须泡包装盒	上海市	褐色
15	慕馨玻璃杯包装盒	浙江省金华市	白色	43	绿箭口香糖包装纸	广东省广州市	白色
16	全聚德烤鸭包装盒	北京市	褐色	44	耐克品牌购物袋	上海市	褐色
17	快递盒 2	北京市	深褐色	45	叱咤小警保温杯包装盒	浙江省宁波市	白色
18	哈菲兔时尚童鞋包装盒	浙江省瑞安市	白色	46	华为耳机包装盒	广东省深圳市	白色
19	千黛百合女装包装盒	浙江省杭州市	浅绿色	47	麦当劳包装袋	北京市	白色
20	硬纸片 2	北京市	白色	48	卡牌包装袋	四川省成都市	白色
21	硬纸片 3	北京市	白色	49	西南交大开学典礼礼包袋	四川省成都市	白色
22	元祖食品包装袋	上海市	白色	50	洋河蓝色经典白酒包装袋	江苏省宿迁市	白色
23	硬纸片 4	北京市	白色	51	梦佩斯专业美发包装袋	北京市	白色
24	校庆礼品包装袋	四川省成都市	白色	52	特步品牌购物包装袋	北京市	白色
25	硬纸片 5	北京市	白色	53	安踏品牌购物包装袋	北京市	白色
26	康奈皮鞋鞋盒	浙江省温州市	白色	54	李宁品牌购物包装袋	北京市	白色
27	鲜烤吐司包装袋	北京市	白色	55	麦的多食品包装袋	四川省成都市	白色
28	好丽友 Q 蒂蛋糕包装盒	北京市	浅褐色	56	礼品包装袋	北京市	白色

1.2 主要设备及仪器

主要仪器：傅里叶红外光谱仪，IT 2000（北京鉴知技术有限公司），光谱范围为400~4 000 cm^{-1} ，分辨率为4 cm^{-1} 。

1.3 实验方法

将56个纸质包装袋样品分别剪成0.5 cm×0.5 cm大小的单层长方形，用无水乙醇棉签对剪取的样品擦拭，并进行晾干待测。将处理后的纸质包装袋样品依次置于样品架上，用红外光谱仪进行检测。

2 分析与讨论

2.1 根据纸质包装盒主要填料分类

纸质包装盒的主要原料一般是植物纤维^[7]，在实际生产过程中，为了满足性能的需求，往往会添加填料，如高岭土、碳酸钙、滑石粉、二氧化钛等^[8]。这些常见填料的红外特征峰如表2所示^[9]。

表2 样品中各成分的红外吸收峰（500~4 000 cm^{-1} ）

Tab.2 Infrared absorption peaks of each component in samples (500~4 000 cm^{-1})

成分	吸收峰
碳酸钙	712、875、1 383、1 460、2 854、2 925 cm^{-1}
滑石粉	668、1 020、3 676 cm^{-1}
高岭土	696、778、798、1 066、3 676 cm^{-1}
二氧化钛	722、1 378、1 463、2 854、2 924 cm^{-1}

由于碳酸钙和二氧化钛的共同峰比较多，因此只能选取各自的特征峰作为判断依据^[10]。如712 cm^{-1} 和875 cm^{-1} 对应碳酸钙，722 cm^{-1} 对应二氧化钛；同理，滑石粉和高岭土也存在共同峰，因此出现668 cm^{-1} 和1 020 cm^{-1} 时，可认定为滑石粉，出现696、778、798、1 066 cm^{-1} 时，可认定为高岭土。

根据上述内容以及红外谱图，可先分为3类，如表3所示。

在每一类别中分别挑选具有代表性的样品谱图，如图1、图2、图3所示。

表3 样品所属类别

Tab.3 Categories of samples

类别	样品编号
I	3、7、8、15、19、22、27、29、40、49、50、51、54
II	4、9、11、12、14、24、26、35、45、56
III	1、2、5、6、10、13、16、17、18、20、21、23、25、28、30、31、32、33、34、36、37、38、39、41、42、43、44、46、47、48、52、53、55

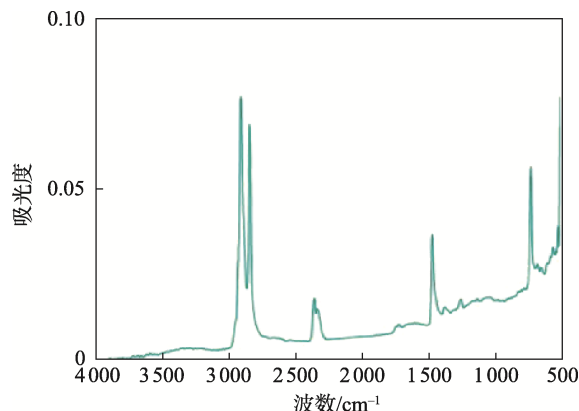


图1 I类7#样品的红外光谱

Fig.1 Infrared spectra of Class I sample 7[#]

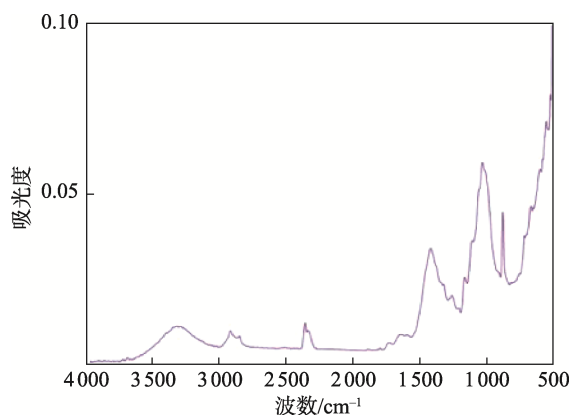


图2 II类4#样品的红外光谱

Fig.2 Infrared spectra of Class II sample 4[#]

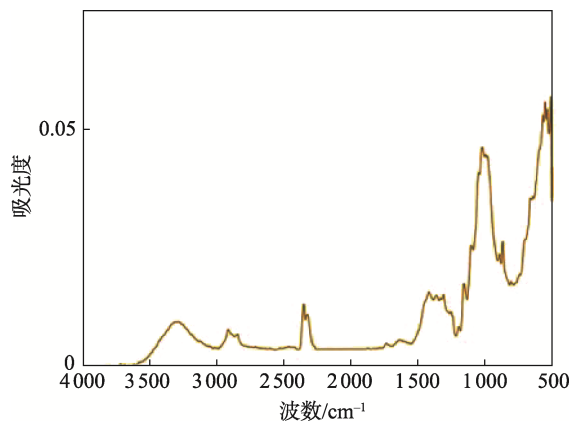


图3 III类6#样品的红外光谱

Fig.3 Infrared spectra of Class III sample 6[#]

对3类样品的填料汇总如表4所示。

表4 样品填料汇总

Tab.4 Summary of sample filler

名称	碳酸钙	滑石粉	高岭土	二氧化钛
I类	无	无	无	有
II类	有	无	有	无
III类	无	无	无	无

2.2 主成分分析法

通过红外谱图对纸盒样品进行分类后, 每一类别中样品数量仍较多。为对样品实现进一步的区分分类, 利用 spss 29.0 软件对数据进行主成分分析 (PCA)。其中前 4 个主成分方差解释率达到 99.142% (见表 5), 说明主成分分析效果较好, 提取这 4 个变量进行后续分析^[11], 为之后的系统聚类法奠定分析基础。

表 5 总方差解释及因子权重分析				
Tab.5 Explanation of total variance and analysis of factor weight				
成分	特征根	方差解释率/%	累积方差解释率/%	权重/%
1	30.915	96.609	96.609	97.445
2	0.514	1.691	98.3	1.705
3	0.184	0.576	98.876	0.581
4	0.085	0.266	99.142	0.268

2.3 系统聚类法

2.1 节提到, 通过分析样品的红外谱图, 结合碳酸钙、滑石粉、高岭土和二氧化钛这 4 种填料的红外特征峰, 来对样本进行初步分类。由于此分类工作为人工进行, 为进一步提高分类精确度和科学性, 在主成分分析 2.2 节进行降维的基础上, 使用系统聚类法对每一类的样本进行进一步分析。使用 spss 29.0 软件进行聚类分析, 以 I 类样本为例, 聚类方法选择组间联接法, 测量区间选择平方欧式距离, 结果如图 4 所示^[12]。

由图 4 可知, 当并类距离为 1 时, 样品可分为 4 类; 当并类距离为 2 时, 样品可分为 2 类; 当并类距

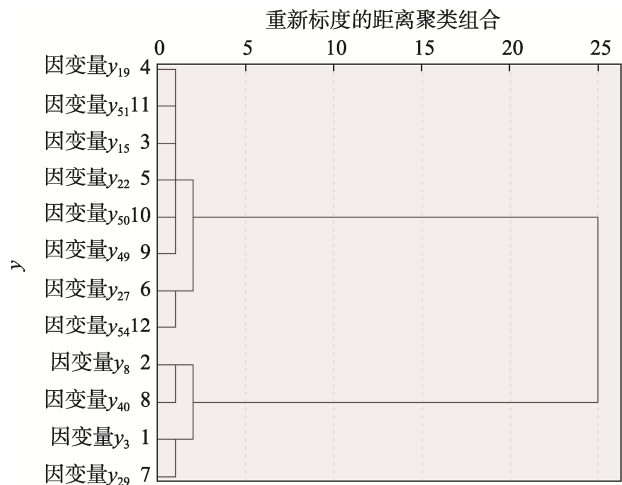


图 4 样本聚类谱系
Fig.4 Pedigree of sample cluster

离为 25 时, 所有样品归为一类。
以类别数量为横坐标, 以聚合系数为纵坐标, 绘制折线图, 见图 5。根据图 5 可知, 当类别数为 2 时, 折线的下降趋势变缓, 故可将类别数设定为 2, 即 I 类样本可以被进一步分为 2 类。
同理, 可将 II 类和 III 类样品分别再用系统聚类法分为 2 类。最终分类结果见表 6, 样本被分为 6 类。

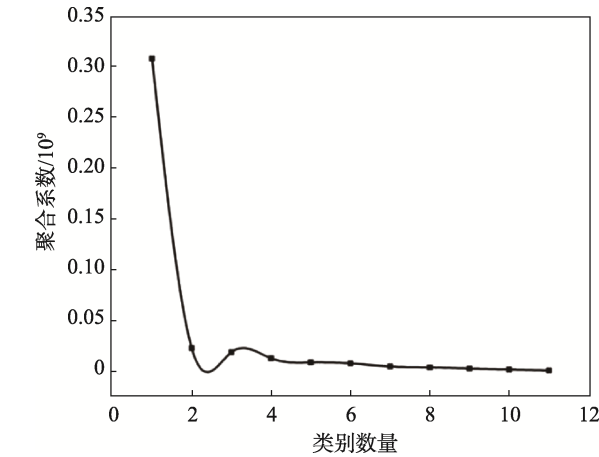


图 5 聚合系数
Fig.5 Clustering coefficients

表 6 样品的最终分类结果	
Tab.6 Final classification result of samples	
类别数	编号
I	15, 19, 22, 27, 49, 50, 51, 54
II	3, 7, 8, 29, 40
III	4, 11, 12, 14, 35, 56
IV	9, 24, 26, 45
V	1, 2, 5, 17, 18, 21, 23, 25, 31, 32, 38, 39, 41, 42, 43, 44, 46, 48, 52, 55
VI	6, 10, 13, 16, 20, 28, 30, 33, 34, 36, 37, 47, 53

2.4 K-means 聚类算法与 Pearson 卡方检验

为验证系统聚类的结果, 利用 K-means 聚类算法进一步处理主成分分析后的数据^[13]。选择聚类数为 6, 对样品进行分类, 结果 (见表 7) 证实通过系统聚类将 56 个样本分为 6 类较为准确。
本文采用了 Pearson 卡方检验, 将系统聚类法得到的最终分类结果与 K-means 聚类分析得到的结果进行比较验证, 从而证实系统聚类结果的准确性。Pearson 卡方检验是用于类别变数有名的卡方检验之一, 用于进行相关性分析。Pearson 卡方检验输出的效应量化分析如表 8 所示, 包括 phi、Crammer's V、列联系数、lambda, 用于分析样本的相关程度^[14-15]。

表 7 样品所属聚类以及距离
Tab.7 Clusters of samples and distance

名称	聚类	距离	名称	聚类	距离
样品 1	5	0.019 08	样品 29	3	0.044 53
样品 2	5	0.023 91	样品 30	6	0.036 42
样品 3	3	0.044 53	样品 31	5	0.025 55
样品 4	2	0.061 09	样品 32	5	0.017 90
样品 5	5	0.054 28	样品 33	6	0.019 55
样品 6	6	0.021 37	样品 34	6	0.010 97
样品 7	3	0.140 02	样品 35	2	0.055 69
样品 8	3	0.042 01	样品 36	6	0.016 44
样品 9	1	0.033 85	样品 37	6	0.029 13
样品 10	6	0.017 39	样品 38	5	0.036 46
样品 11	2	0.036 39	样品 39	5	0.064 60
样品 12	2	0.057 11	样品 40	3	0.042 01
样品 13	6	0.007 81	样品 41	5	0.098 13
样品 14	2	0.026 15	样品 42	5	0.007 90
样品 15	4	0.032 90	样品 43	5	0.000 00
样品 16	6	0.013 40	样品 44	5	0.019 09
样品 17	5	0.031 10	样品 45	1	0.060 36
样品 18	5	0.043 95	样品 46	5	0.031 47
样品 19	4	0.028 47	样品 47	6	0.027 74
样品 20	6	0.028 56	样品 48	5	0.016 95
样品 21	5	0.037 94	样品 49	4	0.091 90
样品 22	4	0.040 54	样品 50	4	0.049 85
样品 23	5	0.010 56	样品 51	4	0.032 28
样品 24	1	0.037 12	样品 52	5	0.021 70
样品 25	5	0.023 27	样品 53	6	0.031 64
样品 26	1	0.067 14	样品 54	6	0.076 51
样品 27	4	0.096 35	样品 55	5	0.026 77
样品 28	5	0.023 71	样品 56	2	0.053 56

表 8 效应量化分析
Tab.8 Quantitative analysis of effects

字段名	K-means 聚类-最终分类结果
Phi	2.166
Crammer's V	0.969
列联系数	0.908
lambda	0.943

phi 系数: phi 相关系数的大小表示两样本之间的关联程度, 当 phi 系数小于 0.3 时, 表示相关较弱; 当 phi 系数大于 0.6 时, 表示相关较强。表 8 中 phi 系数为 2.266, 远大于 0.6, 说明最终的分类结果与 K-means 聚类分析得到的结果相关性极强。

lambda: 用于反应自变量对因变量的预测效果, 一般情况下, 其值为 1 时表示自变量预测因变量效果较好, 为 0 时表明自变量预测因变量较差。表 8 中 lambda 值为 0.943, 非常接近于 1, 说明最终的分类结果对 K-means 聚类分析得到的结果的预测效果非常好。

3 判别预测模型的构建与验证

依据系统聚类(2.3 节)的分类结果, 建立多元判别预测模型, 以实现对新样品的分类预测。构建朴素贝叶斯分类、随机森林、XGBoost 分类 3 种模型, 验证比较 3 种判别模型预测准确率。

3.1 朴素贝叶斯分类

朴素贝叶斯分类是以贝叶斯定理为基础并且假设特征条件之间相互独立的方法。先通过已给定的训练集, 以特征词之间独立作为前提假设, 学习从输入到输出的联合概率分布, 再基于学习到的模型, 输入求出使得后验概率最大的输出^[16]。利用朴素贝叶斯分类模型对 56 个样品进行训练识别, 训练集占比为 70%, 得到分类识别后的混淆矩阵热力图见图 6, 模型评估结果见表 9, 朴素贝叶斯分类模型的综合准确率为 82.9%。

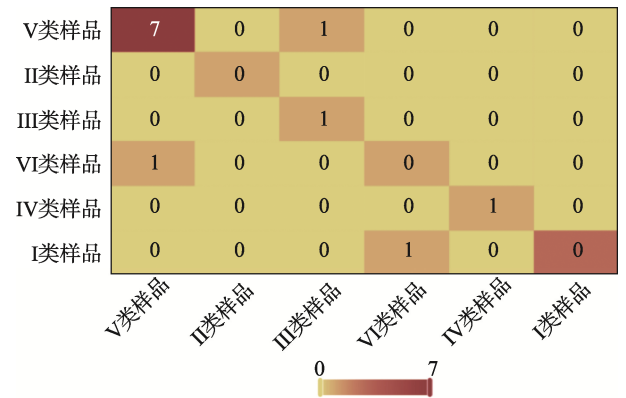


图 6 朴素贝叶斯分类混淆矩阵热力图
Fig.6 Thermodynamic diagram of Naive Bayesian classification confusion matrix

表 9 朴素贝叶斯分类模型评估结果
Tab.9 Evaluation results of Naive Bayesian classification model

类别	准确率	召回率	精确率	F_1
训练集	0.846	0.846	0.9	0.837
测试集	0.824	0.824	0.853	0.829

注: F_1 为精确率和召回率的调和平均, 下同。

3.2 随机森林

随机森林模型是一个包含多棵决策树的分类器, 不同决策树之间没有关联, 是快速分类的监督式聚类的分类方法^[17]。利用随机森林模型对 56 个样品进行训练识别, 训练集占比为仍 70%, 得到分类识别后的混淆矩阵热力图见图 7, 模型评估结果见表 10, 随机森林模型的综合准确率为 94.3%。

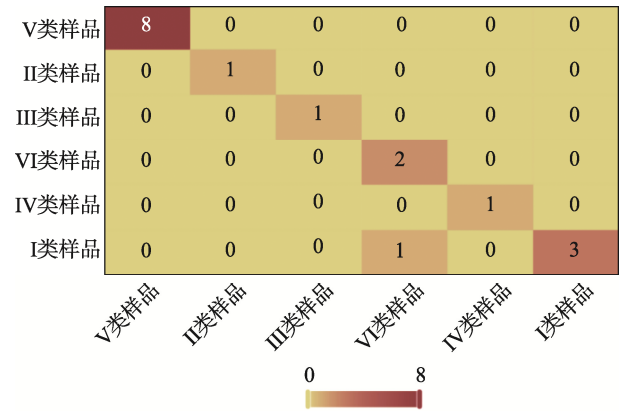


图 7 随机森林混淆矩阵热力图
Fig.7 Thermodynamic diagram of random forest confusion matrix

表 10 随机森林模型评估结果
Tab.10 Evaluation results of random forest model

类别	准确率	召回率	精确率	F_1
训练集	1	1	1	1
测试集	0.941	0.941	0.961	0.943

3.3 XGBoost 分类

XGBoost 分类是一种基于决策树的非参数学习算法, 将数据集分割成多个较小的子集, 然后用这些子集构建出一个决策树, 根据误差函数计算每棵树的分数, 然后对测试数据使用 XGBoost 算法确定最终结果^[18]。利用 XGBoost 分类模型对 56 个样品进行训练识别, 训练集占比仍为 70%, 得到分类识别后的混淆矩阵热力图见图 8, 模型评估结果见表 11, 判别模型的准确率为 84.6%。

将建立的朴素贝叶斯分类模型、随机森林模型、XGBoost 分类模型进行比较, 朴素贝叶斯分类模型准确率为 82.9%, 随机森林模型准确率为 94.3%, XGBoost 分类模型准确率达到 84.6%。3 种判别模型的准确率均较高, 说明该判别方法有较高的可靠性和参考价值。其中随机森林模型准确率高于朴素贝叶斯分类模型和多层感知器判别模型。分析原因可能在于随机森林在处理高维度、非线性特征以及包含异常值的数据时表现较好, 而朴素贝叶斯由于其条件独立假设和对数据分布的假设, 可能在某些情况下表现不如随机森林; XGBoost 是基于梯度提升的决策树, 它可

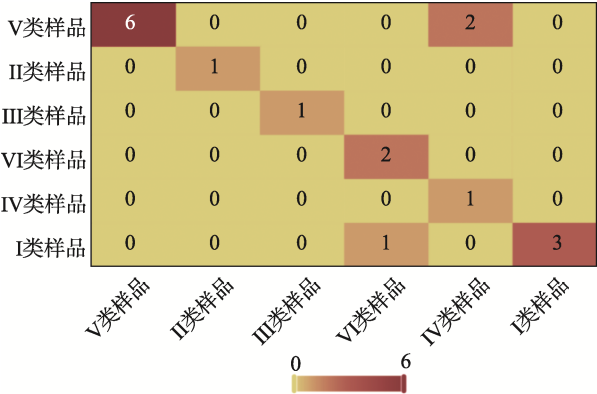


图 8 XGBoost 分类混淆矩阵热力图
Fig.8 Thermodynamic diagram of XGBoost classification confusion matrix

表 11 XGBoost 分类模型评估结果
Tab.11 Evaluation results of XGBoost classification

类别	准确率	召回率	精确率	F_1
训练集	1	1	1	1
测试集	0.824	0.824	0.922	0.846

以处理非线性问题, 但在处理复杂数据集时, 可能没有随机森林强大的非线性建模能力。

4 结语

利用傅里叶变换红外光谱法, 检验不同材质、不同生产地的纸质包装盒, 可以实现无损鉴定, 快速准确。首先根据纸质样品的红外谱图, 将样品大致分为 3 类, 再根据所查阅的特征峰数据, 判断出所包含的填料; 然后在此基础上, 使用主成分分析法对红外光谱测试得到的数据进行降维, 对降维后提取到的主成分进行系统聚类, 将样品最终分为 6 类, 通过 Pearson 卡方检验发现最终分类得出的结果与 K-means 聚类分析结果基本吻合; 最后建立了朴素贝叶斯分类、随机森林、XGBoost 分类 3 种判别预测模型, 对系统聚类下新样品的分类预测结果进行评估, 准确率均较高, 其中随机森林模型能对新样品实现最准确的预测。此方法对犯罪现场所提取的纸盒物证可以进行迅速分类鉴定, 为公安机关处理案件提供帮助。

参考文献:

[1] 郭琦, 姜红, 杨金颐, 等. X 射线荧光光谱结合深度学习算法可视化检验食品包装纸[J]. 激光与光电子学进展, 2022, 59(4): 466-472.
GUO Q, JIANG H, YANG J J, et al. Visual Inspection of Food Packaging Paper by X-Ray Fluorescence Spectroscopy Combined with Deep Learning Algorithm[J]. Laser & Optoelectronics Progress, 2022, 59(4): 466-472.
[2] 林凡琦, 姜红, 张文宇. 高光谱结合多元统计分析鉴

- 别一次性纸杯样品[J]. 化学研究与应用, 2021, 33(10): 1949-1955.
- LIN F Q, JIANG H, ZHANG W Y. Hyperspectral Analysis Combined with Multivariate Statistical Analysis to Identify Disposable Paper Cup Samples[J]. Chemical Research and Application, 2021, 33(10): 1949-1955.
- [3] 姜红, 鞠晨阳, 黄国玺, 等. 拉曼光谱法检验一次性纸杯的研究[J]. 黑龙江造纸, 2018, 46(4): 4-9.
- JIANG H, JU C Y, HUANG G X, et al. Analysis of Disposable Paper Cups by Raman Spectroscopy[J]. Heilongjiang Pulp & Paper, 2018, 46(4): 4-9.
- [4] 付钧泽, 姜红, 陈煜太, 等. 傅里叶变换红外光谱法检验香烟水松纸的研究[J]. 中华纸业, 2017, 38(20): 43-48.
- FU J Z, JIANG H, CHEN Y T, et al. A Study on Test of Tipping Paper with Fourier Transform Infrared Spectroscopy[J]. China Pulp & Paper Industry, 2017, 38(20): 43-48.
- [5] 杨梦恩, 姜红, 陈琨, 等. 差分拉曼光谱和统计学方法在食品包装纸分类中的应用[J]. 理化检验-化学分册, 2022, 58(11): 1303-1308.
- YANG M E, JIANG H, CHEN H, et al. Application of Differential Raman Spectroscopy and Statistical Methods in Food Packaging Paper Classification[J]. Physical Testing and Chemical Analysis (Part B: Chemical Analysis), 2022, 58(11): 1303-1308.
- [6] 徐冰冰. 防水防油牛皮纸的制备及性能表征[D]. 北京: 北京林业大学, 2021: 43-47.
- XU B B. Preparation and Characterization of Waterproof and Oilproof Kraft Paper[D]. Beijing: Beijing Forestry University, 2021: 43-47.
- [7] 马建锋, 杨淑敏, 田根林, 等. 拉曼光谱在天然纤维素结构研究中的应用进展[J]. 光谱学与光谱分析, 2016, 36(6): 1734-1739.
- MA J F, YANG S M, TIAN G L, et al. Study on the Application of Raman Spectroscopy to the Research on Natural Cellulose Structure[J]. Spectroscopy and Spectral Analysis, 2016, 36(6): 1734-1739.
- [8] 姜红, 陆润洲, 段斌, 等. 基于差分拉曼光谱对烟盒物证检验[J]. 包装工程, 2021, 42(21): 79-85.
- JIANG H, LU R Z, DUAN B, et al. Cigarette Case Evidence Based on Differential Raman Spectroscopy[J]. Packaging Engineering, 2021, 42(21): 79-85.
- [9] NA T, ZHU C J, ZHANG C H, et al. Study on Raman Spectra of Aliphatic Polyamide Fibers[J]. OPTIK International Journal for Light and Electron Optics, 2016, 127(1): 21-24.
- [10] 郭鹏, 姜红. 拉曼光谱检验烟用内衬纸的研究[J]. 中华纸业, 2016, 37(12): 53-58.
- GUO P, JIANG H. A Study on Analysis of Inner Liner for Cigarette by Raman[J]. China Pulp & Paper Industry, 2016, 37(12): 53-58.
- [11] 朱晓晗, 胡越宁, 何歆沂, 等. 基于主成分分析法对一次性纸杯的拉曼光谱检验研究[J]. 中国造纸, 2020, 39(9): 38-42.
- ZHU X H, HU Y N, HE X Y, et al. Detection of Disposable Paper Cup by Raman Spectroscopy and PCA[J]. China Pulp & Paper, 2020, 39(9): 38-42.
- [12] 张进, 姜红, 刘峰, 等. 鞋底材料的差分拉曼光谱可视化快速鉴别[J]. 激光与光电子学进展, 2021, 58(8): 0830004.
- ZHANG J, JIANG H, LIU F, et al. Differential Raman Spectroscopy Visualization and Rapid Identification of Shoe Sole Materials[J]. Laser & Optoelectronics Progress, 2021, 58(8): 0830004.
- [13] 吴凤慧, 成颖, 郑彦宁, 等. K-means 算法研究综述[J]. 现代图书情报技术, 2011(5): 28-35.
- WU S H, CHENG Y, ZHENG Y N, et al. Survey on K-Means Algorithm[J]. New Technology of Library and Information Service, 2011(5): 28-35.
- [14] GIANNA B, JEREMY S, KENNETH F. Wallis. Decompositions of Pearson's Chi-Squared Test[J]. Journal of Econometrics, 2004, 123(1): 189-193.
- [15] 张建勇, 高冉, 胡骏, 等. 灰色关联度和 Pearson 相关系数的应用比较[J]. 赤峰学院学报(自然科学版), 2014, 30(21): 1-2.
- ZHANG J Y, GAO R, HU J, et al. Application Comparison of Grey Correlation Degree and Pearson Correlation Coefficient[J]. Journal of Chifeng University (Natural Science Edition), 2014, 30(21): 1-2.
- [16] 阿曼. 朴素贝叶斯分类算法的研究与应用[D]. 辽宁: 大连理工大学, 2014: 2-4.
- A Man. Research and Application on Naive Bayes Classification[D]. Liaoning: Dalian University of Technology, 2014: 2-4.
- [17] 张宇, 胡晓光, 姜红, 等. 红外光谱法结合化学计量学对快递包装纸盒的检验研究[J/OL]. 包装工程: 1-10[2023-08-07]. <http://kns.cnki.net/kcms/detail/50.1094.TB.20230510.1638.002.html>.
- ZHANG Y, HU X G, JIANG H, et al. Infrared Spectroscopy Combined with Chemometrics on the Inspection of Express Cartons [J/OL]. Packaging Engineering: 1-10[2023-08-07]. <http://kns.cnki.net/kcms/detail/50.1094.TB.20230510.1638.002.html>.
- [18] 黄卿, 谢合亮. 机器学习方法在股指期货预测中的应用研究——基于 BP 神经网络、SVM 和 XGBoost 的比较分析[J]. 数学的实践与认识, 2018, 48(8): 297-307.
- HUANG Q, XIE H L. Research on the Application of Machine Learning in Stock Index Futures Forecast—Comparison and Analysis Based on BP Neural Network, SVM and XGBoost[J]. Journal of Mathematics in Practice and Theory, 2018, 48(8): 297-307.