基于判别分析与 ANN 的药品铝塑包装泡罩 XRF 分析

姜红¹,马枭¹,李飞²,李春宇¹,吕航¹,范烨³,满吉⁴

(1.中国人民公安大学 侦查学院,北京 100038; 2.昆明市公安局网络安全保卫支队,昆明 650000;3.广州市公安局白云分局,广州 510420; 4.北京华仪宏盛技术有限公司,北京 100123)

摘要:目的 针对案件现场常见的药品铝塑包装泡罩,为达到对其分类识别的目的,提出系列检验分析、 数据处理方法。方法 采用 X 射线荧光光谱法对 45 个药品铝塑包装泡罩样本所含元素进行检验并讨论 分析。对检验结果进行无监督的系统聚类,利用离差平方和法计算欧氏距离进而将未知样本分为 5 类。 结果 将分类结果作为变量进行判别分析,选取累积方差百分比为 97.8%的 2 个判别函数,其类内平方 和与总平方和之比为 0.015 和 0.394,具有较强的解释能力。绘制的样本判别分类图将 5 类样本类之间 相互区分开来,样本总体判别正确率为 95.6%。提取样本在判别函数上的判别得分构建了人工神经网络, 最终分类正确率为 97.8%。结论 利用 X 射线荧光光谱法对药品铝塑包装泡罩进行检验,将元素种类及 含量作为变量进行了分类,并构建了 45 个药品铝塑包装泡罩样本的人工神经网络分类模型,可借助该 模型进一步实现对于案件现场未知类别的药品铝塑包装泡罩样本的分类识别。 关键词: X 射线荧光光谱法;药品铝塑包装泡罩;元素;系统聚类;判别分析;人工神经网络 中图分类号: TB487; O657.34 文献标识码:A 文章编号: 1001-3563(2021)09-0189-05 DOI; 10.19554/j.cnki.1001-3563.2021.09.026

XRF Analysis of Pharmaceutical Aluminum-Plastic Packaging Blister Based on Discriminant Analysis and ANN

JIANG Hong¹, MA Xiao¹, LI Fei², LI Chun-yu¹, LYU Hang¹, FAN Ye³, MAN Ji⁴

(1.School of Investigation, People's Public Security University of China, Beijing 100038, China;
 2.Kunming Public Security Bureau Network Security Detachment, Kunming 650000, China;
 3.Baiyun District Branch of Guangzhou Municipal Public Security Bureau, Guangdong 510420, China;
 4.Beijing Huayi Horizon Technology Co., Ltd., Beijing 100123, China)

ABSTRACT: A series of inspection, analysis and data processing methods were proposed for common pharmaceutical aluminum-plastic packaging blister at the scene of cases in order to achieve the purpose of classification and identification. X-ray fluorescence spectrometry was used to test and analyze the elements contained in 45 pharmaceutical aluminum-plastic packaging blister samples. Unsupervised systematic clustering was performed on the test results, and the Euclidean distance was calculated by using the sum of squared deviation method to classify the unknown samples into 5 categories. The classification results were observed as discriminant analysis variables. Two discriminant functions with a cumulative variance percentage of 97.8% were selected and the Wilks' lambda is 0.015 and 0.394, which has the strongest explanatory ability. Finally, the five types of samples were distinguished from each other, and the overall discrimination accuracy rate was 95.6%. In order to achieve the purpose of pattern recognition of samples of unknown categories, extract

收稿日期: 2020-09-08

基金项目:国家重点研发计划(2017YFC0822004, 2019YFF0303405);中央高校基本科研业务费(2019JKF427) 作者简介:姜红(1963—),女,硕士,中国人民公安大学教授,主要研究方向为微量物证、理化检验。

通信作者:李春宇(1980—),男,硕士,中国人民公安大学副教授,主要研究方向为光谱检验技术。

the discriminant score of the samples on the discriminant function to construct an artificial neural network. The final classification accuracy rate was 97.8%. X-ray fluorescence spectroscopy was used to test the pharmaceutical aluminum-plastic packaging blister, and the types and contents of elements were classified as variables and an artificial neural network classification model of 45 pharmaceutical aluminum-plastic packaging blister was constructed. This model can be used to further achieve the classification and identification of pharmaceutical aluminum-plastic packaging blister of unknown categories at the scene of cases.

KEY WORDS: X-ray fluorescence spectrometry; pharmaceutical aluminum-plastic packaging blister; element; hierarchical clustering; discriminant analysis; artificial neural network

药品是人们生活中的必需品,在各类案件现场 中,经常能提取到一些药品的包装物。目前,市面常 见的药品包装材料包括塑料、橡胶、玻璃和金属这4 类^[1-3]。塑料药品包装具有良好的力学性能,具有一 定强度和韧性,能有效保护内部药品;具有良好的安 全性能,毒性小且能防虫防腐;具有良好的加工性能, 易于加工和机械化批量生产^[4]。塑料在药品包装行业 中得到了广泛应用,其中药品铝塑包装泡罩主要成分 为聚氯乙烯,具有质地轻、可塑性强、可大规模批量 生产等优点,被广泛应用于包装胶囊剂、片剂和丸剂 等固体药剂^[5]。通过对案件中提取到的药品铝塑包装 泡罩物证进行分析检验,可以为案件的侦破提供线索 并证明犯罪事实,同时也为检测药品包装安全性、打 击制贩假药等提供帮助。

X 射线荧光光谱法(X-Ray Fluorescence, XRF) 具有无损检材、分析速度快、测量元素等特点,被广 泛应用于医药^[6]、地质、考古^[7]、环境^[8—9]、海洋^[10—11] 等各个领域。在前期实验工作中,已通过元素种类以 及简单的无监督分析方法将药品铝塑包装泡罩进行 分类,但分类的效果以及分类准确率有待进一步提 高。文中旨在通过 X 射线荧光光谱法对药品铝塑包 装泡罩进行元素检验,通过多元统计分析的方法实 现对样品的可视化分类,并构造人工神经网络分类 模型实现对未知样本的自动分类,为物证分析提供 新的思路。

1 实验

1.1 原料

主要原料:收集不同种类、厂家的药品铝塑包装 泡罩样品共45个。

1.2 设备及参数

主要设备: 便携式 X 射线荧光光谱仪 X-MET8000,英国 Oxford Instruments,电压为 40 kV, 电流为 60 mA。

1.3 元素数据采集

利用便携式 X 射线荧光光谱仪对 45 个药品铝塑

包装泡罩样本进行测试,经过预实验后设定测试时间 为 60 s,对 3 次测试数据取平均值得到最终数据。同 时随机选取一样本,在测定时间为 60 s 的情况下考察 仪器测量结果的重现性。

2 结果与讨论

2.1 元素结果分析

根据实验结果可知,药品铝塑包装泡罩样品中所 含元素主要为 Cl, Ca, Fe, Zn, Ag, Sn, Ta, Hg, V, Cu, Mn, Ti, 其中元素 Cl 与 Fe 的含量最多。其 原因是药品铝塑包装泡罩的主要成分为聚氯乙烯 (PVC),所含 Cl 元素含量较多。此外,在塑料加工 生产过程中为了达到增强产品韧性、保持合适软硬度 的目的,常常会加入含有 Fe, Ti 等元素的增塑剂。 为了有效阻止和减少塑料的降解,往往会加入含有 Zn, Ca, Sn 等元素的稳定剂。

2.2 系统聚类

近年来,聚类分析法被广泛地应用于理化检验领 域^[12],取得了良好效果。系统聚类可对未知样本进行 聚类分析并分类。离差平方和法是将类与类之间合并 后离差平方和的增加作为类间的平均距离,是一种 应用较为广泛、适用性较好的方法。文中采用离差 平方和法,以及聚类分析中广泛应用的平方欧式距 离作为衡量类别之间亲疏程度的距离。绘制的聚类 结果见图 1。

由图 1 可知,系统聚类将样本之间的分类概率情况映射为 0,5,10,15,20,25。样本之间分为同一组的概率从映射为 0 时开始依次递减,因此,当初始映射即 45 个药品铝塑包装泡罩样本分类最细致时,将样本分为 5 类;当映射值为 5 时,将样本分为 3 类;当映射值为 10 时,将样本分为 2 类;当映射值为 25 时,所有样本归为同一类。对在初始映射时,分出的 5 类样本中的第 1 类数据即对 2[#],3[#],4[#],6[#],14[#]样本数据进行分析,见表 1。

由表 1 可知, 第 1 类样本中的 5 个样本所含 Cl 元素含量均小于 10 g/kg, 而其余样本 Cl 元素含量 均大于 10 g/kg; 5 个样本 Fe 元素含量均小于 1 g/kg,





函数

编号

1

2

3

特征

值

24.656

0.563

0.401

表 1 第 1 类样本元素含量 Tab.1 Element content of the first sample

	表 2	判别分析结果	
Tab.2	Result	of discriminant an	alysis

累积百分

比/%

95.6

97.8

99.4

Wilks' Lambda

0.015

0.394

0.616

方差百分

比/%

95.6

2.2

1.6

样木棺具	元素含量(mg·kg ⁻¹)			
中午细石	Cl	Fe	Mn	Ti
2	220	541	28	0
3	595	877	0	0
4	1 823	746	0	0
6	1 088	0	0	0
14	1 830	785	0	4

而其余样本 Fe 元素含量均大于 1 g/kg, 且 5 个样本 均不含或含有少量 Mn 和 Ti 元素。通过分析第 1 类 样本元素质量分数可知,系统聚类依据较为明确, 类内样本特征较为一致,分类结果准确可靠。

为进一步检验聚类结果是否准确,将无监督模式 下的聚类分析结果作为有监督模式下判别分析的依 据,以此进行系统聚类准确性的验证。

2.3 判别分析

判别分析作为一种常用的有监督学习方法,可实现对样本的分类检验^[13],因此,文中利用系统聚类分类结果进行判别模型的构建,达到对样本分类效果进行检验的目的。判别分析得到 45 个药品铝塑包装泡罩样本分析结果,见表 2。

由表 2 可知,表 2 中总共提取了 4 个判别函数。 特征值以及方差百分比表示各判别函数对于各类别 的样本的区分能力的大小^[14],函数 1 的特征值为 24.656,方差百分比为 95.6%,表示函数 1 对于样品 类别的区分能力十分强。函数 2 特征值为 0.563,方 差百分比为 2.2%,其区分能力较弱,但为了达到可 视化的目的,清楚地区分各类别之间的差异,所以应 至少选取 2 个判别函数构成二维空间分布。函数 3 和
 4
 0.158
 0.6
 100
 0.864

 函数 4 的特征值分别为 0.401 和 0.158, 方差百分比

 分别为 1.6%和 0.6%,表明其区分能力均弱于函数 2。

 由于前 2 个判别函数累积方差百分比已达到 97.8%,

 已经能很好地区分各类别样本,因此,是否保留判别

 函数 3 和判别函数 4 有待进一步考察。

Wilks' lambda 是类别内平方和与总平方和之比, Wilks' lambda 值较大,表示各个类别的均值基本相 等;Wilks' lambda 值较小,表示类别之间差异较为明 显^[15]。判别函数 1 和判别函数 2 的 Wilks' lambda 值 为 0.015 和 0.394,明显小于判别函数 3 和判别函数 4 的 0.616 和 0.864,表明函数 1 和 2 贡献大于函数 3 和 4。综合考虑方差百分比和 Wilks' lambda 值,选择 判别函数 1 和判别函数 2 作为判别函数轴绘制样本判 别分类,见图 2。

由图 2 可知,由于判别函数 1 所占方差百分比占 据主要地位,5 类样本在函数 1 上明显相互区分开来。 在判别函数 2 上的区分效果较弱,5 类样本类间分离 程度较小。组质心是各类样本中所有样本在样本分类 图中的平均质心,各组之间的组质心分布可作为区分 各类别样本的直观依据。图 2 中 5 类样本组质心可直 观区分开来,证明分类效果良好。为进一步检验预测 结果的准确性,对 45 个药品铝塑包装泡罩样本进行 预测分类检验,分类结果见表 3。



图 2 样本判别分类 Fig.2 Classification diagram of samples

表 3 判别分析预测分类结果 Tab.3 Summary of predictive classification

百払八米	预测分类				
尿如刀矢	1	2	3	4	5
1	5	0	0	0	0
2	1	12	0	0	0
3	0	0	6	0	0
4	0	0	0	11	0
5	0	0	0	1	9

由表3可知,依据所构建的判别模型对原始样本 进行分类时,将一个原始分类为第3类的样本划分为 了第2类,将一个原始分类为第5类的样本划分为了 第1类。总体预测正确率为95.6%,分类效果较为良 好。值得注意的是,判别分析仅仅对已知样本进行了 检验,对于未知样本的分类能力有待进一步检验,因 此还需进行进一步的研究。

2.4 人工神经网络

人工神经网络(artificial neural networks, ANN) 由输入层、隐藏层和输出层组成,可通过不断调整输 出的结果连接之间的权重以达到准确分类的目的。输 入层和输出层的神经元数目分别由输入数据数目和 输出种类数量确定,而隐藏层神经元数目Z可根据式 (1)进行设置^[16]。

$$Z = \sqrt{X + Y} + f \tag{1}$$

式中: *X* 为输入层神经元数目; *Y* 为输出层神经 元数目; *f* 为 0~10 的常数, 经验证后此处 *f* 取值为 2, 即隐藏层神经元数目为 5。

在人工神经网络中,样本数据被按照 75%,15%, 15%的比例随机分为训练集、验证集和测试集这 3 个 部分。其中,训练集通过不断地将输出结果与真实结 果相拟合,达到调整权重、提高模型分类能力的目的。 验证集主要是对网络参数的设置以及泛化能力进行 检验,验证网络设置是否合理。测试集内的数据与训 练集内的数据互不重合,以此测试经过训练集训练和 验证集调整得到的人工神经网络对未知陌生样本的 分类与识别效果^[17]。

前期判别分析选取了特征值最大的 2 个判别函数进行分类,取得了 95.6%的正确率,但人工神经网络将所有样本分为了互不重复的训练集、验证集和测试集,对于样本的数据特征信要求更高,因此,提取样本在 4 个判别函数上的判别得分输入人工神经网络,输出分类结果见表 4。

表 4 人工神经网络输出分类结果 Tab.4 Artificial neural network outputs classification results

原始分类		-	预测分类	1	
	1	2	3	4	5
1	5	0	0	0	0
2	0	13	0	0	0
3	0	0	6	0	0
4	0	0	0	11	1
5	0	0	0	0	9

由表4可知,人工神经网络分类正确率为97.8%, 在原有判别分析正确率基础上有所提高。由于划分了 训练集、验证集和测试集,所得网络结果经过了验证 和模拟,因此该人工神经网络模型可直接应用于未知 类别样本的模式识别。

3 结语

文中将无监督的系统聚类与有监督的判别分析 和人工神经网络相结合,在分类过程中进行了多次验 证,避免了拟合不足与过拟合的情况。在对已有样本 快速分类的同时对未知样本进行模拟分类。首先,系 统聚类实现了对未知药品铝塑包装泡罩的分类,通过 检验已分为同一类的样本之间的共同特征是否一致, 达到初步验证分类结果的目的。进而利用聚类结果作 为判别分析依据,通过判别摘要以及可视化样本分类 图验证分类结果是否准确。验证的同时进行判别分 析,总体预测正确率为 95.6%。进而将判别得分作为 数据输入所构建的人工神经网络,最终得到了正确率 为 97.8%。通过 3 种数据分析方法的结合,提高了物 证分类的准确性。

参考文献:

[1] 周健丘,梅丹.药品包装材料对药品质量和安全性

的影响[J]. 药物不良反应杂志, 2011, 13(1): 27—31. ZHOU Jian-qiu, MEI Dan. Influence of Pharmaceutical Packaging Materials on Quality and Safety of Medicines[J]. Adverse Drug Reactions Journal, 2011, 13(1): 27—31.

[2] 姜红,任继伟,鞠晨阳,等.X 射线荧光光谱结合聚
 类分析检验药品铝塑包装片[J].化工新型材,2019,47(11):194—198.
 JIANG Hong, REN Ji-wei, JU Chen-yang, et al. XRF

Combined with Cluster Analysis for Inspection of Pharmaceutical Aluminum-Plastic Package[J]. New Chemical Materials, 2019, 47(11): 194–198.

[3] 梁吉雷,吴萌萌,宋玉鹤,等. 浅议我国药品包装材料现状及发展趋势[J].山东化工,2017,46(6):65—67.
 LIANG Ji-lei, WU Meng-meng, SONG Yu-he, et al.

Progress of Pharmaceutical Packaging Materials in China[J]. Shandong Chemical Industry, 2017, 46(6): 65–67.

- [4] 姜红, 鞠晨阳, 盘永平, 等. 拉曼光谱法检验泡沫塑料的研究[J]. 上海塑料, 2019(2): 48—53.
 JIANG Hong, JU Chen-yang, PAN Yong-ping, et al. Research on Foam Plastics by Raman Spectroscopy[J]. Shanghai Plastics, 2019(2): 48—53.
- [5] 赖仲平,张焱. 透明 PVC 药品包装材料透射率的研究[J]. 塑料助剂, 2017(6): 30—34.
 LAI Zhong-ping, ZHANG Yan. Research on Transmittance of Clarity PVC Drug Packaging[J]. Plastics Additives, 2017(6): 30—34.
- [6] ORTIZ R S, MARIOTTI K C, SCHWAB N V, et al. Fingerprinting of Sildenafil Citrate and Tadalafil Tablets in Pharmaceutical Formulations via X-Ray Fluorescence (XRF) Spectrometry[J]. Journal of Pharmaceutical and Biomedical Analysis, 2012, 58: 7–11.
- [7] BURGIO L, CLARK R J H, HARK R R. Raman Microscopy and X-Ray Fluorescence Analysis of Pigments on Medieval and Renaissance Italian Manuscript Cuttings[J]. Proceedings of the National Academy of Sciences, 2010, 107(13): 5726–5731.
- [8] FELLIN M., NEGRI M, ZANUTTINI R. Multi-Elemental Analysis of Wood Waste Using Energy Dispersive X-Ray Fluorescence (ED-XRF) Analyzer[J]. European Journal of Wood and Wood Products, 2014, 72: 199–211.
- [9] PEARSON D, CHAKRABORTY S, DUDA B, et al. Water Analysis via Portable X-Ray Fluorescence Spectrometry[J]. Journal of Hydrology, 2017, 544: 172–179.
- [10] TURNER A, SOLMAN K R. Analysis of The Elemen-

tal Composition of Marine Litter by Field-Portable-XRF[J]. Talanta, 2016, 159: 262-271.

- [11] TURNER A. Mobilisation Kinetics of Hazardous Elements in Marine Plastics Subject to An Avian Physiologically-Based Extraction Test[J]. Environmental Pollution, 2018, 236: 1020–1026.
- [12] SAJDAK M. Application of Chemometrics to Identifying Solid Fuels and Their Origin[J]. Central European Journal of Chemistry, 2013, 11(2): 151–159.
- [13] TITARENKO V O, KHALAFYAN A A, TEMERDASHEV Z A, et al. Identification of The Varietal and Regional Origin of Red Wines by Classification Analysis[J]. Journal of Analytical Chemistry, 2018, 73(2): 195–206.
- [14] 马枭,姜红,杨佳琦. X 射线荧光光谱结合多元统计 分析塑料打包带(绳)[J]. 激光与光电子学进展, 2019, 56(22): 243—247.
 MA Xiao, JIANG Hong, YANG Jia-qi. Examination of Plastic Pack Belts (Ropes) via X-Ray Fluorescence Spectrometry Combined with Multivariate Statistical Analysis[J]. Laser & Optoelectronics Progress, 2019, 56(22): 243—247.
- [15] 李乃和,汤世强.特大型连续服务供应商的顾客细分与新顾客归类判别研究[J].中国管理科学,2015,23(S1):602—609.
 LI Nai-he, TANG Shi-qiang. Research on Customers' Segmentation and New Comers' Classification by Super-Large Scale Continuous Service Provider[J]. Chinese Journal of Management Science, 2015, 23(S1):602—609.
 [16] 叶树彬,徐亮,李亚凯,等.基于人工神经网络的傅
- [10] 所柄枪, 标先, 争业战, 寻. 坐了火工种空柄站的将 里叶变换中红外光谱法对食用油油烟种类识别研究
 [J]. 光谱学与光谱分析, 2017, 37(3): 749—754.
 YE Shu-bin, XU Liang, LI Ya-kai, et al. Study on Recognition of Cooking Oil Fume by Fourier Transform Infrared Spectroscopy Based on Artificial Neural Network[J]. Spectroscopy and Spectral Analysis, 2017, 37(3): 749—754.
- [17] 王晓宾, 马泉, 王新承. 基于人工神经网络的水彩笔 油墨红外光谱模式识别[J/OL]. 激光与光电子学进展: 1—11[2020-06-07]. http://kns.cnki.net/kcms/detail/ 31.1690.TN.20200601.1754.167.html.
 WANG Xiao-bin, MA Xiao, WANG Xin-cheng. Infrared Spectral Pattern Recognition of Watercolor Pen Ink Based on Artificial Neural Network[J/OL]. Laser & Optoelectronics Progress: 1—11[2020-06-07]. http:// kns.cnki.net/kcms/detail/31.1690.TN.20200601.1754.1 67.html.