

## 基于高光谱成像的纸质文物“狐斑”检测方法研究

戴若辰<sup>1</sup>, 唐欢<sup>2\*</sup>, 汤斌<sup>1\*</sup>, 赵明富<sup>1</sup>, 代理勇<sup>1</sup>, 赵雅<sup>3</sup>, 龙邹荣<sup>1</sup>, 钟年丙<sup>1</sup>

1. 重庆理工大学, 重庆市光纤传感与光电检测重点实验室, 重庆 400054

2. 重庆中国三峡博物馆, 馆藏文物有害生物控制研究国家文物局重点科研基地, 重庆 400015

3. 重庆第二师范学院, 重庆 400065

**摘要** 受保存条件影响, 很多纸质文物表面会形成狐斑(foxing), 如果不能进行有效监测和科学判断, 会进而影响纸质文物安全。纸质文物狐斑病害检测存在滞后性、主观性等问题, 对于书画藏品被墨色、颜料及印章等覆盖的区域更是难以通过肉眼进行识别, 因此, 基于文物的预防性保护理念, 亟待开发对于狐斑高效、精确识别的无损检测技术。可见光-近红外高光谱图像结合了光谱和图像, 包含丰富的空间信息与光谱信息, 可以实现无损批量地平面采集样本光谱信息。该研究提出一种基于高光谱成像技术检测纸质文物狐斑的快速识别方法, 获取模拟纸质文物在 360~970 nm 的高光谱图像, 因 360~450 nm 受噪声影响过大, 所以选择剔除这部分光谱数据; 选取感兴趣区域并获取相应的平均光谱反射率, 比较健康区域与被狐斑感染区域, 发现两者的光谱曲线存在差异; 在 450~600 nm 附近, 受狐斑影响区域比健康区域的光谱反射率偏高, 并在 600 nm 附近出现波峰形态; 而在 600~900 nm 范围内, 被感染区域与健康区域的光谱都趋于平稳, 两者之间差异逐渐减小。选取从特征波长对应的图像中提取的特征信息建立图像识别模型, 运用波段运算观察狐斑图像特征, 狐斑的大小和分布情况都能清晰地显示, 但与印章和墨迹重叠部分, 狐斑被印章和墨迹遮盖, 难以识别; 利用最小噪声分离, 虽然不同部分有重叠, 但能发现仅凭肉眼难以识别的隐藏的狐斑; 180 条高光谱数据(450~970 nm)建立狐斑判别模型, 随机地分为 120 条数据为训练集, 60 条数据为测试集, 应用 K-近邻法与 BP 神经网络建立纸质文物狐斑光谱判别模型, 总体上两种方法对狐斑判别率分别达到 73.3% 和 85%; BP 神经网络相较于 K-近邻模型, 总体判别率更高, 识别效果也更好。结果表明, 利用高光谱成像可高效准确识别纸质文物狐斑, 为后续研究狐斑分布发展提供可靠的技术手段, 也为博物馆馆藏文物的保存提供指导意见。

**关键词** 纸质文物; 狐斑; 高光谱图像; 光谱; 机器学习分类

**中图分类号:** TN29

**文献标识码:** A

**DOI:** 10.3964/j.issn.1000-0593(2022)05-1567-05

### 引言

造纸术的出现使中华民族能够利用纸张保存政治、文化、军事等信息, 宝贵的纸质文物蕴含了中国近 2000 年历史的智慧结晶。但在纸质文物保存过程中, 若保存方法不当, 极易遭受“狐斑”影响。狐斑通常表现为黄色、褐色或铁锈色斑点, 因其狐狸般狡黠的生长特性与毛皮的颜色, 故称之为狐斑(foxing)<sup>[1]</sup>。狐斑不仅掩盖字迹, 影响阅读, 而且影响范围会不断扩大, 甚至渗透到深层的书页中。由于目前针对文物表面狐斑清洗的技术方法十分有限, 因此, 狐斑一旦形

成, 很难去除, 将给文物带来不可逆转的损坏。狐斑的成因还未形成统一的结论, 主要有生物论与无机论两种理论。生物论认为主要是霉菌等真菌造成纸张上产生狐斑, 霉菌以纸张中纤维素为营养源, 通过分泌纤维素酶, 降解与吸收纸张中的纤维素, 导致纸张机械强度下降。同时在霉菌代谢过程中, 产生的草酸、柠檬酸等有机酸使纸质材料内部酸性增加, 分泌的色素在纸质文物表面形成狐斑。无机论认为狐斑是金属离子如铜、铁等在纸张表面沉积的结果。狐斑的形成可能涉及复杂的生物、化学反应, 既有可能与一种或几种真菌生长有关, 也存在无机物沉积共同作用的可能性<sup>[2-4]</sup>。除了成因尚未明确之外, 狐斑在有机质文物, 特别是纸质文物

**收稿日期:** 2021-04-13, **修订日期:** 2021-07-27

**基金项目:** 国家重点研发计划专项(2020YFC1522500), 国家自然科学基金青年科学基金项目(61805029)资助

**作者简介:** 戴若辰, 1997 年生, 重庆理工大学电气与电子工程学院硕士研究生 e-mail: 1434437978@qq.com

\* 通讯作者 e-mail: tanghuan3gm@163.com; tangbin@cqut.edu.cn

表面的侵袭、扩散、发展规律更是少有研究。对于狐斑的传统检测,完全依靠人工目视观察,不仅在狐斑病害发生的识别上具有滞后性,同时也存在检测人员判别的主观性,对于书画藏品被墨色、颜料及印章等覆盖的区域更是难以通过肉眼进行识别,因此,基于文物的预防性保护理念,亟待开发对于狐斑高效、精确识别的无损检测技术,以此为后续的防治提供基础。

高光谱成像技术结合了光谱与图像分析的优点,具有无损、快速、高效等特点,在不损坏样品的情况下,快速完成光谱数据采集,单次采集即可获得样本光谱、形状、颜色和成分等不同方面特征,在检测纸质文物特征研究中具有巨大的优势和潜力,目前已在纸质文物检测方面有初步应用。如:首都博物馆的武望婷等<sup>[5]</sup>利用高光谱相机采集清代张士保作品《论道图》的影像,经过最小噪声分离变换、主成分变换、掩膜等处理对画作的墨线进行了提取,实现了图像增强效果,并且发现了两处人眼无法识别的隐含信息;上海博物馆的周新光等<sup>[6]</sup>为了提高印章的可辨识度,尝试解决印章与画意重合的问题;武锋强等<sup>[7]</sup>利用高光谱成像技术对古画颜料成分进行分析,通过最小噪声分离、纯净像元提取、光谱特征拟合,与标准波谱库匹配,正确识别出彩绘艺术画颜料成分。

针对上述现状,本研究拟通过建立光谱分析方法获得非接触式纸质文物高光谱狐斑检测系统,为狐斑病害风险防控提供有效数据支撑,同时,拓展光谱检测技术在文物保护领域中的应用。

## 1 实验部分

### 1.1 纸质文物狐斑检测系统

狐斑检测系统核心高光谱相机(Gaiatracer-V10,测量范围 360~970 nm)(Zolix Ltd., Beijing, China),如图 1 所示,

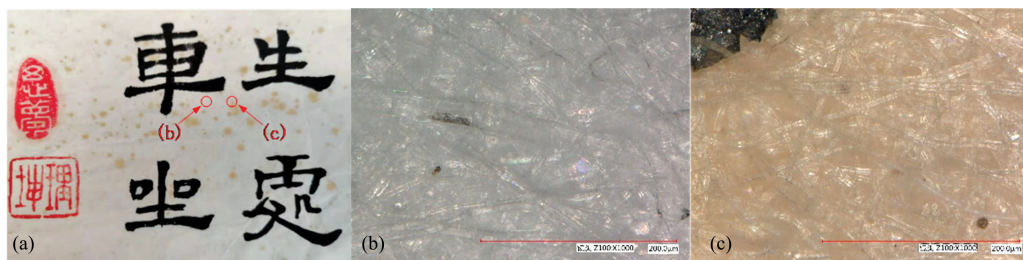


图 2 (a)狐斑样品局部;(b), (c)超景深显微图片

Fig. 2 (a) Part of foxing sample; (b), (c) Super depth of field micrograph

### 1.4 特征波段提取

高光谱数据由 256 个光谱通道数据组成,分辨率达到 2.38 nm,由于 360~450 nm 波段的数据充满噪声信息,不可用于特征判断,于是删除前 38 个波段数据。光谱数据通常表现出高度相关性与大量冗余信息。特征波段提取方法从原有光谱数据中提取与样品狐斑有关的光谱,利用 PCA 减少输入数据维数<sup>[10]</sup>,减少信息的共线性和数据冗余,并对所选特征波段建立的模型进行分析比较,选择出最优的特征波段

系统主要由卤素线型光源(光源光谱范围 350~2 500 nm,功率 300 W)、一维扫描样品台(长宽为 300 mm×300 mm,移动行程 300 mm)、高光谱相机、成像镜头(视场角±7°)及含采集控制软件的 PC 组成,适用于实验室环境下物体鉴别分析。经多次试验,设置参数积分时间 25ms,采集到的图像清晰无失真。

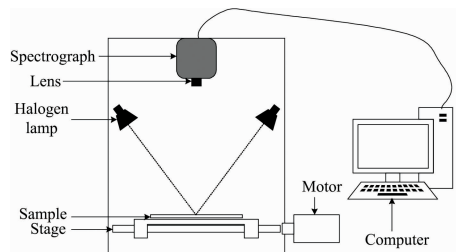


图 1 纸质文物狐斑检测系统

Fig. 1 Paper cultural relics foxing detection system

### 1.2 样本

纸质文物模拟实验样本由重庆中国三峡博物馆提供,样本有轻度狐斑,狐斑为分散的棕黄色点状,样本局部狐斑发生情况较为严重,大面积连接成片如图 2(a)所示。图 2(b)与图 2(c)分别是利用超景深显微系统(VHX-7000,放大倍率×1 000)(Keyence Ltd., Osaka, Japan)拍摄被狐斑侵袭的部分与无狐斑侵袭的部分;两部分均无明显霉菌孢子存在。

### 1.3 图像处理方法

最小噪声分离变换<sup>[8]</sup>(Minimum Noise Fraction Rotation, MNF Rotation)可用于判断图像数据维数,即波段数,分离数据中的噪声,减少处理中的计算量。MNF 本质即是两次叠加的主成分分析<sup>[9]</sup>(principal component analysis, PCA)。第一次 PCA 用于分离数据中的噪声,变换后噪声数据只与最小方差有关。第二次 PCA 是对噪声白化数据的标准主成分变换。

提取方法。

### 1.5 数据建模

随机划分样本集为训练集与测试集后,应用 KNN 算法<sup>[11]</sup>和 BP 神经网络<sup>[12]</sup>,建立模拟纸质文物上狐斑的判别模型;然后将建立的判别模型对测试集样本进行鉴别,并根据鉴别准确率对模型性能进行评价比较,最终选出更优的数据建模方法。

## 2 结果与讨论

### 2.1 模拟纸质文物高光谱图像的光谱分析

图 3 所示为墨迹和健康的纸质区域、被狐斑感染区域的

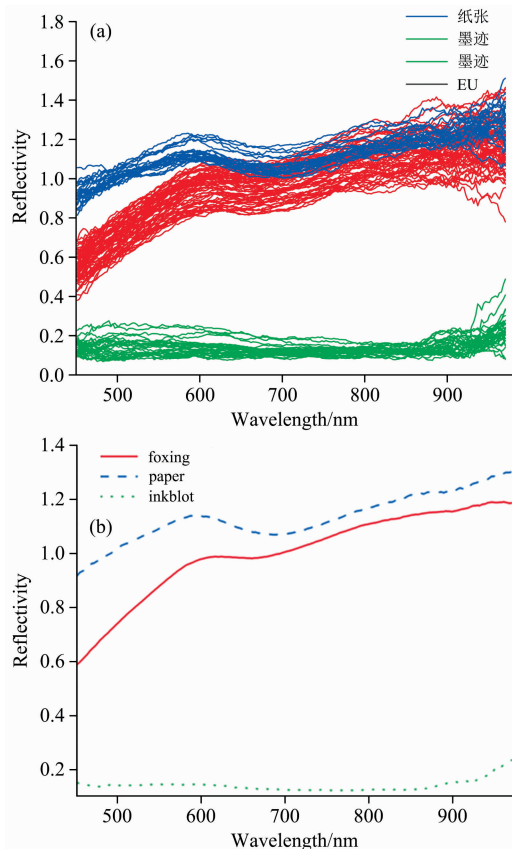


图 3 纸质文物反射光谱与平均反射光谱

Fig. 3 Paper cultural relics reflection spectrum and average reflection spectrum

平均光谱信息，由图可知，健康区域和感染区域的平均反射率差异明显。纸质文物模拟样品光谱在全波长范围内变化较为规律，且在 450~600 nm 波长范围内，模拟样品表面狐斑与背景中的纸张部分、墨迹部分的差异较大。由此可见，基于光谱特性进行纸质文物的狐斑检测分析具有可行性。

纸质文物模拟样本在 464 和 767 nm 波段下的光谱图像如图 4 所示。在图 4(a)464 nm 波段下，样本表面狐斑与纸张背景差异较大，而在图 4(b)767 nm 波段下，表面狐斑已经“埋没”在纸张背景中。这种情况与不同部分的光谱曲线的差异是一致的，但仅凭特殊波段图像不足以提取出纸质文物表面狐斑的特征，需要进一步进行处理。

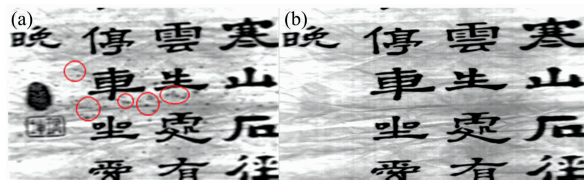


图 4 纸质文物特征波段图像

(a): 464 nm 波段图像; (b): 767 nm 波段图像

Fig. 4 Paper cultural relics feature band image

(a): 464 nm band image; (b): 767 nm band image

### 2.2 基于原始高光谱图像的波段运算

波段运算<sup>[13]</sup>增强可增强目标与背景波谱特征间的微小差别，压制图像中光照差异的影响，突出目标的反射辐射特征，突出目标与背景间的波谱差异。为进一步区分出健康区域与被感染区域，对特征波段图像进行波段运算，如图 5 所示。虽然印章部分与被感染区域有重叠，但被感染区域与健康区域的分布区分比较明显，特别是图 5(c)与图 5(d)，狐斑部分呈现白色，较为突出，易于判别。大多数被感染区域经波段运算后可以被较好地地区分开，波段运算方法纸质文物狐斑识别，有利于实施有效的纸质文物狐斑诊断监测和科学防治。

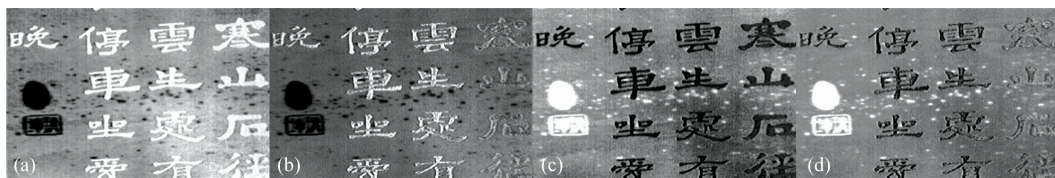


图 5 纸质文物特征波段运算结果

(a): 464~767 nm 图像; (b): 464/767 nm 图像; (c): 767~464 nm 图像; (d): 767/464 nm 图像

Fig. 5 Calculation results of band math of paper cultural relics

(a): 464~767 nm image; (b): 464/767 nm image; (c): 767~464 nm image; (d): 767~464 nm image

### 2.3 基于原始高光谱图像的最小噪声分离变换

图 6 为霉变样本高光谱图像(360~970 nm)进行最小噪声分离(MNF)变换后的图像。选择 MNF 变换后的第一、第二和第七个分量组成 RGB 彩色图像，比起人眼观测的图像，被狐斑感染区域颜色呈现出明显的蓝色，能直观地分辨出狐斑的大小及分布情况，且染病越严重呈现出的颜色越深。作为背景的纸张部分、墨迹部分、印章部分分别呈现出红色、黑色、绿色，呈现出不同的颜色相比波段运算的结果，消除

了印章部分的干扰，对于狐斑的判定更加有利。

### 2.4 纸质文物模拟样本狐斑的判别分析

用狐斑样本的 180 条高光谱数据(450~970 nm)建立狐斑判别模型，随机地分为 120 条数据为训练集，60 条数据为测试集。样本包含纸张部分、墨迹部分与被感染部分共 3 种类型，采用 K-近邻法(KNN)和 BP 神经网络两种分类方式建立狐斑判别模型。将经过预处理以及特征波长提取分析的 120 条高光谱数据作为 KNN 与 BP 神经网络模型的输入构建



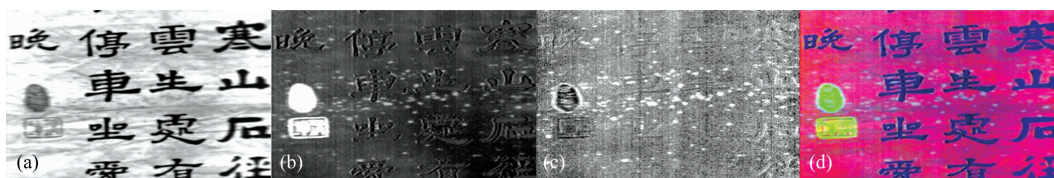


图 6 模拟样本经 MNF 变换后图像

(a): MNF 第一分量; (b): MNF 第二分量; (c): MNF 第七分量; (d): 转换为 RGB 图像

Fig. 6 The images of the simulating sample after MNF transformation

(a): MNF-band 1; (b): MNF-band 2; (c): MNF-band 7; (d): RGB image after MNF transformation

狐斑判别模型, 然后利用测试集的 60 条样本数据对模型进行验证。

结果如表 1 所示, KNN 模型被感染部分测试集判别率为 79.1%、对纸张部分判别率为 57.1%, 对墨迹部分判别率为 60%; BP 神经网络被感染部分测试集判别率为 83.8%、对纸张部分判别率为 63.6%, 对墨迹部分判别率为 83.3%。BP 神经网络判别结果相比 KNN 模型而言, 对染病部分判别正确率更高, 分类效果更好。总体上, 高光谱图像对狐斑部

分判别率更高, 可用于健康样本与狐斑发生样本的识别, 但由于纸张本身存在差异, 对纸张的区分结果欠佳。

### 3 结论

基于高光谱成像研究了受狐斑侵染的纸质文物模拟样本的光谱图像特征, 研究发现狐斑与背景的光谱具有显著差异。对模拟样本不同波段的高光谱图像进行波段运算, 发现波段运算后结果能够提取被狐斑感染区域的特征。对高光谱图像进行最小噪声分离法变换后, 被感染区域区分更加显著。基于 K-近邻法和 BP 神经网络的判别模型结果显示, 高光谱图像对狐斑判别率较高 (>79%), 可用于纸质文物狐斑的判定。研究表明, 高光谱成像可用于纸质文物狐斑的无损检测, 特别是在印章、墨迹、颜料等覆盖后、肉眼难以识别的隐藏的狐斑。本研究可为狐斑的产生及蔓延发展预测、揭示其分布发展规律提供全新的技术手段。后续研究可从光谱图像预处理、有效波段筛选、建模方法优化等方面开展更深入的实验, 以提高检测精度。

表 1 不同部分 KNN 与 BP 神经网络判别分析结果

Table 1 Results of the KNN and BP models for different parts of simulating sample

类型	KNN		BP 神经网络	
	训练集 判别率/%	测试集 判别率/%	训练集 判别率/%	测试集 判别率/%
染病部分	92.2	79.1	95.2	83.8
纸张部分	65.2	57.1	63.2	63.6
墨迹部分	70	60	83.3	83.3
总体	83.3	73.3	88.3	85

### References

- [1] Szulc J, Otlewska A, Ruman T, et al. International Biodeterioration & Biodegradation, 2018, 132: 157.
- [2] CHEN Hui-qiong(陈惠琼). Archives Science Bulletin(档案学通讯), 2018, (5): 81.
- [3] Sclocchi M C, Kraková L, Pinzari F, et al. Microbial Ecology, 2017, 73(4): 815.
- [4] XIE Yan(谢燕). Sciences of Conservation and Archaeology(文物保护与考古科学), 2010, (3): 86.
- [5] WU Wang-ting, ZHANG Chen-feng, GAO Ai-dong, et al(武望婷, 张陈锋, 高爱东, 等). Sciences of Conservation and Archaeology(文物保护与考古科学), 2017, 29(4): 45.
- [6] ZHOU Xin-guang, CHU Hao, WU Lai-ming(周新光, 褚昊, 吴来明). Sciences of Conservation and Archaeology(文物保护与考古科学), 2020, 32(3): 65.
- [7] WU Feng-qiang, YANG Wu-nian, LI Dan(武锋强, 杨武年, 李丹). Acta Mineralogica Sinica(矿物学报), 2014, 34(2): 166.
- [8] Zhang M, Li C, Yang F. Computers and Electronics in Agriculture, 2017, 139: 75.
- [9] Bagnasco L, Zotti M, Sitta N, et al. Talanta, 2015, 144: 1225.
- [10] Kang X, Xiang X, Li S, et al. IEEE Transactions on Geoscience & Remote Sensing, 2017, 55(12): 7140.
- [11] ZHANG Jing-yi, CHEN Jin-chao, FU Xia-ping, et al(张静宜, 陈锦超, 傅霞萍, 等). Spectroscopy and Spectral Analysis(光谱学与光谱分析), 2019, 39(10): 3184.
- [12] Ye Chengming, Liu Xin, Xu Hong, et al. Journal of Zhejiang University-Science A (Applied Physics & Engineering), 2020, 21(3): 240.
- [13] Li J B, Huang W Q, Tian X, et al. Computers & Electronics in Agriculture, 2015, 127: 582.

# Study on Detection Method of Foxing on Paper Artifacts Based on Hyperspectral Imaging Technology

DAI Ruo-chen<sup>1</sup>, TANG Huan<sup>2\*</sup>, TANG Bin<sup>1\*</sup>, ZHAO Ming-fu<sup>1</sup>, DAI Li-yong<sup>1</sup>, ZHAO Ya<sup>3</sup>, LONG Zou-rong<sup>1</sup>, ZHONG Nian-bing<sup>1</sup>

1. Chongqing Key Laboratory of Fiber Optic Sensor and Photodetector, Chongqing University of Technology, Chongqing 400054, China

2. Key Scientific Research Base of Pest and Mold Control of Museum Collections of National Cultural Heritage Administration, Chongqing China Three Gorges Museum, Chongqing 400015, China

3. Chongqing University of Education, Chongqing 400065, China

**Abstract** Affected by preservation conditions, foxing will form on the surface of many paper cultural relics. If effective monitoring and scientific judgment are not carried out, the safety of paper cultural relics will be further affected. For the detection of foxing disease on paper cultural relics, there are problems such as hysteresis and subjectivity. It is difficult to identify the area covered by ink, paint and seals in the painting and calligraphy collection. Therefore, the concept of preventive protection based on cultural relics needs to be developed urgently. Non-destructive testing technology for efficient and accurate identification of foxing. The visible-near-infrared hyperspectral image combines spectrum and image, contains rich spatial information and spectral information, and can achieve lossless batch collection of sample spectral information on the flat. This research proposes a rapid identification method based on hyperspectral imaging technology to detect foxing on paper cultural relics. Obtain hyperspectral images of simulating paper cultural relics at the 360~970 nm. Because the 360~450 nm image is much affected by noise, we choose to exclude this part of the spectral data; select the region of interest, obtain the corresponding average spectral reflectivity, and compare the healthy area with that. In the area of foxing infection, it is found that there is a difference in the spectral curves of the two; near 450~600 nm, the spectral reflectivity of the affected area of foxing is higher than that of the healthy area, and the peak shape appears near 600 nm; and in the range of 600~900 nm, The spectrum of the infected area and the healthy area tends to be stable, and the difference between the two gradually decreases. Select the feature information extracted from the image corresponding to the feature wavelength to build an image recognition model, using band math to observe the image characteristics of foxing, the size and distribution of the foxing can be displayed, but the overlapping parts with the seal and ink, the foxing are covered by the seal and ink, which is difficult to identify; use the minimum noise fraction, although different parts are overlapping, it can find hidden foxing that is difficult to identify with the naked eye; 180 pieces of hyperspectral data (450~970 nm) establish a foxing discrimination model, randomly divided into 120 pieces of data as the training set, and 60 pieces of data as the test set, K-nearest neighbor method and BP neural network are used to establish a paper cultural relics foxing spectrum discrimination model. In general, the two methods have distinguished rates of 73.3% and 85% respectively; Comparing with the K-nearest neighbor model, the BP neural network has a higher overall discrimination rate and a better recognition effect. The results show that hyperspectral imaging can efficiently and accurately identify the foxing of paper cultural relics, provide reliable technical means for the follow-up research on the distribution and development of foxing, and provide guidance for the preservation of cultural relics in the museum.

**Keywords** Paper cultural relics; Foxing; Hyperspectral images; Spectrum; Machine learning classification

(Received Apr. 13, 2021; accepted Jul. 27, 2021)

\* Corresponding authors