NON-DESTRUCTIVE ASSESSMENT OF STONE HERITAGE WEATHERING TYPES BASED ON MACHINE LEARNING METHOD USING HYPERSPECTRAL DATA

Xin Wang[1,2], Yuan Cheng[1,3*], Ruoyu Zhang[1,2], Yue Zhang[1,3], Jizhong Huang[1,3*], Hongbin Yan[4]

[1] Institute for the Conservation of Cultural Heritage, School of Cultural Heritage and Information Management, Shanghai University, Shanghai, China-wangxin321@shu.edu.cn, chengyuan@shu.edu.cn, 21820359@shu.edu.cn, 2019zhangy@shu.edu.cn, hjizhong@shu.edu.cn
[2] School of Mechanics and Engineering Science, Shanghai University, Shanghai, China- wangxin321@shu.edu.cn, 21820359@shu.edu.cn
[3] Key Laboratory of Silicate Cultural Heritage Conservation (Shanghai University), Ministry of Education-chengyuan@shu.edu.cn, 2019zhangy@shu.edu.cn, hjizhong@shu.edu.cn
[4] Yungang Research Institute, Datong, China-435135458@qq.com

Keywords: Stone Cultural Heritage, Hyperspectral Imaging, Machine Learning, Weathering Types, Non-destructive Testing.

ABSTRACT:

Stone cultural heritage is exposed to various environments, resulting in a diverse range of weathering types. The identification of these weathering types is vital for targeted conservation efforts. In this paper, a weathering type classification method based on hyperspectral imaging technology is proposed. Firstly, fresh sandstones are collected from Yungang Grottoes to carry out the simulated weathering experiments, including freeze-thaw cycles and wet-dry cycles with acid, alkali and salt solutions. Subsequently, the hyperspectral imaging system was used to collect the visible-near-infrared (VNIR) and short-wave infrared (SWIR) images of the sandstone samples with different weathering types and degrees. The surface spectral reflectance of sandstone samples with different weathering types were used as training data, with weathering types serving as the labels. Support vector machine (SVM), K-nearest neighbour (KNN), linear discriminant analysis (LDA) and random forest (RF) were used to establish weathering type classification models. The results show that the SVM model and LDA model based on both VNIR and SWIR spectra exhibit outstanding performance, with a best accuracy of 0.994. The framework proposed in this paper facilitates rapid and non-contact assessment of the weathering types of the superficial layers of stone cultural heritage, thereby supporting more targeted conservation work.

1. Introduction

Cultural heritage carries a splendid civilization and transmits history and culture, and stone cultural heritage are an important part of it, which is an important material wealth created by our ancestors and possesses high historical, artistic and scientific values. However, stone cultural heritage have been subjected to severe weathering for a long time under the action of natural camping forces and anthropogenic factors, which directly jeopardizes the historical value and artistic value of this precious historical heritage (Meng et al., 2018). Weathering of stone cultural heritage is caused by a variety of factors, especially by freeze-thaw(Qi et al., 2023), acid (Potysz et al., 2020), alkali (Wray and Sauro, 2017) and salt (Lindström et al., 2016) erosion.

During freezing and thawing, the water in the inner and outer layers of the sandstone undergoes repeated phase transition. This cyclical freezing and thawing process generates accumulated stresses within the sandstone's internal structure over a prolonged period. Consequently, the bonds between mineral particles weaken, leading to particle detachment. This reduces the cementation ability of the sandstone, increases pore space within its structure, and diminishes its mechanical strength (Hall and Thorn, 2011). Acid and alkali gradually dissolve and erode mineral particles and cementation within the sandstone (Liang et al., 2024), This results in the formation of microcracks or pores in sandstone, and the structure is gradually dissolved and damaged (Liao et al., 2024), and the physical and chemical properties of sandstone are also changed, and eventually sandstone produces weathering disease (Dehestani et al., 2020). Salt damage in sandstone occurs due to salt crystallization, primarily involving soluble salts (Jia et al., 2019). These soluble salts dissolve in the water and penetrate into the fissures within sandstone statues. Upon water evaporation, soluble salts precipitate and accumulate on the sandstone surface, exerting crystallization pressure within fissures and leading to internal expansion and micro-crack formation. This process further reduces the cementation ability between sandstone minerals (Jiang et al., 2022). The conservation and restoration strategies for different weathering types of stone cultural heritage are radically different. Therefore, determining the weathering types of stone cultural heritage becomes one of the most vital aspects of carrying out targeted restoration and conservation work. For example, scanning electron microscopy allows observation of the micromorphology of weathered sandstones (Mavris et al., 2012), X-ray diffraction analysis allows detection of mineralogical compositions (Matsuzawa and Chigira, 2020), and FTIR spectroscopy allows identification of chemical bonding of organic and inorganic materials in sandstones (Xin et al., 2021), which are often helpful in determining the type of weathering. However, these detection techniques typically involve destructive sampling. Due to the special characteristics of stone cultural heritage, the application of many destructive and micro-destructive testing techniques is limited. Existing non-destructive techniques, such as ultrasonic method (Liu et al., 2024), magnetization rate determination (Uchida et al., 2021) and X-ray fluorescence analysis (De Deus Ferreira E Silva et al., 2023), have their limitations: these techniques are all contact-based, and they can only detect small, localised areas; they can only obtain basic physical information or chemical composition, making it

* Corresponding author

difficult to capture the weathering types. Therefore, it is necessary to explore and design a fast, contact-free, completely non-destructive weathering type assessment method to provide scientific guidance for more targeted restoration and protection of stone cultural heritage.

Hyperspectral imaging technique enables the acquisition of images across hundreds of spectral channels (Carrasco-García et al., 2023), and its rich spectral information can be used to accurately classify a variety of materials of interest (Guo et al., 2023). Because of this capability, hyperspectral imaging technique has been applied in the field of geology (Chakraborty et al., 2022), but very few studies have explored its use for assessing weathering in stone cultural heritage. Initial work has applied hyperspectral imaging salt-induced weathering of sandstones, demonstrating that the combination of SWIR spectra with machine learning algorithms can effectively assess the degree of salt-induced weathering of sandstones (Yang et al., 2023). Additionally, acid, alkali, and freeze-thaw are also the main triggers of sandstone weathering, but research on these aspects of sandstone weathering using hyperspectral imaging technology is currently lacking. Therefore, we proposed a weathering type evaluation method based on hyperspectral imaging technology and carried out the following work: (i) Fresh sandstone samples were collected from Yungang Grottoes, a world heritage site, to carry out freeze-thaw, acid, alkali and salt simulation weathering tests; (ii) VNIR images and SWIR images of samples with different weathering types and degrees were acquired using a hyperspectral imaging system; (iii) Classification models for weathering types of stone cultural heritage was developed using SVM, KNN, LDA, and RF with the surface spectral reflectance of the sandstone samples with different weathering types as the training features.

2. Materials

2.1 Sample preparation

To carry out the work related to the simulated weathering experiment and spectral data acquisition in the laboratory, sandstones were collected from the Yungang Grottoes site. In order to ensure that the structural properties of the samples closely resemble those of the main body of the grottoes, the samples were collected more than 20 cm away from the surface layer of the sandstone body near the Yungang Grottoes scenic area. The collected sandstone exhibited homogeneous texture with distinct grain characteristics. Using a stone cutter, the samples were prepared into cubic blocks measuring 5×5×5 cm^3.

2.2 Simulated weathering experiments

To investigate the weathering mechanism of Yungang Grottoes sandstones under natural conditions, simulated weathering experiments were carried out involving freeze-thaw, acid, alkali, and salt cycles. The freeze-thaw, acid, and alkali simulated weathering experiments were set to undergo 60 cycles, with sandstone samples subjected to non-destructive testing after every 5 cycles, leaving 3 samples for further analysis. Additionally, 3 fresh blank control samples were included for each experiment, requiring a total of 39 sandstone samples. Due to the severe destruction caused by salt weathering, the number of salt simulation weathering cycles was set to 7 times. Similar to the other experiments, non-destructive testing was performed after each cycle, with 3 samples retained for analysis. This experiment also required 3 fresh blank control samples, totalling 24 sandstone samples. In order to speed up the weathering process, acid, alkali, salt solutions were selected: PH 1 sulfuric acid solution, PH 14 NaOH solution, and saturated Na_2SO_4 solution.

Before the simulated weathering experiments, all sandstone samples were put into an oven to dry at 105°C for 12 h. The simulated weathering experiments were carried out after the samples cooled to room temperature. The steps of freeze-thaw simulated weathering experiment are as follows: (a) Except for the fresh control samples, the remaining 36 sandstone samples were vacuum-saturated with pure water in a desiccator for 4 hours at -0.1 MPa, followed by 4 hours of standing time under atmospheric pressure; (b) The saturated samples were then frozen at -20°C for 4 hours, thawed in pure water at 20°C for 4 hours, completing one freeze-thaw cycle; (c) After every five cycles, samples were dried, subjected to non-destructive testing, and 3 samples were retained for further analysis, while the remaining samples underwent vacuum saturation for the next cycle. The steps for acid, alkali, and salt simulated weathering experiments were similar, involving soaking the samples in corresponding solutions, drying, and non-destructive testing after a certain number of cycles, with 3 samples retained for further analysis until the set number of cycles was completed.

2.3 Hyperspectral image acquisition

The hyperspectral imaging system used in this experiment is shown in Figure 1, which mainly consists of a dark box, VNIR camera, SWIR camera, a linear light source, two 150 W halogen lamps, a computer and a moving platform. VNIR camera has a spatial resolution of 320 × 256 pixels and a spectral resolution of 1.45 nm, capable of capturing hyperspectral images in the range of 400 to 1000 nm, while. The SWIR camera has a spatial resolution of 1632 × 1232 pixels and a spectral resolution of 6.42 nm, capable of capturing hyperspectral images in the spectral range of 1100 to 2500 nm. The height of the camera can be adjusted by a motorized elevation device so that the entire sample can be within the field of view of the spectrometer lens. Spectral acquisition was carried out in a dark environment, with linear light sources and halogen lamps mounted on both sides of the spectrometer lens to provide a uniform light source for VNIR camera and SWIR camera, respectively. In addition, a black light-absorbing cloth was placed underneath the sandstone samples to serve as a background material, which reduces the interference of incident light during data collection. To examine the spectral characteristics of sandstones of different weathering types, spectral images of the sandstone samples were collected after every ten freeze-thaw, acid, alkali and after each salt cycle. When collecting VNIR images, the optimal object distance between the hyperspectral lens and the sample was determined to be 30 cm, the moving platform moving speed was set to 30 mm s^{-1} and the camera exposure time was set to 12 ms; when collecting SWIR images, the optimal object distance was also determined to be 30 cm, the moving platform moving speed was set to 30 mm s^{-1} and the camera exposure time was set to 3 ms.

In order to eliminate the uneven distribution of light intensity in each band and the influence of the noise generated by the dark current in the spectrometer on the spectral data, it is necessary to perform black-white correction on the original image. Standard white and dark reference images are used to correct the original hyperspectral images, and the correction formula is as follows:

$$I_{ref} = \frac{I_{raw} - I_{dark}}{I_{white} - I_{dark}} \quad (1)$$

Where I_{ref} represents the calibrated hyperspectral image, I_{raw} represents the original hyperspectral image, I_{dark} represents the

standard black reference image obtained by completely covering the lens with an opaque black cover, and I_{white} represents the standard whiteboard hyperspectral image.

3. Methods

3.1 Hyperspectral imaging pre-processing

Hyperspectral images contain a large amount of spatial and spectral information, which undoubtedly brings great difficulties to the computation process. Additionally, the inhomogeneity of sandstones also leads to significant differences in the hyperspectral data of each pixel point of the acquired images, which also creates difficulties in using the spectral data of sandstones to identify their weathering types. Spatial binning solves these problems very well by combining a group of adjacent pixels into a single pixel, which can significantly improve the computation speed and improve the signal-to-noise ratio. By combining a group of neighbouring pixels into a single pixel, the computational speed can be significantly improved, and the signal-to-noise ratio can be increased at the same time. Of course, spatial binning may also lead to the loss of detailed features in the image. Taken this into consideration, we fused 100 neighbouring pixels in the image into a single pixel and extracted the mean of the spectral data of the 100 pixels as the feature spectra.

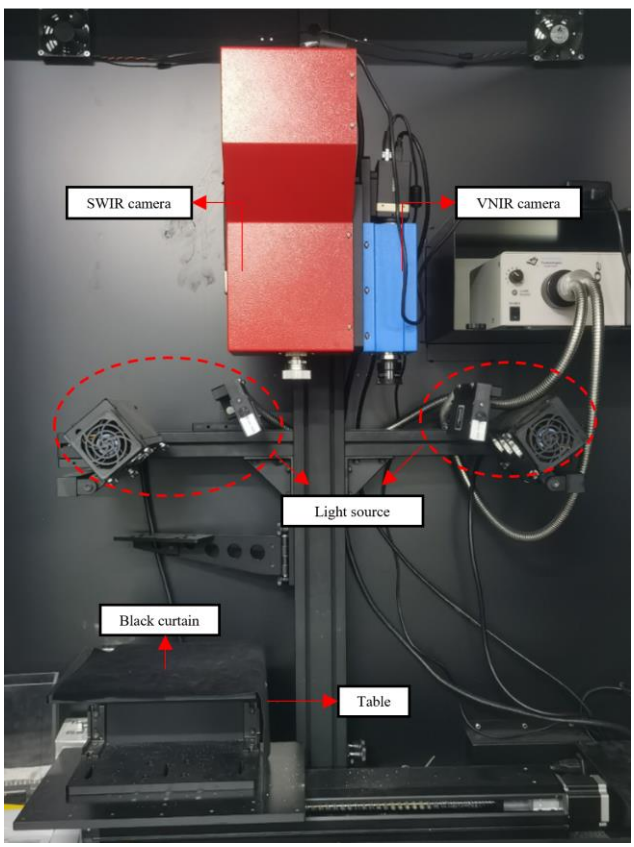


Figure 1. Hyperspectral imaging system

3.2 Spectral data pre-processing

When light interacts with the surface of a sample, it exhibits absorption, transmission, and reflection properties that depend on the surface chemistry of the sample as well as the surface state (Rinnan et al., 2009). The reflectance spectral characteristics of sandstones in the range of 400–2500 nm are primarily determined by their surface chemical composition and microstructure (Uren et al., 2021). Consequently, reflectance spectra can be used as unique identifiers to capture differences in the composition of sandstone surface constituents. However, raw spectral data are often interspersed with noise and some invalid information, which will reduce the accuracy of the analytical results. To minimize the influence of various interfering factors on the spectra, pre-processing of hyperspectral data, such as normalization, multiple scattering correction, and standard normal variation (SNV), is usually performed. By comparing the coefficient of determination and root-mean-square error of the spectral preprocessing methods, SNV was chosen as the spectral data preprocessing method in this paper. The SNV is designed to reduce or eliminate the variations in the spectral data due to the differences in optical path lengths, the particle size effect, the surface scattering and other non-systematic errors, so as to improve the performance of the analytical model, and the basic principle of this method is as follows: The original value will be subtracted from the mean and divided by the standard deviation, and the obtained features satisfy the normal distribution with a mean of 0 and a standard deviation of 1, so that the original distribution of features with large differences in the model has the same weight impact, features between different dimensions have a certain numerical comparability, the size of the resulting parameter values can be a response to the degree of contribution of different features to the sample label, which greatly improves the accuracy of the classifier.

3.3 Algorithms

3.3.1 Support Vector Machine (SVM): SVM is a supervised learning model based on statistical theory, employing interval maximization as its learning strategy, formulated as solving a convex quadratic programming problem. In an N-dimensional space, SVM identifies a hyperplane that distinctly segregates data points from varying categories. A hyperplane is a surface that linearly partitions two segments of n-dimensional data. A line in two dimensions and a plane in three dimensions are referred to as hyperplanes. Theoretical backing for the SVM algorithm is offered in Equation 2.

$$T = \{(x_i, y_i) \mid i = 1, 2, \cdots, n\} \tag{2}$$

In the field of real numbers, the n-dimensional characteristic vectors are x_i and y_i, where x_i, $y_i \in X$ and , $y_i \in \{-1, +1\}$. The linear Eqs. (3), (4) can be used when the data set being analysed is expressed with a linear relationship.

$$w^T x + b = 0 \tag{3}$$

$$w = (w_1; w_2; \cdots; w_d) \tag{4}$$

where w is the hyperplane, b is the distance between the origin and the hyperplane. For this reason, the distance d from the hyperplane to any point X can be expressed according to Equation 5.

$$d = \frac{|\omega^T x + b|}{\|\omega\|} \tag{5}$$

For nonlinear problems, SVMs usually introduce kernel tricks to map the original data into a higher dimensional space to make the data linearly differentiable in this new space. Common kernel functions include linear kernel, polynomial kernel, and radial basis function kernel.

3.3.2 K-Nearest neighbour (KNN): KNN predicts a sample's

category by analysing the categories of its K nearest neighbours (other samples). The algorithm is mathematically described as Equation 6. Consider X as a new observation $x \in X$; the KNN approach initially calculates the distance d(x, x') between x and every training observation $x' \in X$. The distance metric may be any measure of similarity, including Euclidean distance, cosine similarity, or Manhattan distance. Subsequently, the algorithm picks the k training data points closest to x and determines the class of x to be the predominant class among these k neighbours.

$$y' = \text{argmax}(1 \le i \ll k) \sum [y_i = x_i'] \tag{6}$$

where y' is the forecasted output variable of the new observation x, y_i represents the output variable of the ith nearest neighbour, and x_i' the corresponding input variable of the nearest neighbour. The argmax function returns the class label with the highest frequency among the k neighbours. In regression analysis, the predicted output variable y_i is simply the average of the output variables of the k-nearest neighbours.

3.3.3 Linear Discriminant Analysis (LDA): LDA aims to find a linear combination to maximize the inter-class distance between different data while minimizing the intra-class distance between the same class. This method enhances classification accuracy by projecting each sample's output in the dataset with class, minimizing intra-class variance, and maximizing inter-class variance. After the projection process, the projection points of the data of each kind of class are brought closer, while maximizing the distance between different categories of data, which is defined in Equation 7.

$$y_i = W^T x_i \tag{7}$$

Where y_i is the output sample set, W^T is the projection matrix consisting of the eigenvector y_i quantities corresponding to the largest eigenvalues of $S_w^{-1} S_b$ is the input sample set. The intra-class dispersion matrix S_w and the inter-class dispersion matrix S_b, are defined as follows:

$$S_w = \sum_{j=1}^{L} \Sigma_{i=1}^{N_j} (x_j^i - \bar{x}_j)(x_j^i - \bar{x}_j)^T \tag{8}$$

$$S_b = \sum_{j=1}^{L} (\bar{x}_j - \bar{x})(\bar{x}_j - \bar{x})^T \tag{9}$$

where L represents the number of categories, N_j represents the number of samples belonging to the jth category of samples, x_j^i represents the ith sample of the jth category; $\bar{x}_j$ represents the mean vector of the jth category, and $\bar{x}$ represents the vector of the mean values of all samples.

3.3.4 Random Forest (RF): The RF algorithm is a form of ensemble learning where multiple decision trees are generated and combined to improve accuracy. In this process, RF samples and variables multiple times, generates a large number of decision trees (Zhang et al., 2024), and summarizes the results to improve accuracy. For each prediction, all decision trees make predictions sequentially, and the multitude or average of all predictions (corresponding to classification or regression problems, respectively) is used as the final prediction result of the prediction object. Assuming that the size of the original dataset is N, we need to train and construct an RF containing m decision trees, which first needs to be sampled N times from the original dataset (Bootstrap sampling method), constituting a training set with the same size as N (which is not exactly the same as the original dataset). If each sample in the dataset has T attributes, then t ($t \leqslant T$) attributes will be randomly selected during the decision tree splitting inside the RF, and then the split attributes of the nodes will be selected according to a certain strategy, and finally a decision tree will be grown on this training set of size N. This is repeated m times, and a random forest containing m decision trees is trained.

3.4 Model evaluation

Four indicators were used to evaluate the performance of the model: accuracy, precision, recall, and kappa coefficient (K), of which the kappa coefficient ranges from -1 to 1, and the rest of the indicators range from 0 to 1. Accuracy represents the percentage of correctly predicted results in the total number of samples, which is an intuitive reflection of the model's performance; precision indicates the probability of the samples that are actually positive out of the samples predicted to be positive, which reflects the accuracy of the positive class prediction; recall indicates the proportion of all instances that are actually positively classified by the classification model that are correctly identified as positively classified, which reflects the model measure of the model's ability to capture positively classified instances. The definition for the 3 indicators are shown in the following equations.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{10}$$

$$Precision = \frac{TP}{TP + FP} \tag{11}$$

$$Recall = \frac{TP}{TP + FN} \tag{12}$$

where TP is True Positive, the number correctly categorized as positive; TN is True Negative, the number correctly categorized as negative; FP is False Positive, the number incorrectly categorized as positive; and FN is False Negative, the number incorrectly categorized as negative. The Kappa coefficient (K) measures the proportion of actual positive classes that can be correctly identified by the model, which reflects the consistency between the actual effect of the classifier and the effect of random classification. K can be expressed as:

$$K = \frac{p_o - p_e}{1 - p_e} \tag{13}$$

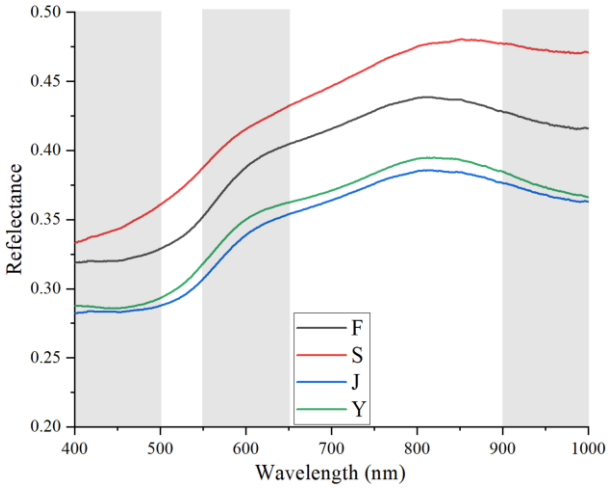
where p_o is predictive consistency, which refers to the number of observations in which two evaluators agree on the assessment as a proportion of the total number of observations, and p_e is chance consistency, which refers to the probability that two evaluators will agree in a randomized situation.

4. Results and discussion

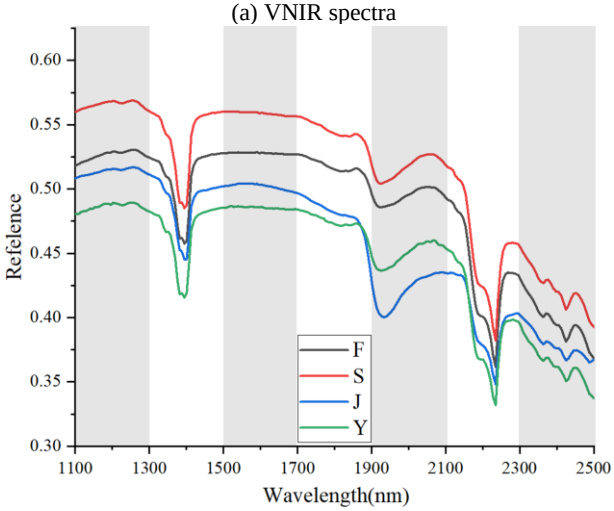
4.1 Spectral characteristics

The corresponding average spectral data of the four weathering types of sandstones are shown in Figure 2. Despite the similarities in morphology among all collected fresh sandstone samples from Yungang Grottoes, distinct spectral differences

exist due to varying weathering types. In the VNIR spectra, alkali-weathered sandstones exhibit the lowest reflectance while acid-weathered sandstones show the highest reflectance. The spectral curve of freeze-thaw-weathered sandstones climbs significantly higher than that of other weathering types in the 550–650 nm wavelength range; the spectral curve of acid-weathered sandstones climbs significantly higher than that of other weathering types in the 400–500 nm wavelength range, and minimizes variation in the 900-1000 nm wavelength range. ; the spectral curve of alkali-weathered sandstones is the smoothest from 400–500 nm wavelength range; the spectral curve of salt-weathered sandstones from 400–500 nm wavelength range is significantly different from that of other weathered sandstones, showing a decreasing and then increasing trend, and the height is similar to that of alkali-weathered sandstones, but the wavelength range after 500 nm is obviously higher than that of alkali-weathered sandstones.



(a) VNIR spectra



(b) SWIR spectra

Figure 2. Average VNIR spectral reflectance curves of sandstones with different weathering types, where F, S, J, and Y represent freeze-thaw, acid, alkali, and salt weathered sandstones, respectively.

In the SWIR spectra, salt-weathered sandstones show the lowest reflectance, acid-weathered sandstones show the highest. Freeze-thaw-weathered sandstones experience the most significant ascent in the range of 1100–1300 nm, maintaining stability in the range of 1500–1700 nm. The spectral curve of acid-weathered sandstones is the most stable in the range of 1100–1300 nm, and shows the trend of levelling off and then decreasing during 1500–1700 nm. The spectral curves of the alkali-weathered sandstones between 1500 and 1700 nm show an increasing and then decreasing trend, and in the 1900–2100 nm and 2300–2500 nm regions are significantly different from the other types of weathering sandstones; the spectral curve of salt-weathered sandstones also show an increasing and then decreasing trend in the 1500–1700 nm region, but the magnitude of the change is smaller than that of alkali-weathered sandstones. There are differences in the spectral characteristics of sandstones of different weathering types, and the differences are more obvious in SWIR spectra, and the further processing of the spectral data can amplify such differences, which also lays a foundation for the establishment of weathering type classification models using sandstones spectral data.

4.2 Modelling based on VNIR spectra and SWIR spectra

In this paper, a total of 19,751 VNIR spectral data and 18,754 SWIR spectral data were acquired and used to build the weathering type classification model, respectively. The dataset was split into a training dataset, comprising 70% of the data, and a test dataset containing the remaining 30%.

The performance of the weathering type classification models based on VNIR spectra is summarized in Tables 1 and 2. Among these models, LDA model exhibited the highest performance and the KNN model demonstrated the lowest performance. Specifically, the LDA model achieved a minimum classification accuracy of 0.83 and 0.81 in the training set and the validation set, the maximum classification accuracy of 0.99, and the average accuracy of 0.922 and 0.912, with the kappa coefficient, the recall and the precision rate all reached more than 0.88, and the model showed high accuracy and stability; the minimum accuracy of the KNN model on the training set and the validation set was 0.67 and 0.55, respectively, and the maximum accuracy was 0.89 and 0.83, respectively, and the average accuracy, kappa coefficient, recall and precision rate were not more than 0.8, and the model has a low accuracy rate.

Algo-rithms	Training set					
	Accuracy			Kappa	Recall	Precision
	Min	Max	Mean			
SVM	0.80	0.99	0.916	0.87	0.90	0.92
KNN	0.67	0.89	0.79	0.71	0.77	0.79
LDA	0.82	0.99	0.922	0.89	0.91	0.93
RF	0.64	0.94	0.80	0.71	0.77	0.84

Table 1. Prediction results of weathering type classification models based on VNIR spectra on training set

Algo-rithms	Validation set					
	Accuracy			Kappa	Recall	Precision
	Min	Max	Mean			
SVM	0.78	0.99	0.904	0.86	0.89	0.90
KNN	0.55	0.83	0.71	0.58	0.68	0.73
LDA	0.81	0.99	0.912	0.87	0.90	0.92
RF	0.60	0.89	0.71	0.58	0.67	0.77

Table 2. Prediction results of weathering type classification models based on VNIR spectra on validation set

The results of the weathering type classification models based on VNIR spectra are shown in Tables 3 and 4, where the LDA model again demonstrates superior performance, while RF model exhibits relatively poorer performance. The LDA model achieves a minimum classification accuracy of 0.99 in the training set and a maximum classification accuracy of 1.0 in the test, with an

average classification accuracy of over 0.995, and the kappa coefficients, recalls, and precision rates of 0.99, which show superior performance. Conversely, the RF model displays lower accuracy, with minimum accuracies of 0.78 and 0.67, and maximum accuracies of 1.0 and 0.99. However, the average classification accuracy remains above 0.89, and the kappa coefficients, recall, and precision rates are 0.84 for both the training and validation sets, indicating relatively good performance.

Among the models construed based on the VNIR and SWIR spectra, those employing SVM and LDA consistently exhibit superior performance, while the RF and KNN models demonstrate relatively poorer performance. The accuracy of the weathering type classification model construed based on SWIR spectra is significantly higher than that of the model construed based on VNIR spectra, which also corresponds to the more obvious differences in SWIR curve features.

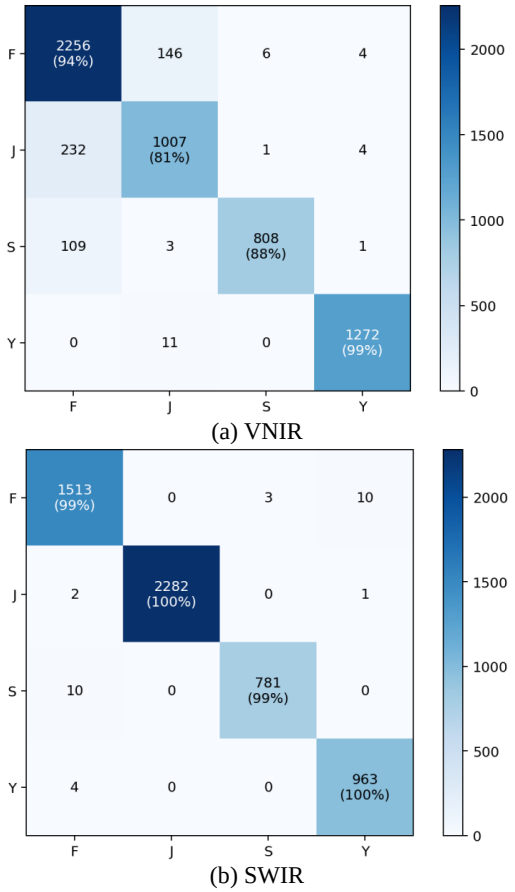
Algo-rithms	Training set					
	Accuracy			Kappa	Recall	Precision
	Min	Max	Mean			
SVM	0.98	1.0	0.988	0.98	0.98	0.98
KNN	0.80	0.99	0.92	0.86	0.89	0.90
LDA	0.99	1.0	0.996	0.99	0.99	0.99
RF	0.78	1.0	0.92	0.86	0.88	0.91

Table 3. Prediction results of weathering type classification models based on SWIR spectra on training set

Algo-rithms	Validation set					
	Accuracy			Kappa	Recall	Precision
	Min	Max	Mean			
SVM	0.97	1.0	0.986	0.98	0.98	0.98
KNN	0.71	0.99	0.89	0.85	0.85	0.87
LDA	0.99	1.0	0.994	0.99	0.99	0.99
RF	0.67	0.99	0.89	0.84	0.84	0.87

Table 4. Prediction results of weathering type classification models based on SWIR spectra on validation set

The LDA model based on both VNIR and SWIR spectra achieved the highest classification accuracy in detecting weathering types. To further analyze the performance of the LDA models, confusion matrices were constructed. As shown in Figure 3, the prediction accuracy of different weathering types is shown on the main diagonal of each matrix. The LDA model based on the full VNIR band exhibited high recognition rates for freeze-thaw and salt weathering, with relatively lower recognition rates for alkali and acid weathering. Salt weathering had the highest recognition rate, with an accuracy of 0.99, possibly due to the significant differences between the spectral curves of salt-weathered sandstones in the 400–500 nm range compared to other weathering types. Alkali weathering had the lowest recognition rate, with an accuracy of 0.81. For misclassified spectral data, most of the alkali weathering was classified as freeze-thaw weathering, while freeze-thaw weathering was misclassified as alkali weathering, and acid weathering was misclassified as freeze-thaw weathering. As observed in Figure 2 (a), the spectral curves of freeze-thaw and alkali weathering types indeed exhibited significant similarities, while the spectral curves of acid -weathered sandstones closely resembled the trend of freeze-thaw-weathered sandstones. The models based on SWIR spectra demonstrated very high classification accuracy, with only a small portion of spectral data misclassified as other weathering types. Additionally, this model exhibited the highest recognition rate for alkali weathering, possibly due to differences in spectral characteristics between alkali-weathered sandstones and other weathered sandstones in the 1900–2100 nm wavelength range.



(a) VNIR

(b) SWIR

Figure 3. Confusion matrix of LDA weathering type classification model based on VNIR and SWIR spectra

5. Conclusions

In this paper, fresh sandstones from Yungang Grottoes were collected to carry out freeze-thaw, acid, alkali and salt simulated weathering experiments. Utilizing hyperspectral imaging technology, spectral data of sandstones with different weathering types were collected. Based on the spectral data, weathering type classification models using machine learning were meticulously constructed and rigorously evaluated. The results indicate that there are differences in the spectral characteristics of sandstones with different weathering types, and the SVM model and LDA model construed based on VNIR spectra and SWIR spectra exhibit superior performance. The framework proposed in this paper is helpful for rapid, non-contact assessment of weathering types in the surface of stone cultural heritage, supporting more targeted conservation efforts. In future work, additional spectral data from sandstones with different weathering types in other regions will be included to enhance the model's universality.

Acknowledgements

This research was supported by Science and Technology Major Special Program Project of Shanxi Province (202201150501024), Research Project of Shanxi Cultural Heritage Bureau (2024KT19) and National Science Foundation of China (42301149, 42007265).

References

Carrasco-García, M.G., Rodríguez-García, M.I., González-Enrique, J., Ruiz-Aguilar, J.J., Turias-Domínguez, I.J., 2023: Hyperspectral technology for oil spills characterisation by using feature selection. *Transportation Research Procedia*., 71, 117-123. doi.org/10.1016/j.trpro.2023.11.065.

Chakraborty, R., Kereszturi, G., Pullanagari, R., Durance, P., Ashraf, S., Anderson, C., 2022: Mineral prospecting from biogeochemical and geological information using hyperspectral remote sensing - Feasibility and challenges. *J. Geochem*., 232, 106900. doi.org/10.1016/j.gexplo.2021.106900.

De Deus Ferreira E Silva, J., Júnior, J.M., Vieira Da Silva, L.F., Chitlhango, A.P., Silva, L.S., De Bortoli Teixeira, D., Moitinho, M.R., Fernandes, K., Ferracciú Alleoni, L.R., 2023: Magnetic signature and X-ray fluorescence for mapping trace elements in soils originating from basalt and sandstone. *Chemosphere*., 341, 140028. doi.org/10.1016/j.chemosphere.2023.140028.

Dehestani, A., Hosseini, M., Beydokhti, A.T., 2020: Effect of wetting–drying cycles on mode I and mode II fracture toughness of sandstone in natural (pH = 7) and acidic (pH = 3) environments. *Theoretical and Applied Fracture Mechanics*., 107, 102512. doi.org/10.1016/j.tafmec.2020.102512.

Guo, Y., Mokany, K., Ong, C., Moghadam, P., Ferrier, S., Levick, S.R., 2023: Plant species richness prediction from DESIS hyperspectral data: A comparison study on feature extraction procedures and regression models. *ISPRS Journal of Photogrammetry and Remote Sensing*., 196, 120–133. doi.org/10.1016/j.isprsjprs.2022.12.028.

Hall, K., Thorn, C., 2011: The historical legacy of spatial scales in freeze–thaw weathering: Misrepresentation and resulting misdirection. *Geomorphology*., 130, 83–90. doi.org/10.1016/j.geomorph.2010.10.003.

Jia, M., Liang, J., He, L., Zhao, X., Simon, S., 2019: Hydrophobic and hydrophilic SiO_2-based hybrids in the protection of sandstone for anti-salt damage. *J. Cult. Herit*., 40, 80–91. doi.org/10.1016/j.culher.2019.06.001.

Jiang, X., Li, C., Zhou, J.-Q., Zhang, Z., Yao, W., Chen, W., Liu, H.-B., 2022: Salt-induced structure damage and permeability enhancement of Three Gorges Reservoir sandstone under wetting-drying cycles. *Int. J. Sandstone Mech. Min. Sci*., 153, 105100. doi.org/10.1016/j.ijrmms.2022.105100.

Liang, Y., Huo, R., Song, S., Song, Z., Li, G., Mu, Y., 2024: Effect of the surface-area-to-volume ratio on dissolution and deterioration of acid-corroded sandstone. *J. Build. Eng*., 86, 108789. doi.org/10.1016/j.jobe.2024.108789.

Liao, R., Chen, W., Liu, P., Wang, L., Xia, Q., Yang, X., He, Y., Guo, Q., 2024: Mineral transformation, element transport and hydrological impact in weathering at the Bingling Temple Grottoes: Implications for weathering in alkaline environments in NW China. *CATENA*., 239, 107966. doi.org/10.1016/j.catena.2024.107966.

Lindström, N., Talreja, T., Linnow, K., Stahlbuhk, A., Steiger, M., 2016: Crystallization behavior of Na_2SO_4–$MgSO_4$ salt mixtures in sandstone and comparison to single salt behavior. *Applied Geochemistry*., 69, 50–70. doi.org/10.1016/j.apgeochem.2016.04.005.

Matsuzawa, M., Chigira, M., 2020: Weathering mechanism of arenite sandstone with sparse calcite cement content. *CATENA*., 187, 104367. doi.org/10.1016/j.catena.2019.104367.

Mavris, C., Götze, J., Plötze, M., Egli, M., 2012: Weathering and mineralogical evolution in a high Alpine soil chronosequence: A combined approach using SEM–EDX, cathodoluminescence and Nomarski DIC microscopy. *Sedimentary Geology*., 280, 108–118. doi.org/10.1016/j.sedgeo.2012.04.008.

Meng, T., Lu, Y., Zhao, G., Yang, C., Ren, J., Shi, Y., 2018: A synthetic approach to weathering degree classification of stone relics case study of the Yungang Grottoes. *Heritage Science.*, 6, 1. doi.org/10.1186/s40494-017-0165-y.

Potysz, A., Bartz, W., Zboińska, K., Schmidt, F., Lenz, M., 2020: Deterioration of sandstones: Insights from experimental weathering in acidic, neutral and biotic solutions with Acidithiobacillus thiooxidans. *Constr. Build. Mater*., 246, 118474. doi.org/10.1016/j.conbuildmat.2020.118474.

Qi, C., Li, Q., Ma, X., Guo, W., 2023: Deterioration of fresh sandstone caused by experimental freeze-thaw weathering. *Cold Regions Science and Technology*., 214, 103956. doi.org/10.1016/j.coldregions.2023.103956.

Rinnan, Å., Berg, F.V.D., Engelsen, S.B., 2009: Review of the most common pre-processing techniques for near-infrared spectra. *TrAC Trends in Analytical Chemistry*., 28, 1201–1222. doi.org/10.1016/j.trac.2009.07.007.

Uchida, E., Watanabe, R., Cheng, R., Nakamura, Y., Takeyama, T., 2021: Non-destructive in-situ classification of sandstones used in the Angkor monuments of Cambodia using a portable X-ray fluorescence analyzer and magnetic susceptibility meter. *Journal of Archaeological Science: Reports*., 39, 103137. doi.org/10.1016/j.jasrep.2021.103137.

Uren, A.L., Laukamp, C., George, A.D., Occhipinti, S.A., Aitken, A.R.A., 2021: Inferring sandstone grain size using spectral datasets: An example from the Bresnahan Group, Western Australia. *Remote Sensing of Environment*., 252, 112109. doi.org/10.1016/j.rse.2020.112109

Wray, R.A.L., Sauro, F., 2017: An updated global review of solutional weathering processes and forms in quartz sandstones and quartzites. *Earth-Science Reviews*., 171, 520–557. doi.org/10.1016/j.earscirev.2017.06.008

Xin, L., Li, C., Liu, W., Xu, M., Xie, J., Han, L., An, M., 2021: Change of sandstone microstructure and mineral transformation nearby UCG channel. *Fuel Processing Technology*., 211, 106575. doi.org/10.1016/j.fuproc.2020.106575

Yang, H., Chen, C., Ni, J., Karekal, S., 2023. A hyperspectral evaluation approach for quantifying salt-induced weathering of sandstone. *Science of The Total Environment*., 885, 163886. doi.org/10.1016/j.scitotenv.2023.163886Zhang, X., Shen, H., Huang, T., Wu, Y., Guo, B., Liu, Z., Luo, H., Tang, J., Zhou, H., Wang, L., Xu, W., Ou, G., 2024. Improved random forest algorithms for increasing the accuracy of forest aboveground biomass estimation using Sentinel-2 imagery. *Ecological Indicators*., 159, 111752. doi.org/10.1016/j.ecolind.2024.111752