Automatic rock classification of LIBS combined with 1DCNN based on improved Bayesian optimization

GUANGDONG SONG,[1] SHENGEN ZHU,[1,*] WENHAO ZHANG,[2] BINXIN HU,[1] FENG ZHU,[1] HUA ZHANG,[1] TONG SUN,[3] AND KENNETH TV GRATTAN [1,3]

[1] *Laser Institute, Qilu University of Technology (Shandong Academy of Sciences)), Jinnan 250014, China*
[2] *Qilu University of Technology, International School for Optoelectronic Engineering, Jinnan 250014, China*
[3] *School of Science & Technology, City, University of London, Northampton square, London, EC1V 0HB, United Kingdom*
**10431210478@stu.qlu.edu.cn*

Abstract: To achieve automated rock classification and improve classification accuracy, this work discusses an investigation of the combination of laser-induced breakdown spectroscopy (LIBS) and the use of one-dimensional convolutional neural networks (1DCNN). As a result, in this paper, an improved Bayesian optimization algorithm has been proposed where the algorithm has been applied to automatic rock classification, using LIBS and 1DCNN to improve the efficiency of rock structure analysis carried out. Compared to other algorithms, the improved Bayesian optimization method discussed here allows for a reduction of the modelling time by about 65% and can achieve 99.33% and 99.00% for the validation and test sets of 1DCNN.

1. Introduction

Coal remains one of the main energy sources in wide use in China. In 2020, China's coal consumption accounted for 56.8% of its total energy consumption,[1] while in 2021, there was a small decrease in consumption (to 56.0% of the total energy consumption), and raw coal power generation to ~67.0% of the total power generation (http://www.stats.gov.cn/xxgk/jd/sjjd2020/202210/t20221008_1888971.html). There is a large number of coal mines operating in China and most involve underground mining. To achieve efficient production from such mining and reduce the risk of accidents for underground workers, intelligent mining and unmanned mining are seen as the way forward. Rock and lithology analysis plays an important guiding role in many aspects of the field, such as mining and geological disaster analysis. Through lithology analysis, real-time geological data can be provided for unmanned mining and intelligent mining.[2] The traditional lithology judgment method used relies on the experience of staff to judge the appearance and physical properties of the ore: a task which requires an extremely high level of professional expertise and identification experience from the staff involved, but unfortunately the efficiency of manual identification is not sufficiently high. With the continuous development of compositional analysis techniques as has been seen in the field of spectroscopy, a series of techniques such as X-ray fluorescence spectroscopy (XRF), X-ray diffraction analysis (XRD), and gamma spectroscopy are widely used for rock classification. However, the sample production process using these methods is relatively complex and the measurement time is relatively long. Therefore, the need is for rapid, accurate and in situ geological rock or mineral identification which the technique discussed in this paper can offer.[3-5]

Laser Induced Breakdown Spectroscopy (LIBS) is a laser-based optical spectroscopy technique which operates by shining high-energy pulses on the surface of the sample to create gasification of the sample, to generate a plasma which is studied. This emits a spectrum during

the process of plasma diffusion and cooling, allowing the collection of spectral data with a spectrometer, and then conducting a qualitative analysis of the material tested, based on the wavelength and intensity of the spectral peaks seen.[6-8] This approach has the advantage of multi-form analysis and is fast in operation. Further, the approach is non-destructive, shows a low detection limit, with there being no need to use vacuum environment. All this points to the LIBS technique being well suited to field detection applications. Therefore, LIBS is a technique that is widely used in the fields of environmental monitoring,[9-11] metallurgy,[12-14] medicine,[15-18] food,[19, 20] heritage science,[21] planetary exploration missions,[22, 23] and mineralogy.[24-32]

Wang C et al. used PCA to reduce the dimension of each spectral signal, and then used linear discriminant analysis, and a random forest and a support vector machine approach to classify the spectral data, after dimensional reduction. The results indicate that SVM could well be applied to LIBS classification of rock.[33] We note that El-Saeid et al. used PCA and Graph Theory methods to classify spectra obtained in rocks using two methods (standard LIBS, and Nanoparticle-Enhanced LIBS), showing that excellent classification of the rocks analyzed (with more than 99% of the spectra correctly classified) could be obtained using standard LIBS, coupled to Graph Theory analysis.[34] Yelameli et al. studied the effect of increasing the number of shots per rock and the detrend operation, showing that the number of dimensions could be effectively reduced by applying PCA. The results obtained indicate that the SVM algorithm with the detrend operation, combined with a specific number of shots, creates a good rock classification effect (with an accuracy greater than 95%).[35] Janovszky et al. demonstrated that LIBS mapping, with spatially resolved local analysis offers an efficient and practical approach for the classification of mineral grains. The results indicate that the classification accuracy obtained is better than 92%, using random forest and linear discriminant analysis. Direct classification by evaluating the presence of feature elements is a powerful approach.[36] In this paper, a classification model optimization algorithm has been proposed and applied to the rock recognition method, combining LIBS and 1DCNN. The positive effect is to reduce the modeling time of the classification algorithm and to improve the efficiency of the rock classification, as an effective means to improve the classification accuracy.

In this paper, LIBS technology, combined with a variety of machine learning algorithms (principal component analysis, grid search cross validation, random search cross validation, Bayesian optimization based on Gaussian process, improved Bayesian optimization, support vector machine, one-dimensional convolutional neural network, etc.), is expanded to use to carry out comparative experiments on rock classification for the mining industry. The goal is to automate rock classification and provide reference data that will enhance lithology identification in the mining industry, through a series of straightforward steps that can easily be applied, from inputting raw LIBS rock spectral datasets to outputting rock classification results. In use, a pre-processing operation first is selected to obtain the relevant LIBS spectral data, where the pre-processed spectral data are ‘dimensioned-down’ through a principal component analysis technique, following which the corresponding number of principal components is selected as the input variable for the classifier, used for different classifier models. Three optimizers were constructed and used to conduct comparative experiments on the classifier model, including the combination of grid search and random search optimization, Bayesian optimization based on Gaussian process and improved Bayesian optimization,[37] which can improve the identification of the samples under consideration.

2. Experiment

2.1 Experimental setup used

Fig.1 and Fig.2 respectively illustrate a schematic diagram and show photographs of the LIBS experimental setup. The laser source used in this work was a Q-switched Nd: YAG (Q-smart 450) pulsed laser (from Quantel, France), which was used to generate a laser beam with pulse repetition rate of 8Hz, a pulse width of 7ns and an output wavelength of 1064nm. A laser beam

is focused on the surface of the rock to be tested (using a quartz lens with a focal length of 100mm) to generate the plasma. When the species in the plasma make a transition from an excited to a low-level or ground state, light is emitted and this is collected by using a lens (with a diameter of 10mm and a focal length of 15mm). Here the optical signal is collected using a fiber optic probe and fed to an 8-channel spectrometer (Avantes-usb8) equipped with a CCD detector, where it is converted into an electrical signal and transmitted to a computer for analysis and processing. The pulsed laser used emits a laser pulse energy of 35mJ, where the wavelength range of the spectrometer used was 181-673nm, thus covering a wide range. Further, the delay time of the acquisition of the spectrum is 160μs, the integration time is 1.05ms, and the spectral resolution of the different channels of the spectrometer is 0.058-0.068nm.

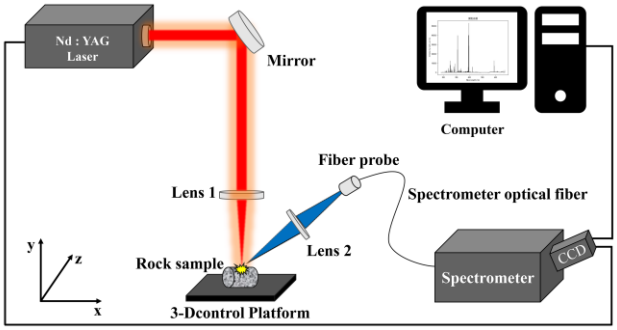

Fig. 1. Schematic diagram of the experimental setup.

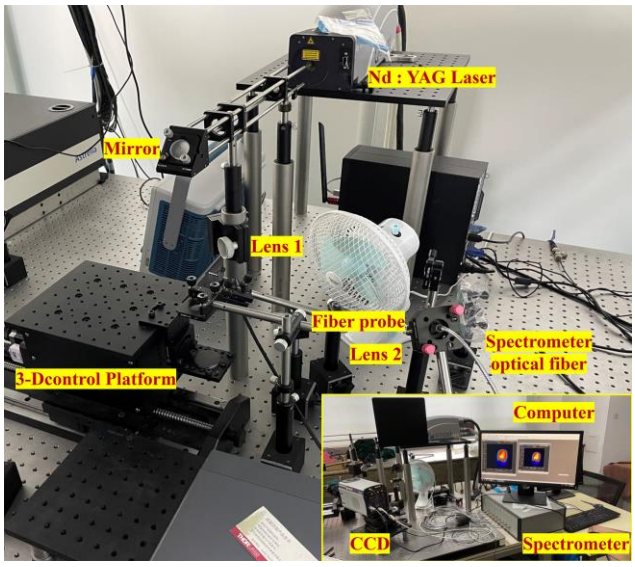

Fig. 2. Photographs of the experimental setup.

2.2 Sample preparation

In this paper, 10 different and representative types of sedimentary rock samples, taken from the rock that forms the roof of Yanzhou coal mine (owned and operated by Shandong Energy Group) were selected for analysis, in consultation with mining experts: these being Siltstone, Oil shale, Argillaceous siltstone 1, Gritstone 1, Argillaceous siltstone 2, Mudstone 1, Red sandstone, Mudstone 2, Gritstone 2, and Fine sandstone, as detailed in Table 1 and Fig.3. To remove dust or any environmental effects from the samples, several ‘cleaning’ shots (i.e. 10 laser pulses applied to six different locations of the samples) were fired at them, before any data sets were acquired. To overcome the heterogeneity of the rocks, 40 laser irradiations were

performed on the same acquisition site in the course of the experimental work – where here the laser focus position was controlled by a laser ranging feedback control 3D platform. In order to eliminate the apparently unstable plasma spectra, it was necessary to eliminate the anomalous spectra and to do so, the specific process was as follows:

① The M spectral information of a measurement sample, after 40 repeated measurements, was used as the independent variable matrix to obtain a 40×M matrix, following which the average value of each column was calculated to obtain a matrix, of row vector 1×M. Each element of this row vector is the average spectral intensity of the corresponding row;

② The sum of the squares of the differences between each measurement value and the mean value were calculated;

③ The measured spectra which were larger than 1.1 times the mean of the sum of squares were removed to complete the screening of abnormal spectral data. This process was found to remove about two-fifths of the original data. Finally, 960 sets of spectral data were obtained as the original data set used in the subsequent analysis.

To eliminate the influence of the time taken for the work to be done on the spectral data collected, 50 sets of data were extracted from each rock sample in sequence. Thus, in total, 500 spectral data sets were formed from the 10 different rock samples (seen in Table 1) to ensure that each was equally sampled.

Table 1. Rock category information: illustrating Rock Category Label and Rock Type Name

Rock Category Label	Rock Type Name
Rock 0	Siltstone
Rock 1	Oil shale
Rock 2	Argillaceous siltstone 1
Rock 3	Gritstone 1
Rock 4	Argillaceous siltstone 2
Rock 5	Mudstone 1
Rock 6	Red sandstone
Rock 7	Mudstone 2
Rock 8	Gritstone 2
Rock 9	Fine sandstone

Fig. 3. Photographs of 10 kinds of rocks.

2.3 Spectral data pre-processing

The experimental system error of the instrument used can interfere with the LIBS spectrum received. This can result from changes in the external environment, random noise in the spectral line signature seen and the diffuse reflection of the solid. As a result, the spectral data collected inevitably contain irrelevant information, such as from the pump light, from stray light that has been collected and molecular vibration effects from the spectral species involved. Spectral data from material other than the rock samples themselves i.e. background spectral data will be superimposed on, and can potentially interfere with, the collected spectral data. All this will negatively affect the accuracy of the classification of the rock samples and slow the speed of iteration of the classification model created. Therefore, it is necessary to eliminate signals which can cause such errors in the spectral signals received from the LIBS process, while still retaining the characteristic spectra of the samples themselves, to provide an appropriate basis for the subsequent analysis of the LIBS data.

In this paper, it is recognized that missing values in the data sets will occur and this is undesirable. However, to compensate, and thus to choose appropriate values to complete the sets, a Savitzky Golay smoothing filter (polynomial order is 4, window width is 5), and a Multiplicative Scatter Correction (MSC) approach, coupled with normalization as the optimal pre-processing method, with basic pre-processing was carried out to enable the results from comparative experiments to be evaluated. The experimental process used is shown schematically in Fig. 4.

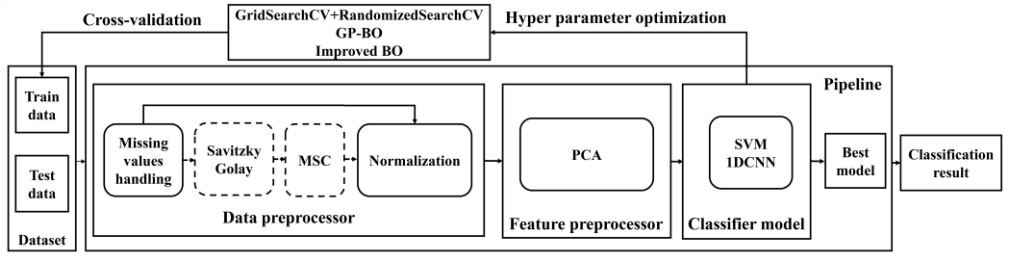

Fig. 4. Flow chart of the experiment.

To prevent data leakage, only the training set, in the pre-processing and cross-validation hyperparameter optimization, is processed. The validation set and test set are used for hyperparameter optimization and model testing experiments through the pipeline mapping transformation. The set is transformed through a pipeline map for hyperparameter optimization and model testing experiments. In this approach, 35 sets of spectral data were randomly selected from the 50 sets of spectral data from each type of rock studied (see Table 1), to form the training set. This training set, comprising 350 sets of spectral data, was used to train the model and to optimize the hyperparameters; the remaining 15 sets of spectral data (from each type of rock) form the test set. As a result, 150 sets of spectral data were used finally to test the accuracy of the classification model used. The algorithms described in this paper were all implemented in Windows10, 64-bit system, Python 3.9.7 version, and using the Jupyter Lab 3.3.2 development environment.

The pre-processing operation used can be described as follows. Firstly, the negative value in the spectral intensity value of the data to be processed is defined as NaN, and then converted to a value of 0, following which the Savitzky Golay was used for smoothing filtering, and then the effect of scattering was eliminated by use of the MSC process. Here Multiplicative Scatter Correction (MSC) is one of the common methods of spectral data pre-processing used. Due to the diffuse reflection and surface inhomogeneity of the solid sample on which the measurement was carried out, the spectral differences caused by the presence of the different scattering levels could be eliminated by the use of the MSC algorithm. Thus, the phenomenon of baseline drift in the spectrum could be dealt with, thereby enhancing the correlation between the spectrum obtained and the original data. The specific process by which MSC is carried out can be described as follows.

1) The mean value of the spectral data used was taken as the 'standard spectrum'. Thus, the LIBS standard spectra for the 10 rock types used in this study are shown in Fig. 5 below.

$$\bar{A} = \frac{\sum_{i=1}^{n} A_i}{n} \tag{1}$$

2) A univariate linear regression is performed on the average spectrum by importing the Linear Regression module in the sklearn library, and the linear translation and tilt offset of each spectral data set, relative to the average spectrum are obtained, by solving the least squares problem.

$$A_i = m_i \bar{A} + b_i \tag{2}$$

3) A Multivariate Scattering Correction is applied for each spectral data point: the corrected spectral data are obtained by subtracting the linear shift from the spectral data and dividing by the regression coefficient.

$$A_{i(MSC)} = \frac{(A_i - b_i)}{m_i} \tag{3}$$

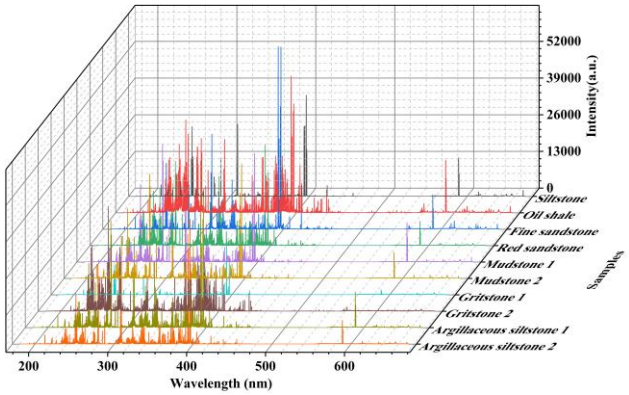

Fig. 5. LIBS standard spectra for 10 rock types.

Each spectral data point is corrected, with reference to the average spectrum (this process does not affect the information each absorbs), to improve the signal-to-noise ratio of the spectral data overall. The spectral data are then normalized by use of the following formula:

$$x_1 = \frac{x - x_{min}}{x_{max} - x_{min}} \tag{4}$$

$$x_2 = x_1 * (Max - Min) + Min \tag{5}$$

In the above formula, x represents the value of the intensity of each group of spectral data obtained at each wave point; x_{min} is the minimum intensity of this group of spectral data; x_{max} is the maximum intensity of this group of spectral data; this is normalized to the interval 1, and x_2 represents the normalized spectral data.

2.4 Models and Algorithms

Principal Component Analysis (PCA) allows for creating a data dimensionality reduction, and thus feature extraction, by mapping n-dimensional features to k-dimensions, which are brand-new orthogonal features, also known as principal components. In this experiment carried out, each group of spectra contains 10239 wave

points. If this is used as the input of the classification model, problems such as dimensionality disaster, low efficiency of the classification model, and long training time can be seen.

The experimental work carried out and reported in this paper uses two machine learning classification models (SVM, 1DCNN), combined with three hyperparameter optimization methods (GS-RS, GP-BO and Improved BO) to conduct comparative experiments on data that are pre-processed, or not pre-processed. Support Vector Machine (SVM) is a machine learning algorithm, based on Statistical Learning Theory (STL) (and created by Vapnik) to deal with binary classification problems. The SVM approach realizes the classification of samples by finding the optimal classification hyperplane that satisfies the constraints. In the linear classification, the classification surface is selected by taking the farthest distance from the two samples. In the case of a nonlinear classification, it is processed through a transformation in high-dimensional space. To reduce the amount of calculations needed and the complexity of the model, a kernel function is introduced to replace the dot product process in the high-dimensional feature space. The kernel function formula is given by.

$$K\left[x^{(i)}, x^{(j)}\right] \equiv \Phi\left[x^{(i)}\right]^{T} \cdot \Phi\left[x^{(j)}\right] \tag{6}$$

In the above Equation (6), $\Phi\left[\mathrm{x}^{(\mathrm{i})}\right]^{\mathrm{T}}$ and $\Phi\left[x^{(j)}\right]$ are $\mathrm{n}-$dimensional features $x^{(i)}$ and $x^{(j)}$ are mapped to $\bar{n}$ dimensional features. Following that, the SVM classifier model was optimized by adjusting the penalty coefficient, C, and the kernel function parameter, gamma. In this paper, a Radial Basis Function (RBF) kernel function was used, which is given as shown below (where δ is the kernel width, $\delta > 0$)

$$K\left[x^{(i)}, x^{(j)}\right] = e^{-\frac{\|x^{(i)} - x^{(j)}\|^2}{2\delta^2}} \tag{7}$$

A Convolutional Neural Network (CNN) is a deep neural network model with the characteristics of neuron weight sharing and local connection. It is important because it has powerful feature extraction capabilities in high-dimensional data. When the traditional fully connected network model processes data, due to the large number of parameters present, problems such as a large amount of calculation needed, the low model efficiency, and local invariance are seen. Convolutional Neural Networks reduce the number of parameters by locally connecting each neuron, sharing weights for each group of connections, and adding pooling layers, while improving the robustness of the model and the ability of the network to generalize. For one-dimensional spectral data from the rock samples considered, it creates a kind of multivariate sequence data, using a one-dimensional convolutional neural network (1DCNN) model. By performing the one-dimensional convolution operation on the input spectral data, the feature information can then be extracted to create an effective classification from the different types of rocks considered.

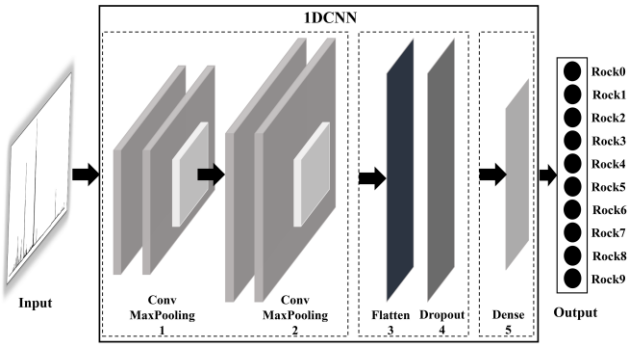

Fig. 6. The structure of one - dimensional convolutional neural network.

The model, 1DCNN uses 4 convolutional layers to extract features, according to the spectral data from the rocks investigated. Each convolutional layer adds an activation function to improve the ability of the neural network to express an effective classification model. After each of the 2 convolutional layers, a maximum pooling layer was added. To preserve the main features and reduce the computational cost, a flattening layer and a Dropout Layer were added in the middle, where the flattening layer converted the high-dimensional data into one-dimensional data and acts as a transition from the convolutional layer to the fully connected layer. Further, the Dropout Layer randomly assigns 50% of the neurons in the network to zeros. By resetting the weight to zero, the weight reduces the sensitivity of the convolutional neural network to any small changes experienced in the data and further improves the accuracy of the processing of unknown data. After that, a fully connected layer was then added and the softmax function was further added as the output layer, to obtain each feature, with data matching the feature category with the highest probability. The Cross Entropy Loss function was selected as the model loss function, where the formula used is as shown in Equation (9). The corresponding network structure of the 1D-CNN was as is shown in Fig. 6 above, where the detailed parameters of the network structure used are as shown in Table 2. The Softmax function formula employed was as follows.

$$Softmax\left(x_i\right)=\frac{\exp\left(x_i\right)}{\sum_{i=1}^{n}\exp\left(x_i\right)}\left(i=1,2,\ldots n\right) \tag{8}$$

In Equation (8), x_i represents the input of the Softmax function, and the n-dimensional real vector was mapped to the (0, 1) interval through the Softmax function. The sum of all probabilities is 1 and consequentially the probability distribution of multi-classification can be obtained.

$$\text{Loss}=-\sum_{i=1}^{n}y_i\cdot\log y_i \tag{9}$$

Table 2. The parameters of 1D - CNN

Number	Network layer	Parameter	Conv kernel	Step size
1	Input layer	145×1	-	-
	Conv-1	5×1	100	1
	Conv-2	5×1	100	1

	Max-Pool-1	3 ×1	-	1
	Conv-3	5×1	150	1
2	Conv-4	5×1	150	1
	Max-Pool-2	3 ×1	-	1
3	Flatten	-	-	-
4	Dropout	0.5	-	-
5	Dense	10 outputs	-	-

Grid Search Cross Validation (GS) is a basic hyperparameter optimization technique. Here, through defining an n-dimensional grid, each grid has a hyperparameter map, and the optimal hyperparameters of the model are obtained by exhausting all the hyperparameter combinations, through cross-validation. In this technique, assuming that there are n hyperparameters, and each hyperparameter has P_i values, then with N-fold cross-validation, the number of hyperparameters to be evaluated is $\mathrm{N}\times(\prod_{i=1}^{n}P_i)$. When this number of hyperparameters is large, although the optimal hyperparameter combination can be obtained in the end, this will lead to dimension explosion and low model efficiency. Randomized Search Cross Validation (RS) obtains the optimal hyperparameter combination in the sampling group by randomly selecting the hyperparameter combination in the hyperparameter space for a given number of iterations n_iter with cross-validation. The use of a random search solves the problems of a large number of grid search combinations, a large amount of computation, and a long time over which it is done, but only suboptimal solutions can be obtained. In this paper, the SVM classification model is used to optimize the super parameters, C and gamma through both grid search cross validation and random search cross-validation. First, the grid range of the optimal combination is determined through a grid search, following which an optimal solution is obtained by use of a random search. Therefore, the use of grid search cross-validation combined with random search cross-validation can be used to shorten the time taken and thus obtain better optimization results.

Bayesian Optimization (BO) regards the function of the optimized object as a random process that satisfies the prior distribution and allows obtaining a new distribution by updating the posterior probability with the Bayesian formula, by solving the function value. Following this, it then judges the most likely value, according to the new distribution. The extreme point and its function value can be calculated to form a new function value observation history. The next stage is to set the number of iterations, repeat the above process, and thus to obtain the optimal solution. It can be noted that the formula from Bayes' theorem can be given, as follows.

$$P(f|\ D_{1:t}) = \frac{P(D_{1:t}|\ f)P(f)}{P(D_{1:t})} \tag{10}$$

Where in the above formula f is the objective function; $D_{1:t} = \{(x_1, y_1), \ldots (x_t, y_t)\}$ represents the observed set, x_t represents the decision vector, $y_t = f(x_t) + \varepsilon_t$ represents the observation value, ε_t represents the observation error; $P(D_{1:t}|\ f)$ represents the likelihood distribution of y; $P(f)$ represents the prior probability distribution of f, which is used to describe the assumption made about the state of the unknown objective function; $P(D_{1:t})$ represents the posterior probability distribution of f which is used to describe the confidence of the unknown objective function after the prior probability is corrected by the measured data set.

Gaussian Process-Based Bayesian Optimization (GP-BO) uses Gaussian Processes (GP) to build a surrogate model to simulate the objective function for black-box optimization and in that way speed up the convergence. The specific idea behind its use is as follows: by obtaining the posterior probability of the observed point x, the mean u and variance σ of the $x$point can be calculated, following which the value of x is determined according to the extraction function u. After that, the process continues to sample the objective function and evaluate the objective function value, then integrate the data and update the Gaussian surrogate model for fast convergence. The formula of the function, u, is given as shown below in Equation (11).

$$u(x) = \underset{x \in D}{\operatorname{argmax}}\, \mu_{t-1}(x) + \beta_t^{\frac{1}{2}} \sigma_{t-1}(x) \tag{11}$$

In this paper, an improved Bayesian optimization algorithm (a Bayesian fast automatic hyperparameter optimization method based on simplex optimization domain) is proposed. In this approach, first, we construct the simplex optimization domain and sample the partition in an internal division. Then create a work queue through the heappop function of the heapq stack column module in Python; As shown in Fig. 7 below, the two division methods of simplex and square are respectively selected to optimize the domain partition, and on the assumption that each partition can be represented by the sampling points of its vertices, the partition efficiency can be defined as E. The formula for $E$is given as Equation (12):

$$E = \frac{En'}{En} \tag{12}$$

In the above formula, En' is the number of new partitions, and En is the number of sampling points required to divide the new area. It can be noted that when bisecting the square optimization domain, two sampling points are required; when bisecting from the inner third of the simplex or an edge, only one sampling point is required. The next step is to divide the interior of the simplex and the sides of the simplex and square respectively, and the three partition efficiencies E are $(n+1)/1$, $2/1$, $2/2^{n-1}$, This gives $E_a > E_b > E_c$. Therefore, it is seen as most efficient to divide the simplex optimization domain internally.

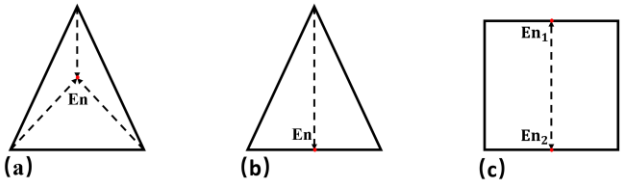

Fig. 7. Optimized domain partitioning efficiency comparison.

Following this, the hyperparameter space can be sampled according to Inverse Distance Weighting (IDW), and local interpolation is then performed in a simplex manner. The inverse distance weighted interpolation is a spatial interpolation method that can be used to estimate the location of the next point to be measured by using the measured sample points with corresponding weights. The inverse distance weighted interpolation formula used is shown below.

$$x^* = \frac{w_1 x_1 + w_2 x_2 + w_3 x_3 + \ldots + w_n x_n}{w_1 + w_2 + w_3 + \ldots + w_n} \tag{13}$$

Where in the above formula, x^* is the next sample point to be measured; $x_i \left(i = 1, 2, .., n\right)$ is the i-th measured sampling point; w is the weight. The formula for the weight of the i-th measured sampling point used is shown below.

$$w_i = \frac{1}{d_{ix^*}^{p}} \tag{14}$$

Here, the weight w is inversely proportional to the p power of the distance d (from the point to be measured x^* to the known sampling point x_i), where the Euclidean Metric is selected to calculate the distance d, and the distance formula is as given in Equation (15) shown. As a result, the corresponding weighting decreases as the distance increases and the rate of descent depends on the pre-set constant p. When p =0, the relative weight is 1, and the prediction point is the average value of the measured data in the search field; and when p is too large, the weight decreases rapidly. It is only the sampling points of the nearest points which have an effect on the predicted points. Therefore, the default value is p =2, and the interpolation method is the inverse distance square weight interpolation.

$$d = \sqrt{\left(x_2 - x_1\right)^2 + \left(y_2 - y_1\right)^2} \tag{15}$$

In the above formula, d is the Euclidean distance between the point $\left(x_2, y_2\right)$ and the point $\left(x_1, y_1\right)$.

Next, the number of iterations is set and the local interpolation value which gives the highest information gain is selected from the priority queue of the points to be measured through each iteration, and the acquisition function is used to prevent the acquisition of a sub-optimal local interpolation value, to ensure the faster convergence of the objective function. The acquisition function used is shown below.

$$f\left(x^*\right) = x^* + C*T \tag{16}$$

In the above formula, C is the weight of the optimization domain exploration, which is used to inform the optimizer of how much attention should be paid to the current exploration optimization domain to prevent the pursuit of the first Sub-optimal value explored by the optimizer. C is an adjustable constant, and the default is C =0.1, where T is the exploration cost. The formula used for the value of T is as shown below.

$$T = \left(X_1 - X_2\right) * \log_{EN}\left(F_1 * F_2\right) \tag{17}$$

In the above formula, X_1 and X_2 are the best sampling point and the worst sampling point in the simplex sampling domain respectively, EN is the number of all the vertices in the simplex optimization domain, and F_1、F_2 are the scores of the simplex parent domain and subdomain, respectively. Through local IDW interpolation, the simplex parent domain can be divided into several simplex subdomains, and the content of each simplex subdomain then comes from the parent domain, so the exploration cost is the cost of each simplex space domain containing the collection points to be measured. A relative measure of how much has been explored in this work.

The specific process employing the improved Bayesian optimization algorithm is as follows: the classification accuracy of the classification model is used as the objective function of the optimizer, constructing a simplex optimization domain from the hyperparameter space and creating a simplex work queue. Then, the surrogate model of the optimizer is constructed

by dividing the simplex optimization domain into independent local interpolations, according to IDW. The local interpolation with the highest acquisition function value is then obtained from the simplex work queue, and the objective function can be evaluated by this local interpolation. As the exploration information is updated, the parent interpolation is split into smaller and more accurate child interpolations, and then added to the simplex worklist. With a given number of iterations, when the model training meets the number of iterations or the optimal local interpolation gets the optimal hyperparameter combination.

3. Results and analysis

Building on the above, the average spectrum of 10 types of rock samples was obtained and then the characteristic wavelengths of each element were determined according to the NIST atomic spectrum database. It is found through experiments that the elements affecting lithology identification include Nb, Si, Al, Mg, Ca, Ti, Na, Ba, H, Li, Mn, Fe, etc. It is not easy to quickly select the analysis line as the identification and classification model under the principle of characteristic spectral line screening (which involves high spectral line intensity, high element transition probability, and no interference and overlap of other element spectral lines around the spectral line). Typical LIBS spectra of rock samples and emission spectra of main elements in the LIBS spectra are shown in Fig. S1 and Table S1 respectively.

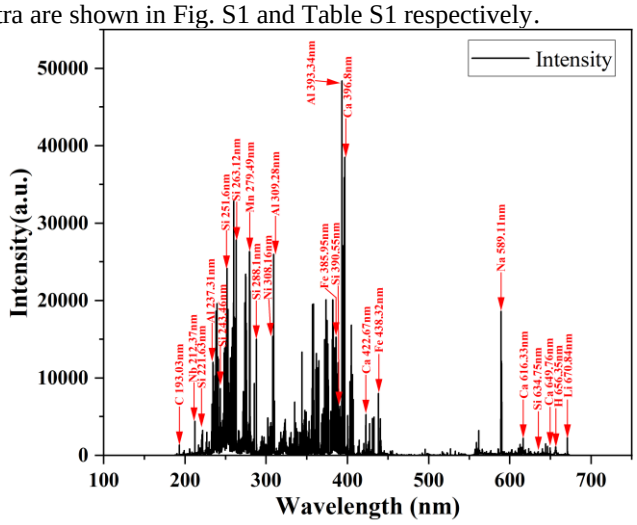

Fig. S1. Typical LIBS spectrum of rock samples.

Table S1. Emission lines of main elements in the LIBS spectrum

Species	Wavelength (nm)	Species	Wavelength (nm)
Si	221.63, 243.46, 251.6, 263.12, 288.1, 385.56, 390.55, 634.75	Al	237.31, 309.28, 393.34, 394.39, 396.09
Fe	385.95, 438.32	Ca	396.8, 422.67, 616.33, 649.76
Na	261.18, 288.1, 589.11,	Ni	308.16
Mn	279.49	Nb	212.37, 260.06
H	656.35	Li	256.23, 670.84
Mg	279.55, 350.07	Ba	273,93, 614.25
O	240.62, 373.68, 404.6	Cr	275.52, 383.42

C	193.03	Ti	384.06, 455.34

3.1 Principal Component Analysis

In this experiment, the PCA has been used to extract features from the spectral data, and the pre-processed spectral data are used as the input of PCA for dimensionality reduction processing, to achieve rapid convergence of the classification model, on the premise of retaining most of the original spectral information. Fig. 8 shows the interpretation rate and cumulative interpretation rate of the first 20 principal components, and Fig. 9 shows the two-dimensional scatter plot of the 10 types of rocks when the first two principal components were selected.

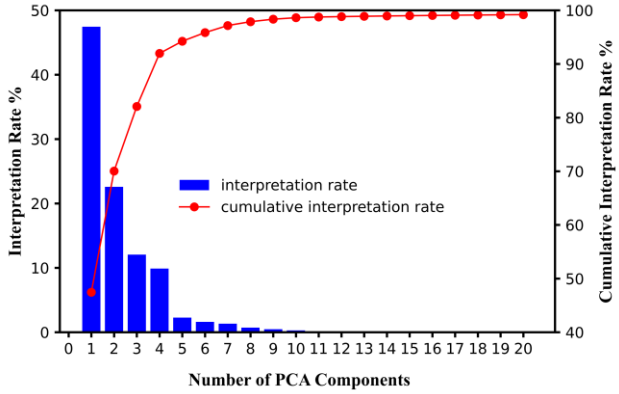

Fig. 8. PCA analysis results of LIBS data for 10 types of rocks.

By observing Fig. 8 above, it can be seen that the cumulative interpretation rate of the first four principal components has the fastest growth rate, and the cumulative interpretation rate reaches 91.96%; when the first ten principal components were taken, where the cumulative interpretation rate reaches 98.62%, and the cumulative interpretation rate of the principal components rate growth is extremely slow.

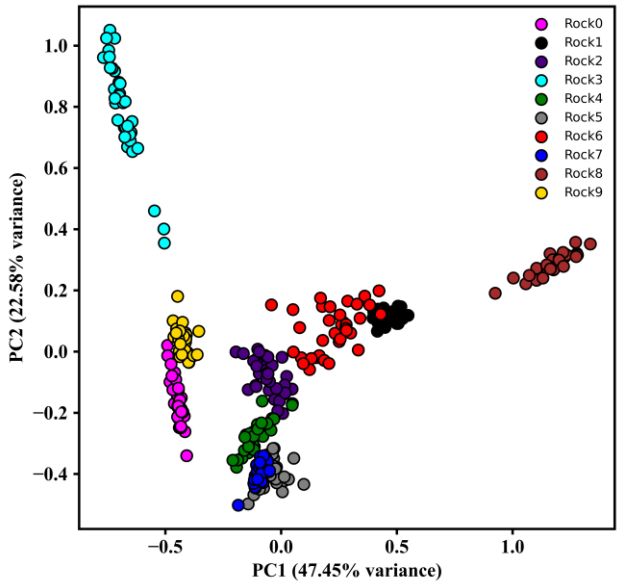

Fig.9. Two-dimensional scatter plot of the first two principal components.

By observing Fig. 9 above, it can be seen that when the first two principal components were selected as features after dimensionality reduction by using the PCA, it is found that ten types of rock data points appear to show a regional clustering phenomenon, among which Rock2 and Rock4, Rock1 and Rock6 have very little information which is overlapping in nature, but most of the information on Rock5 and Rock7 overlap, where most of the sample intervals are small. This is because the cumulative explanation rate of the first two principal components is only 70.03%, and it is necessary to increase the number of principal components and use the classification algorithm to further analyse the data after feature extraction has taken place.

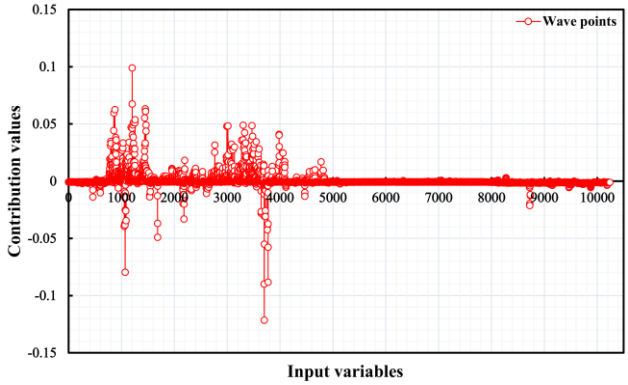

Fig. 10. The loadings of PC1.

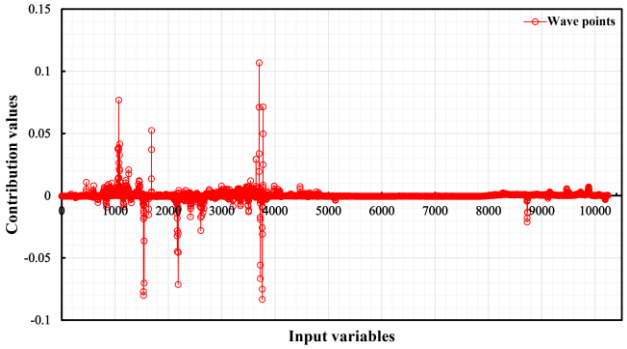

Fig. 11. The loadings of PC2.

Table 3. Contributions of the elements to the PC

Contribution values	Si	Fe	Al	Ca	Nb	Mn	Na	C	Li	Others
PC1	0.183	0.264	0.045	0.152	0.068	0.081	0.049	0.098	0.015	0.045
PC2	0.151	0.166	0.019	0.122	0.015	0.025	0.111	0.007	0.036	0.348

By calculating the loads of PC1 and PC2, the contribution values of the input variables (10,239 wave points) to the first two principal components were obtained. Following that and with reference to the NIST atomic spectrum database, the contribution value of the main elements in the rock to the PC was obtained, where the loads on PC1 and PC2 are, respectively, as shown in Fig. 10 and Fig. 11 above, and the contribution of elements to the PC is as shown in Table 3. It can be seen from Table 3 that Si, Fe, Al, Ca, Nb, Mn, Na, C and Li are the key elements of the first two principal components. Among them, Si, Fe, Al, Ca, Nb, Mn and C provide the major contributions to PC1. When the number of principal components is increased, the contribution value of Na, and Li and other elements to PC2 increases, and this process adds

new important features. Therefore, the appropriate number of principal components should be selected as the input to the classification model.

3.2 Classifier Model Optimization

The LIBS spectral data of 10 types of rocks can be seen as basic pre-processing operations and PCA dimensionality reduction with different numbers of principal components is used as the input of the classification model – and there are 500 sets of spectral data used in total. Among them, there are 350 groups in the training set and 150 groups in the test set, and the proportion of each type of rock in both the training set and the test set is equal. The classification accuracy R_{ACC} was used as the evaluation index of the classification model, the model was evaluated with ten-fold cross-validation, and the confusion matrix was selected for the final evaluation of the classification effect of the test set. The classification accuracy formula is given by:

$$R_{\mathrm{ACC}} = \sum_{i=1}^{10} T_i \Big/ \left(\sum_{i=1}^{10} T_i + \sum_{i=1}^{10} F_i \right) \tag{18}$$

In the above formula， T_i is the total number of correct classifications of class i rocks, F_i is the total number of incorrect classifications of class i rocks, $\sum_{i=1}^{10} T_i$ represents the total number of correct classifications of 10 types of rocks, and $\sum_{i=1}^{10} F_i$ represents the total number of incorrect classifications of 10 types of rock samples.

The SVM in this experiment can be implemented, based on the sklearn library in Python 3.9.7. The classification accuracy was used as the objective function of GS and RS, and the SVM model was optimized by adjusting the hyperparameter space and the number of iterations. Table 4 shows the 10-fold average classification and recognition accuracy of the sub-test set when different numbers of principal components were selected as the input variables of the SVM classification model, and the training set was subjected to ten-fold cross-validation as the sub-training set and the sub-test set, respectively.

Table 4. SVM recognition results with different input variables

Numbers of variables	Cumulative interpretation rate /%	Average accuracy rate /%
1	47.45%	55.14%
2	70.03%	74.57%
3	82.08%	79.14%
5	94.23%	83.31%
10	98.61%	87.00%
15	98.99%	86.20%

It can be seen from Table 4 above that with the increase of the principal component score, the cumulative interpretation rate is increasing, and the average classification accuracy of the SVM classification model is also improving. When ten principal components were selected as the input of the classification model, the average classification accuracy of the SVM reaches 87.00%. When the principal component scores continue to increase, the average classification accuracy of the classification model was seen to decrease. This arose because with the increase of principal component fraction, the increase of noise and dimension will reduce the classification accuracy.

Through the comparative analysis in Table 4, the top 10 principal components are selected as the inputs of GSCV-RSCV-SVM and GP-BO-SVM models. Firstly, the range of penalty factor C in SVM is [0.01, 0.1, 1, 10, 100, 1000, 10000], and the range of parameter gamma was [0.001, 0.01, 0.1, 1, 10, 100]. Through grid search cross validation, the best parameter combination C=1000, and gamma=0.1 is obtained. At this time, the average accuracy of cross

validation of the classification model is 87.20%. Then set 10<C<1000, 0.01<gamma<1, were set, and through 30 iterations of random search cross validation, the optimal parameters of SVM were found to be C=743, gamma=0.38. At this time, the average accuracy of the cross validation set of the classification model was 87.71%.

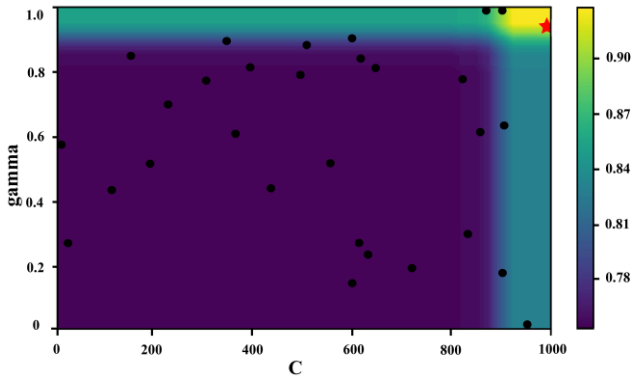

Fig. 12. BO-GP-SVM optimal parameters optimization process.

To further improve the classification efficiency of SVM, Bayesian optimization based on the Gaussian process was used to optimize the hyperparameters of SVM. Following that, it was necessary to establish the hyperparameter C and gamma two-dimensional hyperparameter space. As shown in Fig. 12 above, the best parameter combinations C=990.101 and gamma=0.919 were obtained after 30 iterations. At this time, the average accuracy of cross validation of the classification model was 90.20%. Although BOGP-SVM has improved the classification accuracy compared with GS-RS-SVM, the modelling times of the two models were 300.611s and 298.023s respectively, so the model efficiency of the two methods was not high.

Next, the Bayesian optimization based on the Gaussian process was used to optimize the hyperparameters of 1DCNN, and thus construct a four-dimensional parameter space for the convolutional layer activation function, optimizer, batch size, and number of model training rounds through the Bayesian optimizer. The process searches within the parameter space serve as a surrogate model. The blue area in the figure is the area with a 'poor model' effect, the yellow area is the 'better area', and the black point is the sampling position of the Bayesian optimizer. The Bayesian optimizer fitted the model by using the prior probability at the higher sampling density. parameters and the red five-pointed star was obtained as the best hyperparameter combination position. The optimal hyperparameter combination (pre-processing) was as follows: the activation function of the convolutional layer was Relu, the optimizer was Adam, the batch size was 23, and the number of model training rounds was 35. At this time, the average cross-validation accuracy of the classification model was 98%. The hyperparameter optimization process is shown in Fig. 13 below.

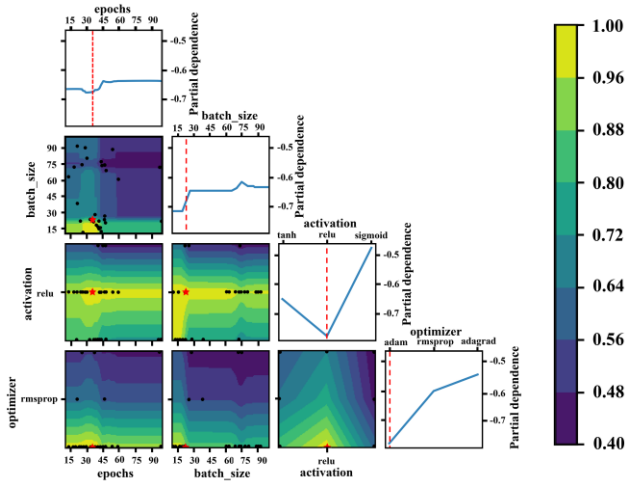

Fig. 13. BO-GP-1DCNN optimal parameters optimization process.

As shown in Fig. 14 below, a Bayesian optimizer was constructed, based on a simplex optimization domain to determine a two-dimensional hyperparameter space for the batch size and the number of model training epochs in the 1DCNN classification model, where the Batch_size $\in$ [1,100], Epoch $\in$ [1,100]. Then the simplex optimization domain was constructed according to the hyperparameter space, and the optimal hyperparameter combination was obtained through 15 iterations. Here the batch size was 17, and the number of model training rounds was 69. At this time, the average cross-validation accuracy of the classification model was 99.33%, and the modelling time was 103s. The hyperparameter optimization process is as shown in Fig. 15 below.

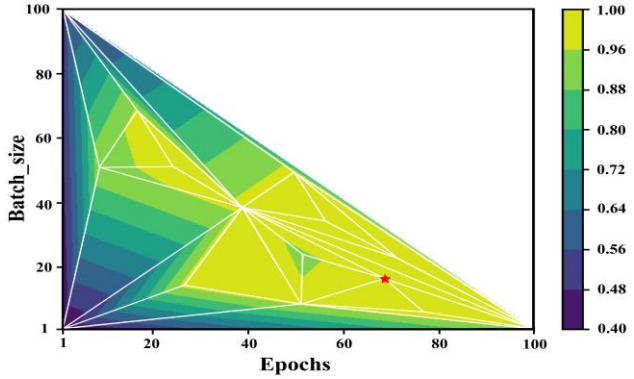

Fig. 14. Improved BO-1DCNN optimal parameters optimization process.

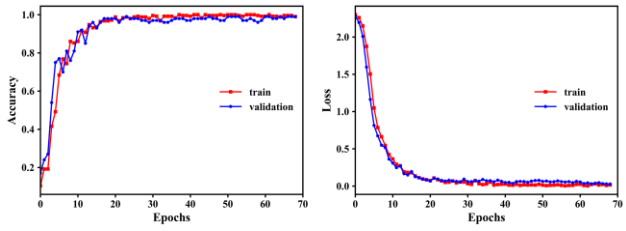

Fig. 15. Improved BO-1DCNN model training process.

By examining the training process of the Improved BO-1DCNN model in Fig. S2, it can be seen that when the network was trained through ~20 rounds, the cross-entropy loss

difference of the network tends to be stable, and the network accuracy at this time was 0.001. Although the actual set network error value was met at this time, the training set and the classification accuracies of the validation set were 97.2% and 98%, respectively, falling into a local optimum. The global search was carried out through the improved Bayesian optimizer. When the network was trained to 69 rounds, the network accuracy is 0.00098, which is the global optimal value. The confusion matrix for the classification and recognition of 10 types of rocks in the test set after model optimization is shown in Fig. S2. It can be seen that the Improved BO-1DCNN model can recognize and classify 10 types of rocks. The accuracy rate can reach 99.33%, and the model recognition time was significantly reduced, compared to the first three models, which can be used to quickly identify and classify 10 types of rock samples.

3.3 Comparative and analysis

To further verify the ability of the Improved BO-1DCNN model to automatically classify rocks, the four classification models are compared with the basic pre-processing and optimally pre-processing rock LIBS spectral data. Fig. 16 below shows the validation set accuracy and modelling time of the four methods obtained. Table 5 is the validation set efficiency comparison of the four models. (1) in Table 5 is pre-processing, and (2) is no pre-processing. It can be seen that SVM has lower classification accuracy than 1DCNN. This is because SVM performs classification and recognition by mapping the input to a high-dimensional space. Moreover, when the SVM input variable dimension is too high, the efficiency of classification recognition and hyperparameter optimization was low. The 1DCNN adopted a convolution calculation and deep architecture and achieves high-accuracy classification by extracting spectral features to form high-level semantic category information.

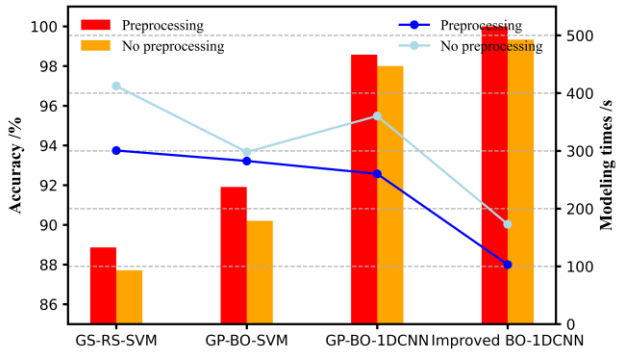

Fig. 16. The validation set accuracy and modelling time of the four methods.

Table 5. Comparison of the validation set efficiency of the four models

Models	Accuracy /%	Modelling times /s
GS-RS-SVM (1)	88.86	300.611
GS-RS-SVM (2)	87.71	412.391
GP-BO-SVM (1)	91.91	282.493
GP-BO-SVM (2)	90.20	298.023
GP-BO-1DCNN (1)	98.57	260.145
GP-BO-1DCNN (2)	98.00	360.381
Improved BO-1DCNN (1)	99.46	103.858
Improved BO-1DCNN (2)	99.33	173.299

By comparing the classification accuracy and modelling time of different classification models, it can be seen that the optimization efficiency of the three optimizers follows the following pattern: GS-RS<GP-BO<Improved BO. By comparing the basic pre-processing experiment and the optimal pre-processing experiment, it has been found that although the classification accuracy of the optimal pre-processing experiment was higher, the modelling time also increased, and a large number of experiments were found necessary to obtain the optimal pre-processing combination. Fig. 17 below shows the test set accuracy and modelling time of the four methods. Table 6 is the test set efficiency comparison of the four models. Therefore, compared with other methods, the Improved BO-1DCNN model proposed in this study not only needed no optimal pre-processing, but also has the highest classification efficiency, used to realize LIBS for the automatic and rapid classification of rock samples.

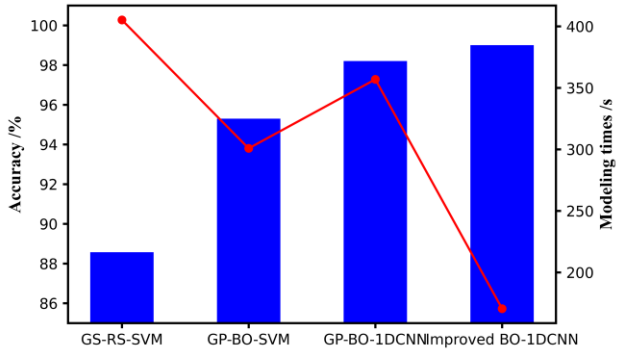

Fig. 17. The test set accuracy and modelling time of the four methods.

Table 6. Comparison of the test set efficiency of the four models

Models	Accuracy /%	Modelling times /s
GS-RS-SVM	88.57	405.231
GP-BO-SVM	95.30	300.824
GP-BO-1DCNN	98.20	356.912
Improved BO-1DCNN	99.00	170.521

4. Conclusions

In this paper, an improved Bayesian optimization algorithm has been put forward and proposed – then it has been applied to create an identification method combining laser-induced breakdown spectroscopy with a one-dimensional convolution neural network. It has been used and evaluated for the classification of 10 different types of rocks, such as siltstone, oil shale, mudstone, argillaceous siltstone, red sandstone, coarse sandstone, fine sandstone, etc., to solve the shortcomings of traditional Bayesian algorithms, such as global dynamic optimization and low efficiency. By selecting different classifiers and optimizers to build classification models, including GS-RS-SVM、GP-BO-SVM、GP-BO-1DCNN and improved BO-1DCNN, comparative experiments were carried out on 10 types of rock spectral data that were pre-processed and unprocessed respectively. The experimental results obtained from that work were then able to show that LIBS technology, combined with the 1DCNN classification model, does not need to select the optimal pre-processing combination, and the improved Bayesian optimization algorithm can achieve high accuracy and fast classification between rocks. Through global dynamic optimization, the modelling time of the improved Bo algorithm has been seen to be greatly reduced, and the classification accuracy obtained was higher after optimization, which improves the classification efficiency of rocks.

5. Back matter

5.1 Funding

This research is financially supported by the "20 New Universities" project of Jinan City (2021GXRC037).

5.2 Acknowledgments

Thanks to the "20 New Universities" project of Jinan City for helping identify collaborators on this work.

5.3 Disclosures

The authors declare no conflicts of interest.

5.4 Data availability statement

Data availability. Data underlying the results presented in this paper are not publicly available at this time but may be obtained from the authors upon reasonable request.

References

1. Y. J. J. o. C. P. Zhang, "Analysis of China's energy efficiency and influencing factors under carbon peaking and carbon neutrality goals," **370**, 133604 (2022).
2. G. F. Wang, H. W. Ren, G. R. Zhao, D. S. Zhang, Z. G. Wen, L. Y. Meng, and S. X. Gong, "Research and practice of intelligent coal mine technology systems in China," Int. J. Coal Sci. Technol. **9**, 17 (2022).
3. T. Koralay and Y. K. Kadioglu, "Reasons of different colors in the ignimbrite lithology: Micro-XRF and confocal Raman spectrometry method," Spectroc. Acta Pt. A-Molec. Biomolec. Spectr. **69**, 947-955 (2008).
4. X. Z. Zhao, X. G. Pu, W. Z. Han, L. H. Zhou, Z. N. Shi, S. Y. Chen, and D. Q. Xiao, "A new method for lithology identification of fine grained deposits and reservoir sweet spot analysis: A case study of Kong 2 Member in Cangdong sag, Bohai Bay Basin, China," Petroleum Explor. Dev. **44**, 524-534 (2017).
5. R. Poormirzaee, S. Hosseini, and R. Taghizadeh, "Smart mining policy: Integrating fuzzy-VIKOR technique and the Z-number concept to implement industry 4.0 strategies in mining engineering," Resour. Policy **77**, 14 (2022).
6. J. J. Yan, P. Yang, R. Zhou, S. H. Li, K. Liu, W. Zhang, X. Y. Li, D. Z. Wang, X. Y. Zeng, and Y. F. Lu, "Classification accuracy improvement by data preprocessing in handheld laser-induced breakdown spectroscopy," Anal. Methods **11**, 5177-5184 (2019).
7. S. Muller and J. A. Meima, "Mineral classification of lithium-bearing pegmatites based on laser-induced breakdown spectroscopy: Application of semi-supervised learning to detect known minerals and unknown material," Spectroc. Acta Pt. B-Atom. Spectr. **189**, 15 (2022).
8. J. X. Chen, J. Pisonero, S. Chen, X. Wang, Q. W. Fan, and Y. X. Duan, "Convolutional neural network as a novel classification approach for laser-induced breakdown spectroscopy applications in lithological recognition," Spectroc. Acta Pt. B-Atom. Spectr. **166**, 7 (2020).
9. Y. Zhang, T. L. Zhang, and H. Li, "Application of laser-induced breakdown spectroscopy (LIBS) in environmental monitoring," Spectroc. Acta Pt. B-Atom. Spectr. **181**, 17 (2021).
10. P. Sun, X. J. Hao, Y. W. Yang, Y. K. Liu, W. Y. Hao, and Y. Tian, "Effect of Baseline Correction on the Quantitative Analysis of Soil LIBS," in *Conference on Plasmonics VI*, Proceedings of SPIE (Spie-Int Soc Optical Engineering, 2021),
11. J. Ren, Y. R. Zhao, and K. Q. Yu, "LIBS in agriculture: A review focusing on revealing nutritional and toxic elements in soil, water, and crops," Comput. Electron. Agric. **197**, 16 (2022).
12. X. Y. Song, K. H. Li, K. J. Dai, X. Q. Wang, H. J. Du, and H. L. Zhao, "A random-forest-assisted artificial-neural-network method for analysis of steel using laser-induced breakdown spectroscopy," Optik **249**, 9 (2022).
13. A. K. Myakalwar, C. Sandoval, M. Velasquez, D. Sbarbaro, B. Sepulveda, and J. Yanez, "LIBS as a Spectral Sensor for Monitoring Metallic Molten Phase in Metallurgical Applications-A Review," Minerals **11**, 22 (2021).
14. N. Ahmed, J. A. Awan, K. Fatima, S. M. Z. Iqbal, M. Rafique, S. A. Abbasi, and M. A. Baig, "Machine learning-based calibration LIBS analysis of aluminium-based alloys," Eur. Phys. J. Plus **137**, 14 (2022).
15. C. W. Zhu, J. X. Lv, K. Liu, J. Chen, K. Liu, G. Q. Li, B. Lu, and X. Y. Li, "Rapid Determination of Arsenic in Traditional Chinese Medicine by Laser-Induced Breakdown Spectroscopy (LIBS)," Anal. Lett., 11.
16. D. M. Wayua, H. K. Angeyo, A. Dehayem-Kamadjeu, and K. A. Kaduki, "Direct Analysis of Blood for Diagnostic Metals for Malaria by Peak-Free Laser-Induced Breakdown Spectroscopy (LIBS) with Artificial Neural Networks (ANN) and Partial Least Squares (PLS)," Anal. Lett., 14.

17. Z. Gazali, R. Kumar, P. K. Rai, P. K. Rai, A. K. Rai, and S. N. Thakur, "Discrimination of gallbladder stone employing Laser-Induced Breakdown Spectroscopy (LIBS) and Photoacoustic Spectroscopy (PAS)," Spectroc. Acta Pt. A-Molec. Biomolec. Spectr. **260**, 11 (2021).
18. I. Cherni, M. Nakkach, H. Ghalila, R. Nouir, M. Somai, F. Daoued, I. Rachdi, F. Boussema, N. Jaidane, and S. Hamzaoui, "Noninvasive diagnosis of type 2 diabetes mellitus by hair analysis using laser-induced breakdown spectroscopy (LIBS)," Instrum. Sci. Technol., 16.
19. D. Stefas, N. Gyftokostas, P. Kourelias, E. Nanou, C. Tananaki, D. Kanelis, V. Liolios, V. Kokkinos, C. Bouras, and S. Couris, "Honey discrimination based on the bee feeding by Laser Induced Breakdown Spectroscopy," Food Control **134**, 9 (2022).
20. B. Sezer, A. Unuvar, I. H. Boyaci, and H. Koksel, "Rapid discrimination of authenticity in wheat flour and pasta samples using LIBS," J. Cereal Sci. **104**, 7 (2022).
21. V. Detalle and X. S. Bai, "The assets of laser-induced breakdown spectroscopy (LIBS) for the future of heritage science," Spectroc. Acta Pt. B-Atom. Spectr. **191**, 23 (2022).
22. F. Yang, L. N. Li, W. M. Xu, X. F. Liu, Z. C. Cui, L. C. Jia, Y. Liu, J. H. Xu, Y. W. Chen, X. S. Xu, J. Y. Wang, H. Qi, and R. Shu, "Laser-induced breakdown spectroscopy combined with a convolutional neural network: A promising methodology for geochemical sample identification in Tianwen-1 Mars mission," Spectroc. Acta Pt. B-Atom. Spectr. **192**, 15 (2022).
23. C. Q. Liu, Z. C. Ling, J. Zhang, Z. C. Wu, H. C. Bai, and Y. H. Liu, "A Stand-Off Laser-Induced Breakdown Spectroscopy (LIBS) System Applicable for Martian Rocks Studies," Remote Sens. **13**, 15 (2021).
24. J. L. Yu, Z. Y. Hou, S. Sheta, J. Dong, W. Han, T. J. Lu, and Z. Wang, "Provenance classification of nephrite jades using multivariate LIBS: a comparative study," Anal. Methods **10**, 281-289 (2018).
25. W. J. Xu, C. Sun, Y. Q. Zhang, Z. Q. Yue, S. Shabbir, L. Zou, F. Y. Chen, L. Wang, and J. Yu, "Accurate determination of structural H2O in rocks using LIBS coupled with machine learning algorithms extensively exploring the characteristics of the H-alpha line," J. Anal. At. Spectrom. **37**, 317-329 (2022).
26. X. Wang, S. Chen, M. F. Wu, R. Q. Zheng, Z. Liu, Z. J. Zhao, and Y. X. Duan, "Low-cost smartphone-based LIBS combined with deep learning image processing for accurate lithology recognition," Chem. Commun. **57**, 7156-7159 (2021).
27. J. A. Meima, D. Rammlmair, and M. Junge, "The use of Laser Induced Breakdown Spectroscopy for the mineral chemistry of chromite, orthopyroxene and plagioclase from Merensky Reef and UG-2 chromitite, Bushveld Complex, South Africa," Chem. Geol. **589**, 15 (2022).
28. P. Janovszky, K. Jancsek, D. J. Palasti, J. Kopniczky, B. Hopp, T. M. Toth, and G. Galbacs, "Classification of minerals and the assessment of lithium and beryllium content in granitoid rocks by laser-induced breakdown spectroscopy," J. Anal. At. Spectrom. **36**, 813-823 (2021).
29. P. Jahoda, I. Drozdovskiy, S. J. Payler, L. Turchi, L. Bessone, and F. Sauro, "Machine learning for recognizing minerals from multispectral data," Analyst **146**, 184-195 (2021).
30. J. H. Hu, T. Zhou, S. W. Ma, D. J. Yang, M. M. Guo, and P. L. Huang, "Rock mass classification prediction model using heuristic algorithms and support vector machines: a case study of Chambishi copper mine," Sci Rep **12**, 20 (2022).
31. T. Chen, L. X. Sun, H. B. Yu, W. Wang, L. F. Qi, P. Zhang, and P. Zeng, "Deep learning with laser-induced breakdown spectroscopy (LIBS) for the classification of rocks based on elemental imaging," Appl. Geochem. **136**, 10 (2022).
32. G. Alix, E. Lymer, G. L. Zhang, M. Daly, and X. Gao, "A comparative performance of machine learning algorithms on laser-induced breakdown spectroscopy data of minerals," J. Chemometr., 16.
33. C. Wang, J. Wang, H. Du, and J. J. L. P. Wang, "Classification of 13 original rock samples by laser induced breakdown spectroscopy," **31**, 035601 (2021).
34. R. El-Saeid, Z. Abdel-Salam, S. Pagnotta, V. Palleschi, and M. J. S. A. P. B. A. S. Harith, "Classification of sedimentary and igneous rocks by laser induced breakdown spectroscopy and nanoparticle-enhanced laser induced breakdown spectroscopy combined with principal component analysis and graph theory," **158**, 105622 (2019).
35. M. Yelameli, B. Thornton, T. Takahashi, T. Weerakoon, and K. J. J. o. C. Ishii, "Classification and statistical analysis of hydrothermal seafloor rocks measured underwater using laser‐induced breakdown spectroscopy," **33**, e3092 (2019).
36. P. Janovszky, K. Jancsek, D. J. Palásti, J. Kopniczky, B. Hopp, T. M. Tóth, and G. J. J. o. A. A. S. Galbács, "Classification of minerals and the assessment of lithium and beryllium content in granitoid rocks by laser-induced breakdown spectroscopy," **36**, 813-823 (2021).
37. Y. Wu, L. Ozdamar, A. J. J. o. c. Kumar, and a. mathematics, "TRIOPT: a triangulation-based partitioning algorithm for global optimization," **177**, 35-53 (2005).