International Journal of Mining Science and Technology 32 (2022) 309–322

Contents lists available at ScienceDirect

International Journal of Mining Science and Technology

journal homepage: www.elsevier.com/locate/ijmst

# Machine learning-based classification of rock discontinuity trace: SMOTE oversampling integrated with GBT ensemble learning

Jiayao Chen [a,b], Hongwei Huang [a,*], Anthony G. Cohn [b,c,d,e,f], Dongming Zhang [a], Mingliang Zhou [a]

[a] Key Laboratory of Geotechnical and Underground Engineering of Ministry of Education, Department of Geotechnical Engineering, Tongji University, Shanghai 200092, China
[b] School of Computing, University of Leeds, LS2 9JT Leeds, United Kingdom
[c] Department of Computer Science and Technology, Tongji University, Shanghai 211985, China
[d] School of Civil Engineering, Shandong University, Jinan 250061, China
[e] Luzhong Institute of Safety, Environmental Protection Engineering and Materials, Qingdao University of Science and Technology, Zibo 255000, China
[f] School of Mechanical and Electrical Engineering, Qingdao University of Science and Technology, Qingdao 260061, China

ARTICLE INFO

*Article history:*
Received 10 March 2021
Received in revised form 19 July 2021
Accepted 25 August 2021
Available online 13 September 2021

*Keywords:*
Tunnel face
Rock discontinuity trace
Machine learning
Gradient boosting tree
Generalization ability

ABSTRACT

This paper presents a hybrid ensemble classifier combined synthetic minority oversampling technique (SMOTE), random search (RS) hyper-parameters optimization algorithm and gradient boosting tree (GBT) to achieve efficient and accurate rock trace identification. A thirteen-dimensional database consisting of basic, vector, and discontinuity features is established from image samples. All data points are classified as either "trace" or "non-trace" to divide the ultimate results into candidate trace samples. It is found that the SMOTE technology can effectively improve classification performance by recommending an optimized imbalance ratio of 1:5 to 1:4. Then, sixteen classifiers generated from four basic machine learning (ML) models are applied for performance comparison. The results reveal that the proposed RS-SMOTE-GBT classifier outperforms the other fifteen hybrid ML algorithms for both trace and non-trace classifications. Finally, discussions on feature importance, generalization ability and classification error are conducted for the proposed classifier. The experimental results indicate that more critical features affecting the trace classification are primarily from the discontinuity features. Besides, cleaning up the sedimentary pumice and reducing the area of fractured rock contribute to improving the overall classification performance. The proposed method provides a new alternative approach for the identification of 3D rock trace.



## 1. Introduction

Discontinuity trace, generated by the intersections of various rock surfaces and discontinuities, is a fundamental yet significant parameter for characterizing the strength of rock masses [1–3]. Discontinuities as major rock mass features are most widely used in rock mass classification systems, such as rock mass rating (RMR) [4], rock mass index (RMi) [5] and blastability index (BI) [6,7] for various applications such as failure criteria of rock mass [8,9], rock fragmentation by blasting [10,11], stability analysis of rock structures [12,13], permeability [14–16], etc. To date, the rock trace mapping manners are broadly classified into two categories, namely, contact and non-contact methods [17,18]. The contact manner, usually carried out with hand-held devices such as geological compass, roughness profiler and measuring tape, is a direct and common approach but is dangerous, non-reproducible, time-consuming and labour-intensive, meanwhile hard to employ in less accessible regions [19]. The emerging non-contact manners, such as photogrammetry, and LiDAR (Light Detection and Ranging), etc., can afford alternative methods for mapping field rock trace geometry by using the considerable resolution 2D digital images or 3D point clouds [7,20]. To this point, they have been considered to promote the efficiency and security of data collection and ensure the objectivity of the subsequent rock trace identification.

The common image processing approaches have been applied for extracting various information about the discontinuity geometry by capturing the changes in grid-, pixel-, and colour-level features [21,22]. The representative algorithms mainly include edge detection, threshold segmentation, and image enhancement, etc., [23,24]. Nevertheless, such approaches have brought inevitable obstacles for engineers in practice since each approach is proposed for a specific application scene [15,25]. Once the scene alters, the previously applied approach may no longer exhibit optimal perfor-

* Corresponding author.
*E-mail address:* huanghw@tongji.edu.cn (H. Huang).

https://doi.org/10.1016/j.ijmst.2021.08.004
2095-2686/

mance. Given aforementioned limitations, novel image analysis techniques/software have been thus used for assessing rock mass block size and improving the rock fragmentation model [11,26]. In particular, deep learning-based approaches are brought considerable concerns in many studies because they can extract the target features at a semantic level by using the special structures of convolutional layers and pooling layers [27,28]. Though high-level segmentation results have been gradually realized, some irreparable shortcomings for 2D image samples are still existing, e.g., it may be problematic for reflecting global information due to the limited image size [12]. Serious statistical error maybe brought by single angle shooting, even the information of the same scene counted from various angles are also very different [29]. Undoubtedly, mastering 3D coordinate information is essential for comprehensively carrying out the core geological characterizations.

Nowadays, numerous scientists have performed rock trace mapping using the 3D point clouds of exposed rock masses [30–33]. It should be noted that the corresponding 3D coordinate information of rock masses can be acquired using either digital photogrammetry or LIDAR approaches. LIDAR is obviously more accurate and accessible than digital photogrammetry in 3D data collection [33–35]. Nevertheless, LIDAR-based instruments are costly and difficult to cover core viewing directions due to the irregularity of rock bulge in inhospitable environments, especially for the underground scenes (e.g., tunnel sections, and mines). Furthermore, the use of LIDAR in some smaller projects is often limited by project resources, even operation level and domain knowledge of engineering personnel [31]. Additionally, the approach of digital photogrammetry is generally more flexible since users can freely select the required viewing directions and the number of shots until all the required details are captured [36]. Moreover, the colour information (R, G, B) can be naturally obtained from 3D reconstruction using image patterns instead of being realized by tediously fitting additional cameras in LADAR [37]. Considering the economic cost and colour requirement, the digital photogrammetry approach is chosen for data acquisition in this study though it also has some shortcomings, such as insufficient accuracy, and slight degradation due to dust and uneven illumination [38].

To date, the approaches to extract rock trace from point cloud data can be mainly summarized into three types:

(1) Define the rock trace as the intersection lines between different rock discontinuities [37,39]. Then, some representative clustering or region growing algorithms are presented to extract the corresponding discontinuities to capture the target intersections [40,41].
(2) Detect the rock traces from the reconstructed digital rock surface by searching the vertices which constitute the real trace line [32,42]. The vertices are usually judged by the principal curvatures.
(3) Identify the rock trace by human–computer interactions [43]. The representative software used is CloudCompare [44].

The above three approaches have attracted considerable research attention and been certified ideal performances for corresponding study cases. However, the first approach is proven more competent for extracting the significant rock trace rather than the fractured rock masses [45]. The second approach is undoubtedly time consuming due to the enormous computations involved in determining the initial vertices and then traversing the trace points [37,46]. The third approach has the most significant feature that the accuracy of manual marking can be best ensured among the three approaches, while the high amount of work required for marking might discourage engineers/experts [30,47]. Thus, it is crucial to explore a new approach that can automatically learn the domain knowledge of expert judgment, and apply the learned knowledge to quickly and accurately assess the remaining data of the same sample type [48–51].

Machine learning (ML) approaches, which have emerged as the most popular data-driven methods, are capable of autonomously learning the features of labelled data from multidimensional databases [52–56]. The flexibility of MLs makes them effective in approaching and solving engineering issues, especially those which are highly complex and nonlinear [57]. For an ML approach, the built-in hyper-parameters are learned automatically through the "black box" operation, the only manual operation being the creation of an appropriate training set by domain experts [58]. To date, MLs are more widely applied in 2D images than 3D point clouds since the data structure in 2D images is relatively simple [59]. Although MLs have now unconsciously penetrated into numerous technical fields, the subsurface field is still lagging behind due to the evident difficulties in the original data acquisition [53]. Therefore, academia and industry are still committed to exploring the possibility and effectiveness of MLs in characterizing 3D information, especially in complex environments [60]. To the best of our knowledge, MLs have not yet been applied in the domain of 3D rock trace classification during the literature review.

In ML-based classification, the occurrence of imbalanced data identified as a common issue has been receiving increasing body of literature devotion [61]. For example, the number of samples in "trace" and "non-trace" data are inevitably imbalanced due to the different occurrences of the rock trace. One can achieve a relative low error rate by classifying all samples into members of majority classes when the frequency of the observed class labels is highly imbalanced [62]. It has been proven that the imbalanced data seriously hinders the comprehensive performance of MLs [63]. Thus, it is urgent to apply an effective method to alleviate the poor classification caused by imbalanced data.

Considering the above limitations, an ensemble classifier named gradient boosting tree (GBT) [64] is applied for accurate rock trace identification using 3D point clouds collected from a tunnelling project under-construction in Yunnan, China. A thirteen-dimensional database is established from the obtained image samples by a 3D reconstruction method and subsequent feature extraction methods. Then, a synthetic minority oversampling technique (SMOTE) [65] is used to alleviate the data imbalance caused by the "trace" and "non-trace" data. Next, two hyper-parameters optimization algorithms, namely grid search (GS) and random search (RS), are utilized for an optimal ML classifier by selecting an ideal evaluation index [66]. For comprehensive comparisons, three other algorithms, namely, random forest (RF) [67], decision tree (DT) [68], and multiple layers perceptron (MLP) [69] are used for performance evaluation. A total of sixteen hybrid ML classifiers are generated and compared comprehensively. Finally, the feature importance, generalization ability and classification error of the optimal classifier are discussed for comprehensive evaluation.

## 2. Database establishment

As depicted in Fig. 1, digital photography is used to obtain the raw image samples of rock tunnel face in Mengzi-Pingbian Highway (MPH) project in Yunnan, China. The site of MPH is located at the edge of the South China plate, and north of the southern section of the Honghe fault. The geological conditions are determined by the lithology of the strata and landforms. Most tunnel sites in MPH belong to a landform of tectonic denudation with high mountains, and the terrain elevations are considerable. The equipment for digital photography consists of a digital camera (Canon 750D), a measuring ruler, a tripod, and two LED light sources with 1000 W adjustable power.

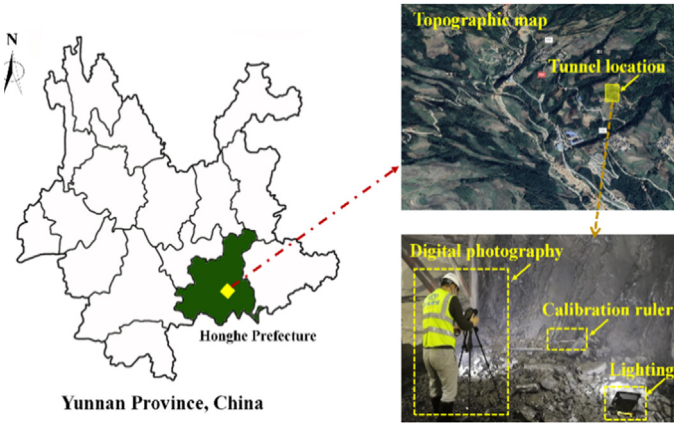


**Fig. 1.** Location of MPH project and the adopted digital photography in rock tunnel.

Then, a structure from motion (SfM)-based technique is applied to create the raw 3D database with the collected image samples (Fig. 2a). The SfM-based method can efficiently realize 3D reconstruction approach since it uses multiple overlapping images to calculate the directional parameters of the camera, thus avoiding additional parameter calibration [70]. The 3D reconstruction process is realized with a software package named PhotoScan in this study [71]. Fig. 2b illustrates the obtained three 3D point cloud datasets, namely, tunnel sections 1, 2 and 3 (TS-1, TS-2, TS-3). The three datasets are selected not only as they contain the typical rock categories (i.e., shale, mudstone, limestone and dolomite) and rock structures (i.e., bedding, and block structures), also they represent the geologic fracture-prone zone widely concerned by the project. Moreover, the statistical information on site conditions of the three selected tunnel faces is listed in Table 1, from which it is seen that complete operations from image acquisition to 3D model reconstruction can be efficiently performed via SfM-based technology.

## 3. Methodology

### 3.1. Feature creation and labelling

#### 3.1.1. Feature creation

Individual point clouds dataset obtained by the SfM-based technique contain only 3D coordinates (*X, Y, Z*) and colour information (*R, G, B*), so they are not adequate descriptors for the deep information they delegate. In isolation, these aforementioned six-dimensional information are insufficient to indicate to which categories these points belong. Hence, it is indispensable to further expand the feature dimensions of each point data to fully reflect the actual properties of the data, especially those belonging to various properties. Fig. 3 describes the establishment of multi-dimensional point features, the methods applied, and the corresponding feature maps. It is found that three series of features are divided with a total of thirteen dimensions including basic features (*X, Y, Z, R, G, B*), vector features ($N_x$, $N_y$, $N_z$, dip (*D*), dip direction (*E*)), discontinuity features (set label (*S*), plane label (*P*)). Taking TS-3 as an example, the detailed generation processes are as follows:

**Basic features:** Fig. 3a presents the main processes of the SfM-based technique using the acquired image samples of a particular tunnel face. During the image acquisition process, a white rigid metre ruler (two endpoints are taken) and two prominent feature points are applied as ground control points to characterise the local coordinate system of a rock tunnel face. The generated 3D point cloud models of the three selected tunnel faces are presented in Fig. 3a. Among them, the format of 3D coordinates can be employed to reveal the relative position relationship and fed into the next-stage features calculations. The fracture occurrences can be calculated for further comparisons whenever initial directions are defined for the collected sections. In addition, the format of colour can reflect the surface texture of the real tunnel, which can be used for the basic judgements for the tunnel characterizations. Overall, the basic features include the above six features, which are directly formed from 3D reconstruction processes.

**Vector features:** Fig. 3b illustrates the main processes of the vectors features generation. A plane-based fitting method first assesses the local surface of the 3D point cloud model. It mainly

**Table 1**
Statistical information of the selected rock tunnel faces.

| Tunnel face cases | TS-1 | TS-2 | TS-3 |
|---|---|---|---|
| Light intensity (LUX) | 258 | 302 | 236 |
| Images employed | 124 | 115 | 105 |
| Shooting time (min) | 11.8 | 8.7 | 12.1 |
| Surface area ($m^2$) | 47.23 | 35.74 | 40.19 |
| Reconstruction time (min) | 84 | 86 | 91 |
| Number of points | 2197472 | 1975082 | 1537018 |

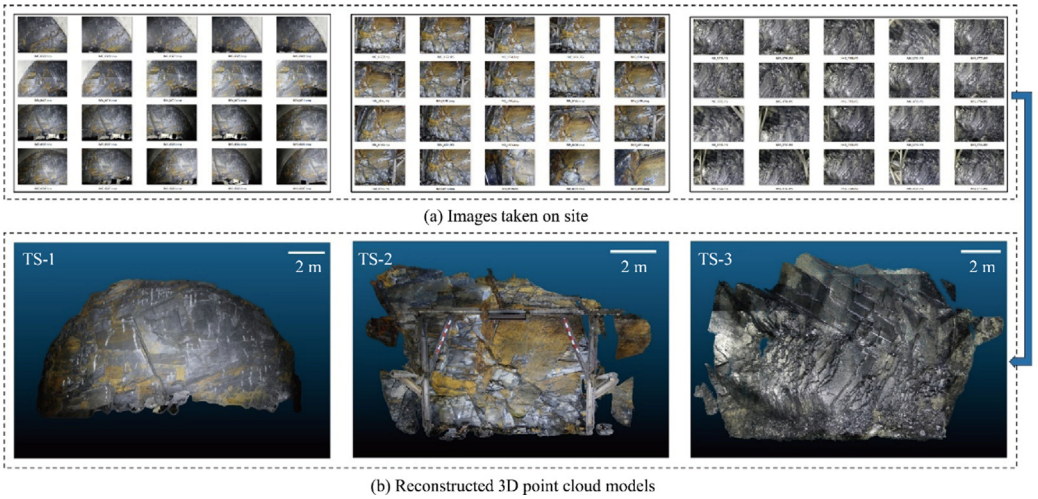


**Fig. 2.** The establishment of 3D point cloud datasets using image samples.

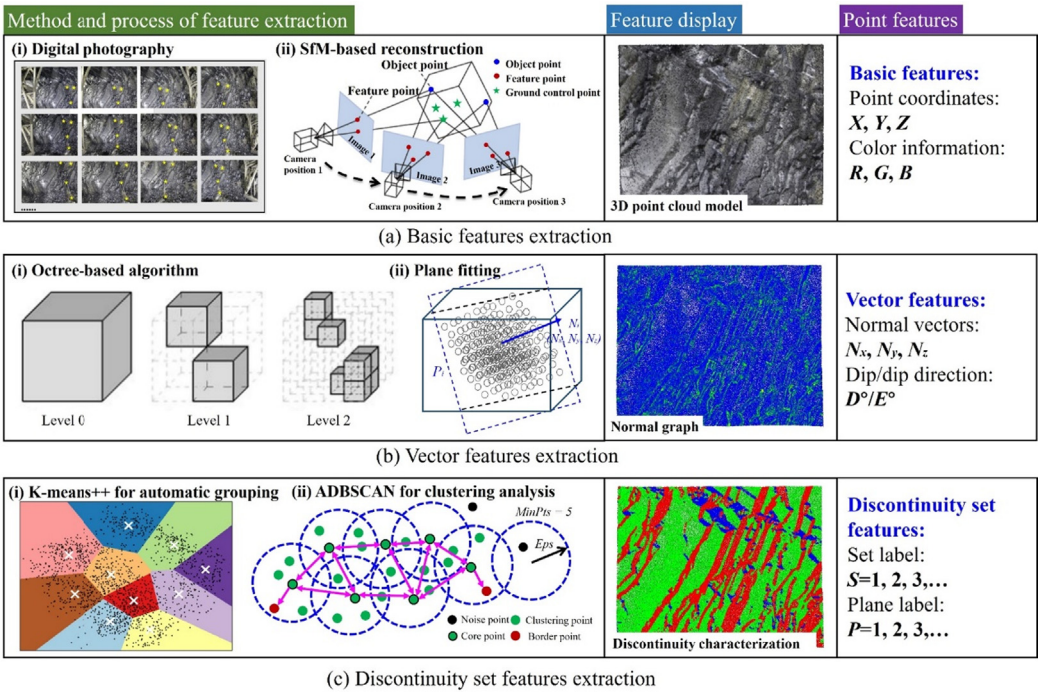


**Fig. 3.** Sketch of the extraction process of 3D point cloud features.

includes two main processes: neighbourhood clustering and orientation calculation. Fig. 3b shows a neighbourhood clustering process by an octree-based algorithm. Applying the octree, the users can retrieve all radius neighbours for an arbitrary query point and radius with regard to a norm [72]. Then, the neighbourhood radius is determined by judging the number of points applied to compute the local surface model. A principle that achieving full coverage search in a relative small-scale radius is needed to follow for the value selection. Next, a minimum spanning tree inspired *k*-nearest neighbour (MST-KNN) method is used to calculate the plane orientation [73,74]. The MST-KNN algorithm is first proposed for outlier detection in 3D point clouds. In this algorithm, a minimal spanning tree is constructed for each set of points selected by *k*-nearest neighbour-based clustering. The minimal spanning tree can detect the directional consistency of the segmented local surface. With this algorithm, the normal vectors (i.e., $N_x$, $N_y$, and $N_z$) of each point for the clustered plane can be calculated, while the corresponding orientations (i.e., dip and dip direction) can be counted for these coplanar points. A typical dip nephogram is selected and plotted in Fig. 3b for the resulting orientations. It is concluded that the vector features mainly contain five relative features: $N_x$, $N_y$, $N_z$, dip, and dip direction.

**Discontinuity set features:** The discontinuity set features of 3D point clouds shown in Fig. 3c contain two main aspects: set labels and plane labels. For the set label, the vector features generated from the former procedure are combined for automatic set classification by a modified *K*-means++ algorithm [75] (see in Fig. 3c). An ultimate set number can be determined by selecting a maximum Silhouette validity index (SVI) value [76]. The SVI is first defined to quantitatively assess the clustering effects: the larger the SVI value is, the closer the sample points are to samples of the same category, and the farther to those of different categories, i.e., the better the clustering results. The optimized total set quantity can be determined in this way, so that the set labels for each judged point can be automatic output. To further improve the overall efficiency, a strategy of random selection for SVI is applied in this work. The results of the set classification are displayed and illustrate the intuitive performances. For the plane labels, an approach called adaptive density-based spatial clustering of applications with noise (ADBSCAN) is employed for a further clustering analysis for each classified discontinuous set [77]. ADBSCAN can automatically detect the optimized edge contour and control the total number of point clouds used for clustering. Thus, for each discontinuity sets, different planes can be detected and numbered, thus the contained point clouds are labelled. Overall, the discontinuity sets have the aforementioned two-dimensional features.

### 3.1.2. Trace labelling

Data labelling with the aim of selecting the relative "ground truth" data is a crucial part of the MLs processes, as it directly affects the self-learning effect of the applied algorithm. The generated rock trace points are assigned the category label 0 "trace", whilst the remaining points are labelled 1 "non-trace". Unfortunately, due to the objective accuracy limit of the extraction algorithm or the subjective difference of manual labelling, it is challengeable to select the so called "ground truth" data. In current researches, the manual labelling by the expert is still widely applied to evaluate the comprehensive classification performances [53]. The possible reason is that an experienced expert can accurately judge the target objects and even perform better than some simple algorithms. Thus, this method is used to produce the original training data and further comparisons of the classified results. However, a single expert may lead to labelling instability. Inspired by the selection strategy used by Weidner et al. [53], a simple expert scoring method (ESM) is proposed to ensure the quality of the labelling information. In ESM, an initial rock mass model is labelled. Next, three experienced experts are asked to rate the overall accuracy of the labelled model using a hundred-point evaluation criterion. The three scores are then averaged and tallied. Another round of modification for the labelled model is required until the mean score is higher than a predetermined threshold (90 in this study). One can adjust the threshold value according to the research requirements. In this way, the trace labelling of training data can meet the certification of experts. Meanwhile,

CloudCompare software is used to thin out the disordered and dense points and display the skeleton of the traces. The main processes of ESM are illustrated in Table 2, where it is seen that a satisfactory score of 90.37 has been obtained when the 6th round was conducted. Accordingly, the labelled and thinned map of the 6th round modified TS-3 model are depicted in Fig. 4, from where it is found that the main traces are generally marked.

### 3.2. Classifiers

#### 3.2.1. Gradient boosting tree (GBT) classifier

As shown in Fig. 5, a gradient boosting tree (GBT) algorithm is adopted to perform the rock mass trace classification experiments. The GBT enhances a decision tree (DT) by a boosting algorithm [64], of which the core concept is to aggregate weak models to form a single strong consensus model, rather than building a new optimized model. In a DT [78], the feature space is firstly classified into sub-regions in order to model the dependent variable for each region [79,80]. Then, each individual region is further divided into new sub-regions to model the relative variables. Repeating these processes until all the completion conditions are met. Within each region, the best fitness can be achieved by selecting the split point and variables. Among them, the end node count is defined as the size of a single tree. In this paper, the GBT is taken to calculate the classification value $p$ by training a model $F$ via a least-squares regression. Adding an estimator further improves the model in a forward stage-wise strategy:

$$F_t(x) = F_{t-1}(x) + \gamma_t z_t(x) \tag{1}$$

where $F_t$ refers to the GBT model consisting of $t$ DTs, $t$ donates the total number of DTs; $\gamma_t$ the learning rate; and $z_t(x)$ the weak learners. Next, a new DT is added to original GBT model for each boosting iteration $t$. The value of $z$ can be calculated:

$$F_t(x) = F_{t-1}(x) + z(x) = p \tag{2}$$

$$z(x) = p - F_{t-1}(x) \tag{3}$$

Then, a weighted sum function $\hat{F}(x)$ is applied for approximating the classified $z_t(x)$ as given by:

$$\hat{F}(x) = \sum_{i=1}^{t} \gamma_i z_i(x) + \text{const} \tag{4}$$

For each DT, the GBT model applies the variables that contribute most to the reduction of loss function $L$ to minimize the ultimate $L$, that is, applying the empirical risk minimisation principle. Specifically, the operation of empirical risk minimization is started with a constant function $F_0(x)$ and gradually optimized:

$$F_0(x) = \arg\min_{\gamma} \sum_{i=1}^{n} L(p_i, \gamma) \tag{5}$$

Next, the initial empirical risk of the $F_0(x)$ is gradually minimized as the boosting iteration $t$ progresses. The following formula is obtained:

$$F_t(x) = F_{t-1}(x) + \arg\min_{z \in Z} \sum_{i=1}^{n} L(p_i, F_{t-1}(x_i) + z(x_i)) \tag{6}$$

where $Z$ refers to the set of total weak learners $z_t(x)$. If the continuous case is considered in the boosting process, that is, $Z$ is the set of differentiable functions, the minimisation process can be addressed by the steepest descent method as follows:

$$F_t(x) = F_{t-1}(x) - \gamma_t \sum_{i=1}^{n} \nabla_{F_{t-1}} (L(p_i, F_{t-1}(x_i))) \tag{7}$$

$$\gamma_m = \arg\min_{\gamma} \sum_{i=1}^{n} L(p_i, F_{t-1}(x_i) - \gamma \frac{\partial L(p_i, F_{t-1}(x_i))}{\partial F_{t-1}(x_i)}) \tag{8}$$

where $\gamma_m$ denotes the learning rate obtained by minimizing the loss value of the leaf node of the $m$-th CART.

Meanwhile, it is learned that a relative small learning rate can promote the models' generalization ability over boosting without shrinkage [64]. Unfortunately, it comes at an increasing computational cost by using more DTs. Other parameters such as the number of trees $T$ and a maximum number of the splits which closely relate to the ultimate model's complexity and structure, must also be fine-tuned to optimize the performance.

#### 3.2.2. Other comparative classifiers

To conduct a comprehensive comparative analysis, DT, RF, and MLP are adopted in this study. Among them, DT is applied for comparison as it is a common root algorithm to build various ensemble models by various integration methods, such as boosting, and bagging [81]. RF is proposed as a typical model with bagged integration of the random sub-space and boot-strap aggregation methods [67,82]. In this approach, the number of samples required in training bootstrap would be optional that is lower than that of the original set. Next, all the bootstrap sets are used to form a single homogenized DT. Each node in a DT denotes a prediction criterion so the tree is related to the output labels, and thus a DT can ultimately predict the bootstrap sample via obtaining the random characteristics of all the nodes. The number of leaves of each DT remains constant through the entire RF model. Next, the output classification results can be further determined and then designated as out of bag (OOB) classification. Subsequently, the final output classification can be computed by averaging the outputs of all DTs [68]. MLP is one of the types of feed-forward artificial neural networks frequently used to produce regression analysis mathematical models [83]. Consisting of an input layer, hidden layers, and an output layer, it can distinguish data that cannot be linearly separated. Except for the input layers, each node in the following layers is a single neuron using a nonlinear activation function. Training the MLP uses back-propagation, a supervised learning procedure that classifies outputs from inputs. The multiple layers and nonlinear activation are the main features of the MLP that distinguish it from a linear perceptron.

### 3.3. Synthetic minority oversampling technique (SMOTE)

In Table 3, the number of point clouds in different categories is roughly counted by utilizing labelled data, and the approximate data imbalance rates of three cases are calculated. It is seen that the category imbalance is commonly existed in these tunnel sections, making the classified results prone to the classes with major-

**Table 2**
The main processes of expert scoring method.

| Score round | 1st | 2nd | 3rd | 4th | 5th | 6th |
|---|---|---|---|---|---|---|
| Expert-1 | 65 | 73 | 85 | 93 | 83 | 92 |
| Expert-2 | 74 | 78 | 83 | 87 | 91 | 86 |
| Expert-3 | 76 | 73 | 92 | 85 | 81 | 94 |
| Average | 71.7 | 74.7 | 86.7 | 88.3 | 85.0 | 90.7 |

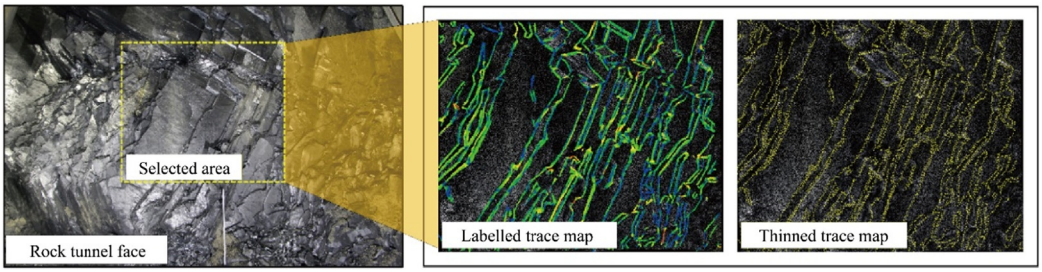


**Fig. 4.** Labelled and thinned map of the 6th round's TS-3 model.

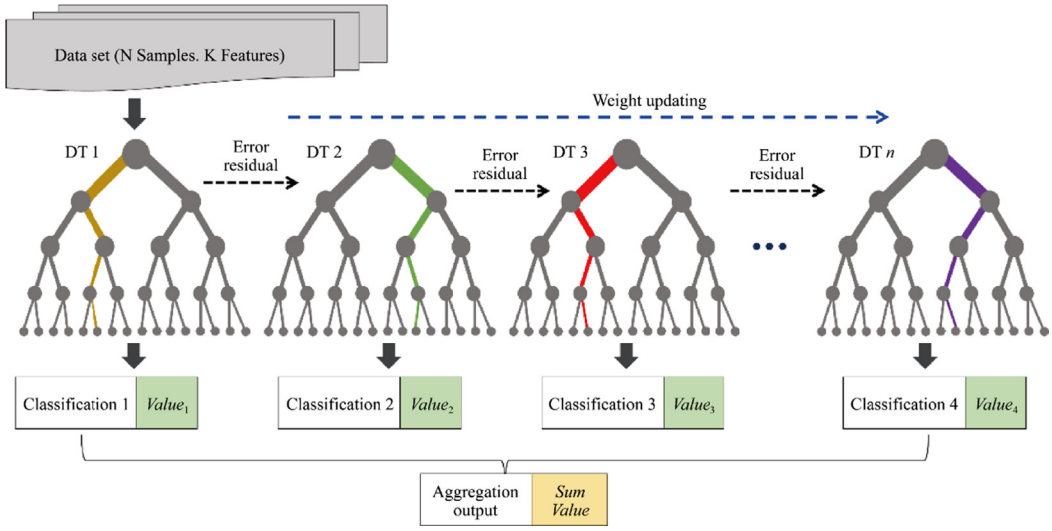


**Fig. 5.** Flowchart of the gradient boosting tree model.

ity samples, thus reducing the overall recognition accuracy. For example, for TS-2, the applied model is implemented on large labelled where about 12.5% of the obtained points are "trace". If this dataset is directly applied for classification, the classifier achieves about 87.5% accuracy even if it always predicted "non-trace". Thus, increasing the classification rate of minority samples is vital for the applications. Subsampling for majority samples and oversampling for minority samples are two common methods to handle the imbalanced samples [84]. Since there are few trace samples, the oversampling technique is undoubtedly more suitable. The original oversampling technique can easily lead to overfitting since it directly copies the minority class samples. It is not subservient to increase the classification rate of the minority category.

The SMOTE algorithm (see in Fig. 6) can simply and effectively reduce the imbalance phenomenon by generating new samples of minority categories without repetition [85]. The main processes are as follows:

**Step 1:** For each sample $x_i$ in the minority category, the Euclidean distance is taken as an index to calculate from $x_i$ to all samples in the same class sample set. Thus, the corresponding $k$-nearest neighbour is obtained.

**Table 3**
Imbalanced data ratio between "trace" and "non-trace" data for three tunnel sections.

| Tunnel sections | TS-1 | TS-2 | TS-3 |
|---|---|---|---|
| Imbalance ratio (trace: non-trace) | 1:7 | 1:8 | 1:9 |

**Step 2:** In accordance with the imbalance ratio of the applied dataset, a sampling ratio $M$ is set to determine the new proportion of various samples. For $x_i$, a few samples are randomly chosen from its $k$-nearest neighbours, assuming that the chosen nearest neighbour is $x_m$. For each randomly chosen neighbour $x_m$, a new sample is generated from the original sample on the basis of Eq. (9).

$$x_{\text{new}} = x + \text{rand}(0,1) \times |x_i - x|, \text{new} \in 1, 2, ...., M \tag{9}$$

where rand (0,1) denotes a random digit in the range of 0 to 1.

**Step 3:** Repeat the above steps for $M$ times to generate $M$ new samples. That is, if the minority category has a total number of $T$, $MT$ new samples are generated. After merging the new samples with the few original samples, a new balanced dataset is generated.

### 3.4. Hyper-parameters tuning

Two commonly applied algorithms (i.e., grid search (GS), and random search (RS)) are selected to optimize the hyper-parameters [86]. Among them, GS can adjust the parameters according to the step size within the specified parameter range, and use the adjusted parameters to train the classifier to find the parameter with the highest accuracy in the verification set from all parameters. This is actually a cycle and comparison process. Meanwhile, the RS replaces GS with random sampling in parameter space. Its theoretical basis is that if the random sample set is large enough, then one can find the global maximum or minimum or their approximations. It can sample parameters with continuous variables as a distribution, which GS cannot.

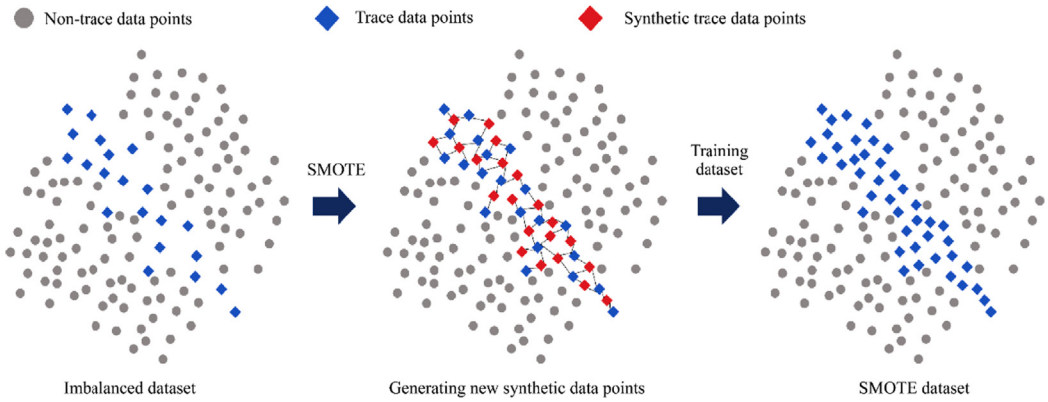


**Fig. 6.** The principle of synthetic minority oversampling technique (SMOTE).

Moreover, the *k*-fold cross-validation (CV) algorithm is proposed for robustness improvement and overfitting avoidance [87]. In *k*-fold cross-validation, the dataset is divided into *k* equal parts, one of which is the validation data and the other *k*-1 data is the training data. Then, in cross-validation, the experiment is repeated for *k* times. In each experiment, a different data part from *k* parts is selected as the validation data (to ensure that the data of *k* parts have been validated separately), and the remaining *k*-1 are considered as the training data. Finally, the *k* experimental results are divided equally. In this study, a 10-fold CV, i.e., *k* is equal to 10, is chosen for the trace classification task.

### 3.5. Evaluation indicators

It is vital to assess the classification models using the appropriate performance indices. In this study, the F score indicator is applied to assess the comprehensive performances of different classifiers by contrasting the classified category with the true category [88]. F score is a definition of harmonic mean of recall and precision and the corresponding expression is given by:

$$\text{F score} = \frac{2 \times (\text{Precision} \times \text{Recall})}{\text{Precision} + \text{Recall}} \tag{10}$$

## 4. Results

### 4.1. Oversampling analysis using SMOTE algorithm

In the whole processes, the 3D rock mass samples are divided into a training set and a testing set with a fixed ratio of 8:2 by applying the mentioned 10-fold-CV method. The SMOTE-GBT algorithm is applied for the trace classification for the established datasets. In the training process, the corresponding datasets are fed into the proposed model to construct basic SMOTE-GBT classifiers. Then, the validation processes are performed to further optimize the weights and biases of the models. Next, the testing processes of the three cases are conducted using the validated classifiers. To comprehensively compare the performances of the SMOTE algorithm, the F score values of various imbalance ratios of the three cases are counted and shown in Fig. 7. It is found that the classification performance of the three cases in trace and non-trace improves significantly with the gradual alleviation of data imbalance. In general, the effect of using the SMOTE oversampling algorithm is mainly divided into two phases: the fast ascent phase and the stable phase. For the three cases of this paper, a cut-off point can be applied as a critical evaluation of SMOTE technology. This is because once the ratio exceeds the cut-off point, the F score values no longer vary greatly with the change of imbalance rate. It is undoubtedly evident from Fig. 7 that a ratio of 1:5 to 1:4 should be selected as the ideal cut-off point in these 3D rock mass cases.

In addition, it is found that the F score of non-trace classification for each imbalanced ratio is higher than that of trace in all selected tunnel faces. The potential reason is that the initial number of non-trace points is much larger than that of trace points which directly affects the calculation of the total loss of SMOTE-GBT classifier. Meanwhile, the judgment of linear trace points is undoubtedly more challenging than non-trace points since there may be a considerable number of misjudgements for ambiguous points. Moreover, it is almost impossible to avoid errors in trace points labelling although the manual process has been extensively evaluated by experts. Overall, applying an SMOTE technology can effectively improve the classification performance caused by data imbalance to some extent.

### 4.2. Hyper-parameters tuning analysis

To further select an optimal classifier, a comparative analysis is then performed by applying grid search- and random search-based optimization algorithms. The two optimization algorithms are programmed in Python language. In accordance with the conclusion of Section 4.1, the ratio of trace points and non-trace points is set to 1:5 by SMOTE technology to improve the overall F score values. Once the algorithms reach their maximum F score for trace data, the optimal hyper-parameters are recorded.

Table 4 lists the setting of hyper-parameters for the SMOTE-GBT classifier, the optimized hyper-parameters for the GS-SMOTE-GBT, and RS-SMOTE-GBT classifiers, and the computational costs for all classifiers. It is seen that the ultimately selected hyper-parameters are different for the three hybrid classifiers although the searching space and steps are controlled the same. As for the number of hyper-parameter searches, GS-SMOTE-GBT had the largest value, followed by RS-SMOTE-GBT, and SMOTE-GBT. Accordingly, the classifiers GS-SMOTE-GBT and SMOTE-GBT have the longest and least runtime, respectively. Owing to the large number of TS-2 data samples, no matter which optimization method is used, it consumed more time than the other two tunnel faces of the same classifier. All the hyper-parameters of SMOTE-GBT classifier for the three tunnel faces are determined by the default value of the original GBT classifier and remain the same. Meanwhile, the detailed hyper-parameters for different tunnel face samples performed with the same classifier are different. The potential reason is that different datasets may produce different weights and biases of

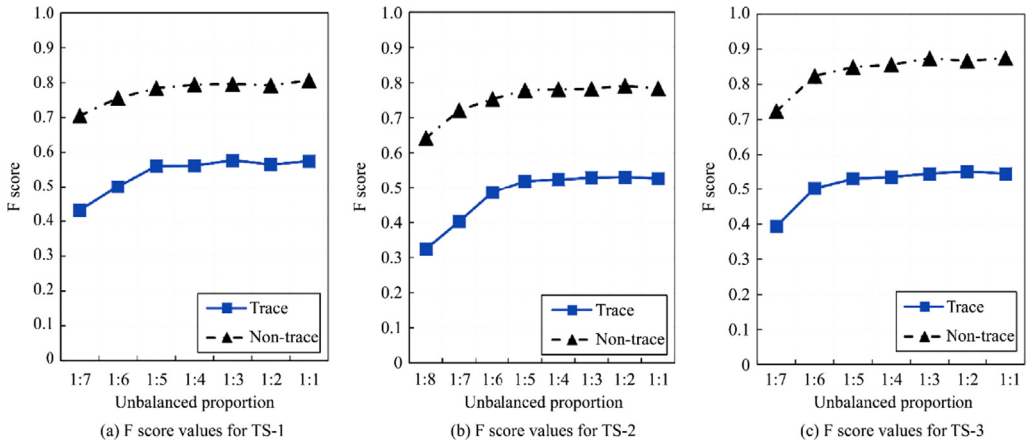


**Fig. 7.** The F score values of both trace and non-trace classification for various cases by using SMOTE oversampling algorithm.

**Table 4**
Hyper parameters optimization of the proposed SMOTE-GBT classifier.

| Hyper-parameters | Searching space | Searching step | SMOTE-GBT | | | GS-SMOTE-GBT | | | RS-SMOTE-GBT | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | TS-1 | TS-2 | TS-3 | TS-1 | TS-2 | TS-3 | TS-1 | TS-2 | TS-3 |
| tree_num | [10, 1000] | 10 | 35 | 35 | 35 | 80 | 630 | 540 | 420 | 510 | 390 |
| max_features | [1, 15] | 1 | 2 | 2 | 2 | 2 | 4 | 8 | 5 | 11 | 8 |
| max_depth | [5, 50] | 1 | 15 | 15 | 15 | 31 | 42 | 29 | 24 | 19 | 11 |
| min_samples_split | [2, 11] | 1 | 5 | 5 | 5 | 4 | 8 | 6 | 6 | 2 | 3 |
| max_samples_split | [1, 11] | 1 | 5 | 5 | 5 | 9 | 5 | 10 | 7 | 8 | 9 |
| learning_rate | [0.001, 0.1] | – | 0.001 | 0.001 | 0.001 | 0.005 | 0.092 | 0.087 | 0.036 | 0.064 | 0.075 |
| Number of searches | | | – | – | – | 1762 | 1104 | 523 | 36 | 74 | 42 |
| Accumulated running time (s) | | | 62.3 | 71.9 | 48.2 | 2704 | 2197 | 1840 | 407 | 682 | 352 |

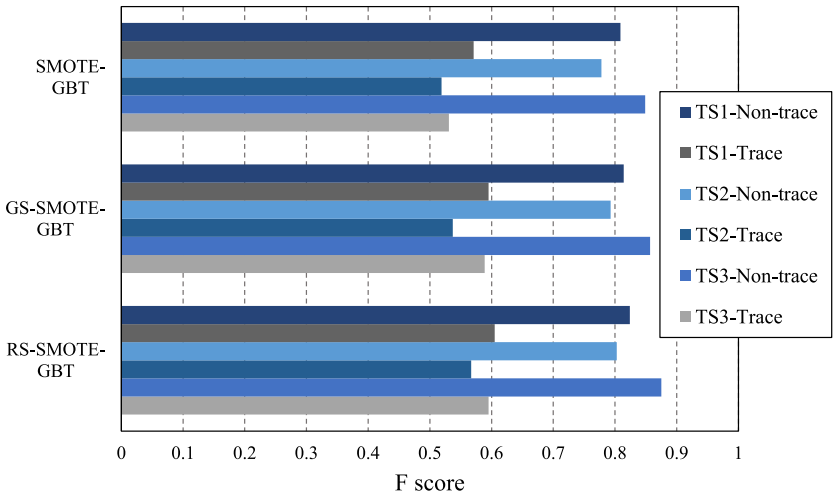


**Fig. 8.** The F score values for various tunnel faces by using hyper-parameters optimization algorithm.

the hybrid ML models, which affects the assignment of the hyper-parameters. Also, the corresponding classification performances for the three models are shown in Fig. 8, where it is found that the SMOTE-GBT has the worst F score performance among the three classifiers. The RS undoubtedly outperforms the other two optimization algorithms in terms of F score for both trace and non-trace classification. Additionally, it is observed that the F score of non-trace classification is still higher than that of trace classification for all the optimized classifiers for any imbalanced ratio. In summary, applying the random search-based algorithm demonstrated an ideal hyper-parameters optimization for the hybrid SMOTE-GBT classifier.

### *4.3. Analysis for various hybrid ML approaches*

Subsequently, another ensemble learning classifier (i.e., RF) and two single machine classifiers (i.e., DT and MLP) are applied to perform comprehensive comparisons. During the training processes,

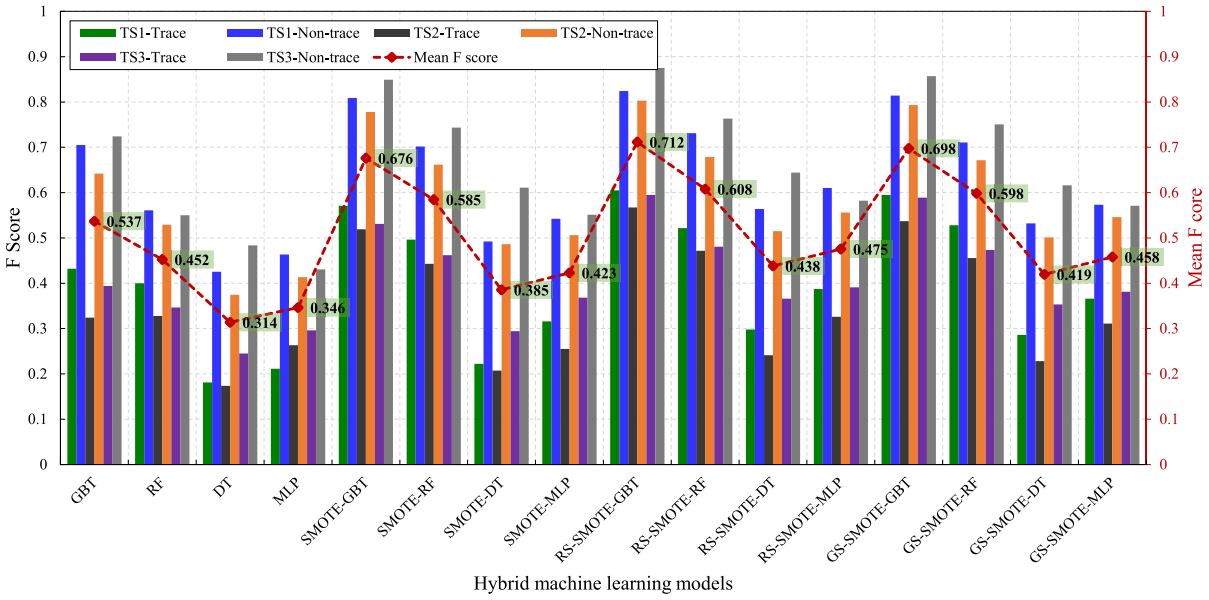


**Fig. 9.** The performance of both trace and non-trace classification for various cases by applying various machine methods.

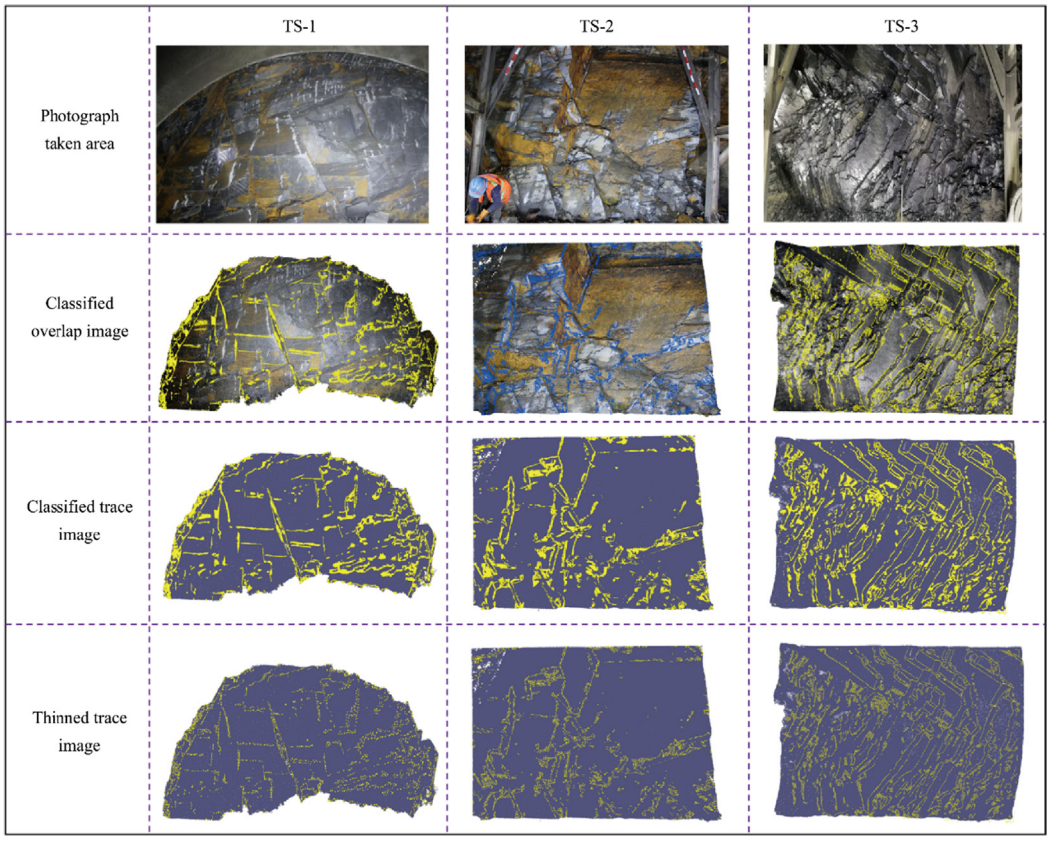


**Fig. 10.** RS-SMOTE-GBT classifier for rock trace classification results using samples from three rock tunnel faces, including: photograph taken area, classified overlap image, classified trace image, and thinned trace image.

the SMOTE technology is applied to generate the class imbalanced samples to control the ratio between the number of trace points and non-trace points to 1:5, and then GS and RS are used to search for the optimal hyper-parameters sequentially. Once the optimization of the maximum F score is finalised, the training process is terminated and the corresponding pre-trained classifiers are recorded for the following testing process. In total, sixteen models (see in Fig. 9) are used for the classification tasks of three tunnel faces. In Fig. 9, the F score values of all classifiers for the three tunnel sections are computed and then the corresponding mean values for each classifier are then obtained. As can be observed, the mean F score of RS-SMOTE-GBT is 0.712, which is higher than other classi-

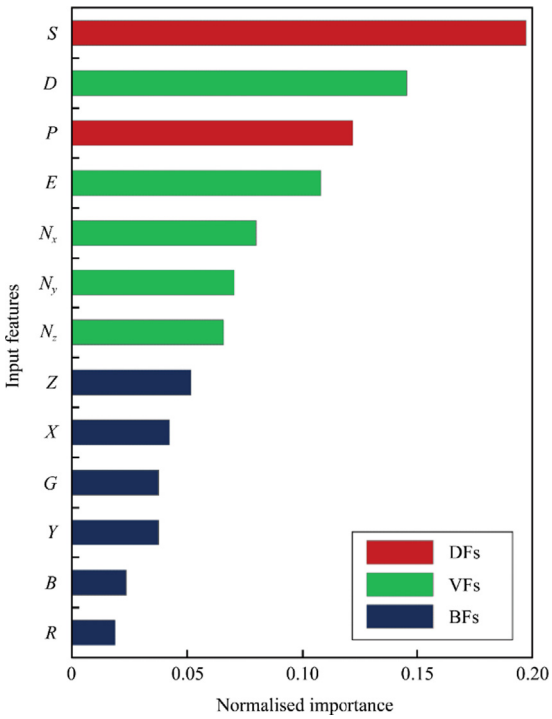


**Fig. 11.** Variable importance of rock trace classification using RS-SMOTE-GBT.

fiers, followed by the GS-SMOTE-GBT, SMOTE-GBT, GS-SMOTE-RF, RS-SMOTE-RF, and SMOTE-RF, etc. The performance ranking of the four basic ML models (i.e., GBT, RF, MLP, and DT) basically determines the comprehensive performance ranking of hybrid classifiers, that is, GBT outperforms RF, MLP, and DT, orderly. Whilst the SMOTE algorithm is proven an excellent improvement for the rock trace classification in all the applied methods in this work. By adjusting the weights of the models to obtain the optimal F score, the hyper-parameters optimization approaches (i.e., RS, and GS) can further enhance the classification performances. Also, the classification performances of non-trace data are significantly better than that of trace data for three tunnel faces. Moreover, for most proposed hybrid classifiers, the rock traces in TS-3 are obviously easier to be classified than those in TS-1 and TS-2.

### 4.4. Qualitative observations

A visual observation can be applied to confirm the general feasibility and effectiveness of a classifier in extracting and interpreting real rock morphology. In this section, the RS-SMOTE-GBT classifier is selected since it is proven an optimal classification performance for rock trace from various tunnel faces. Correspondingly, the qualitative observation of RS-SMOTE-GBT is further evaluated for a comprehensive assessment. Fig. 10 illustrated the classification results of rock trace for various tunnel sections by applying the hybrid RS-SMOTE-GBT classifier. The corresponding overlap and thinning trace maps are also plotted. It is observed from Fig. 10 that most rock traces are correctly classified as expected, especially for TS-3. However, there are also some areas with relatively poor classification performance that need further optimization. Notably, it can be seen that the obviously and sparsely distributed rock trace lines are much easier to be accurately classified than dense and ambiguous trace lines. Meanwhile, the classification performance of the proposed model for the samples of regular layered rock is better than that of irregular rock block. Typically, there are frequent misclassifications in the fractured areas, gravel deposits and angular areas in tunnel faces. The specific analysis for the misclassifications is discussed further in Section 5.

## 5. Discussions

### 5.1. Feature importance evaluation

The measure of variable importance is defined as the extent of contributions to the output from various input variables in ML model [89]. By applying the optimized RS-SMOTE-GBT classifier, variable importance is calculated by examining the effects of variations in the input variables on the Gini index [79]. The Gini indexes are then normalized so that the sum of all importance values equals 1. A larger variable importance value means a more momentous influence on the classifiability of model. Fig. 11 describes the importance values across the trace classification variables, this presents that more critical influence features are primarily from the discontinuity features (DFs), although certain ones from the vector features (VFs) also score highly. Along these features, the highest scoring variables (above 0.10) are dip (*D*), dip direction (*E*), set label (*S*), and plane label (*P*), whereas colour variables (*R, G, B*) suggested lower correlations with the trace classification. It can be seen that the importance of the basic features (BFs) is greatly weakened as they are transformed into new features, which indirectly illustrates that the newly generated features in data pre-processing are beneficial for the ultimate classification.

### 5.2. Generalization ability evaluation

To further test the generalization of the proposed RS-SMOTE-GBT classifier, the trained optimal models are used to test samples from various tunnel faces (i.e., TS-1, TS-2, and TS-3). Seven trained models, namely M1, M2, M3, M1+2, M1+3, M2+3, and M1+2+3, are obtained by training different TS combinations. Among them, Mi refers to the trained model with only TS-*i* data, M*i*+*j* refers to the trained model with TS-*i* and TS-*j* data. The F score values for the corresponding rock trace and non-trace classification by using various combinations of training and testing sets are summarized in Fig. 12. The average columns on the right side and bottom of Fig. 12 can reflect the generalization ability of each classifier and the recognition difficulty of each data set, respectively. For the model trained with one TS data (see in Figs. 12a and b), it is observed that the trained M1 has the highest generalization ability for rock trace points classification with an average F score of 0.392, closely followed by M3 and M2, respectively. However, there is a difference in the ranking of generalization ability for non-trace classification, with the order of strong to weak being M3, M1, and M2. TS-3 is identified as the easiest database to identify in both trace and non-trace classification. It may be concluded that sections with similar surface characteristics (i.e., TS-1 and TS-3) will be better suited for the final model generalization. Meanwhile, it is also found that the test performance of a classifier on the approximate database significantly outperforms that on the data with significant differences. For the model trained with two TSs (see in Figs. 12c and d), it is seen that the overall generalization ability improves with the enrichment of the training sets. The trained M2+3 has the highest generalization ability for classifying both trace and non-trace points, followed by M1+2 and M1+3. Meanwhile, TS-3 is still identified as the easiest database to identify for classifying trace and non-trace, followed by TS-1, and TS-2. As shown in Figs. 12e and f, the recognition difficulties of the three datasets have the same ranking as in Figs. 12 c and d. The overall generalization ability of the trained M1+2+3 reaches the highest value compared to the previous model. Thus, increasing


**Fig. 12.** The F score values of trace and non-trace classification for various combinations of training and testing using RS-SMOTE-GBT classifier.

the diversity of the training samples undoubtedly promotes the overall generalization ability of the classifiers.

### 5.3. Classification error discussion

As for classification error, a potential yet hard to quantify factor is domain bias in the data labelling process [48]. Though the whole labelling process has been finely assessed by ESM, the "trace" points manually labelled by the experts still hard to guarantee the accuracy for completely representing rock trace in reality. In particular, in areas with fractured rocks and heavy deposition, labelling errors may inevitably occur in practice due to the complexity of the surface. Meanwhile, previous literature has shown that there are unavoidably several variabilities in human interpretation of basic rock trace features in 3D point clouds and in more general geological data, even among experts [90]. Thus, the above two factors are reflected in the interpretations of the proposed classifier trained on imperfect labels. Taking TS-3 as a typical example for the analysis of error reduction, the hard-to-label areas determined by experts are eliminated step by step. Seven elimination areas are manually selected and marked in Fig. 13. During the experiment, the point clouds in these areas are eliminated in sequence, and the corresponding F score values are recorded. For each elimination attempt, the proposed RS-SMOTE-GBT algorithm is re-run with an 8:2 ratio of training and test sets, and a 1:5 imbalance rate of trace and non-trace points. The counted F score values of trace classification for the partially eliminated TS-3 datasets are indicated in Fig. 13. It is observed that the F score increases nonlinearly with the elimination of difficult labelled- point clouds (1st to 6th area), whereas it does not fluctuate obviously for the elimination of easy labelled- point clouds (7th area). In general, eliminating difficult labelled areas can improve the overall performance (i.e., F score) of the classifier by about 25%. The potential reason is that the difficulty of training and testing is greatly reduced by the above elimination. It is inferred that cleaning up the sedimentary pumice and reducing the area of fractured rock are conducive to improving the overall performance of the classifier. To some extent, enhancing the accuracy and resolution of the point cloud can also help improve the accuracy of the data labelling accuracy to ensure the effectiveness of active learning of the classifier. Thus, another factor influencing the classification error is the quality of the collected datasets. This is consistent with "the importance of reliable and diverse samples" as pointed out in the previous literature [91,92]. However, collecting ideal training data is perhaps one of the most challengeable procedures in developing a suitable classifier, especially in an underground space context. Undoubt-

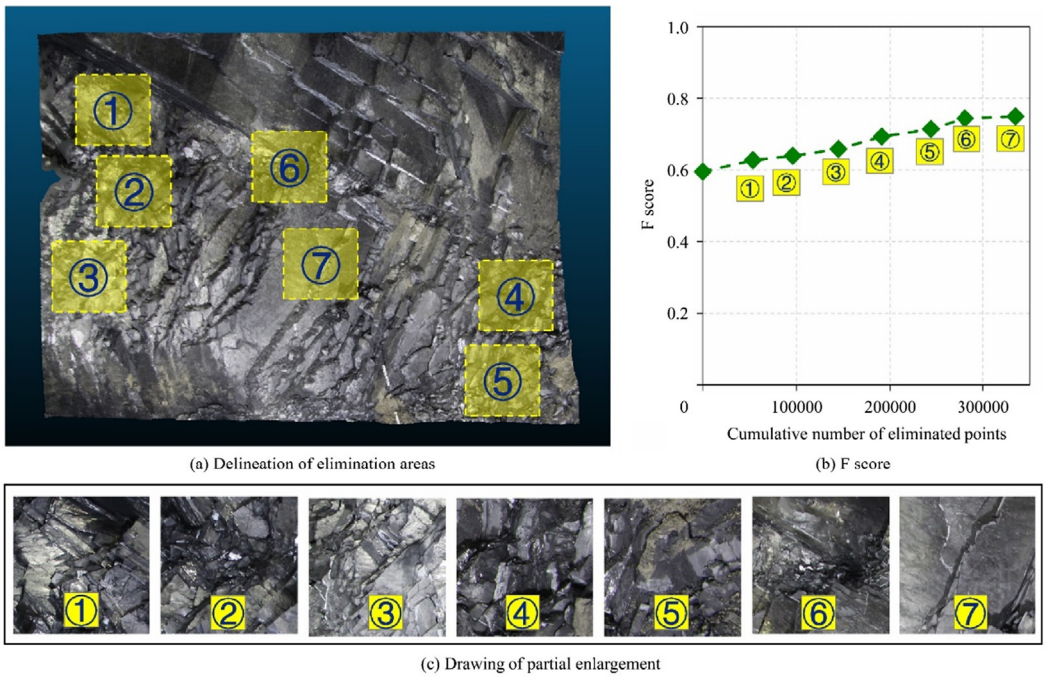


**Fig. 13.** The F score values of trace classification by eliminating fractured rock areas.

edly, the importance of a global database is self-evident for a system that aims at robust discrimination. Therefore, more high precision datasets from different tunnel face should be added continuously to the established database. An enhanced classifier can then be trained and tested in future work to enhance its practical application range.

## 6. Conclusions

In this paper, a hybrid ensemble algorithm named RS-SMOTE-GBT is presented to achieve the classification of rock trace despite the existing imbalance of categories in the datasets used. By applying the SMOTE technique, hyper-parameters optimization algorithms, and 10-fold cross validation, the performance of sixteen hybrid classifiers for rock trace identification is systematically investigated. In addition, the generalization considerations for ML-based point cloud classifiers are illustrated. The main findings are presented below:

(1) The SMOTE technology can effectively improve the classification performances caused by data imbalance to some extent. Given full consideration of classification and computational cost, it is recommended that the ratio of trace and non-trace points should be selected with a value of 1:5 to 1:4 for the applied tunnel section data.
(2) The RS-based hyper-parameters optimization algorithm is proven a more ideal improvement for the hybrid SMOTE-GBT classifier than GS. The RS-SMOTE-GBT classifier outperforms the other fifteen hybrid ML algorithms for both trace and non-trace classification with a highest mean F-score value of 0.712.
(3) The sensitivity analysis indicates that more critical influencing features are primarily from the discontinuity features (DFs), which illustrates that the newly generated features are beneficial for the ultimate results.
(4) Increasing the diversity of training samples can promote the overall generalization ability of classifiers. TS-3 is closed as the most easily identifiable database in all classification combinations.
(5) Cleaning the sedimentary pumice and reducing the area of fractured rock are conducive to improving the overall performance of the classifier for trace classification. The overall F score value can be improved by about 25%.

It should be noted that domain biases inevitably affect the generalization ability of the proposed classifier. This may be caused by variance in the labelling interpretations or the inherent differences in the collected 3D point cloud samples. Thus, the importance of a global database is self-evident for a system aiming at robust discrimination. More high precision datasets are urgent to be added continuously to further enhance the practical scope of the classifier.

## Acknowledgements

The research presented in this paper is supported by Key innovation team program of innovation talents promotion plan by MOST of China (No. 2016RA4059), Natural Science Foundation Committee Program of China (No. 51778474), and Science and Technology Project of Yunnan Provincial Transportation Department (No. 25 of 2018).

## References

[1] Barton N. Suggested methods for the quantitative description of discontinuities in rock masses. Int J Rock Mech Min Sci Geomech Abstr 1979;15(6):319–68.
[2] Hudson JA, Harrison JP, Popescu ME. Engineering rock mechanics: an introduction to the principles. Appl Mech Rev 2002;55(2):B30.
[3] Cai WQ, Zhu HH, Liang WH, Zhang LY, Wu W. A new version of the generalized Zhang-Zhu strength criterion and a discussion on its smoothness and convexity. Rock Mech Rock Eng 2021:1–17.

[4] Bieniawski Z. Engineering classification of jointed rock masses. Civil Engineer in South Africa 1973;5(12):353.
[5] Palmstrom A. RMi – a rock mass characterization system for rock engineering purposes Doctoral dissertation. Norway: University of Oslo; 1995.
[6] Lilly P. The use of the blastability index in the design of blasts for open pit mines. In: Proceedings of Western Australian conference on mining geomechanics. Kalgoorlie, West Australia; 1992. p. 8–9.
[7] Moomivand H, Vandyousefi H. Development of a new empirical fragmentation model using rock mass properties, blasthole parameters, and powder factor. Arab J Geosci 2020;13(22):1–17.
[8] Hoek E, Carranza C, Corkum B. Hoek-Brown failure criterion. 2002 ed. University of Toronto Press; 2002.
[9] Hoek E. Strength of rock and rock masses. ISRM News J 1994;2:4–16.
[10] Cunningham C. Fragmentation estimations and the Kuz-Ram model-four years on. In: Proceedings of the 2nd international symposium on rock fragmentation by blasting. Colorado: SEM; 1987. p. 475–87.
[11] Azizi A, Moomivand H. A new approach to represent impact of discontinuity spacing and rock mass description on the Median fragment size of blasted rocks using image analysis of rock mass. Rock Mech Rock Eng 2021;54 (4):2013–38.
[12] Chen JY, Yang TJ, Zhang DM, Huang HW, Tian Y. Deep learning based classification of rock structure of tunnel face. Geosci Front 2021;12 (1):395–404.
[13] Jing HW, Wu JY, Yin Q, Wang K. Deformation and failure characteristics of anchorage structure of surrounding rock in deep roadway. Int J Min Sci Technol 2020;30(5):593–604.
[14] Ajayi KM, Schatzel SJ. Transport model for shale gas well leakage through the surrounding fractured zones of a longwall mine. Int J Min Sci Technol 2020;30 (5):635–41.
[15] Chen JY, Zhou ML, Zhang DM, Huang HW, Zhang FS. Quantification of water inflow in rock tunnel faces via convolutional neural network approach. Autom Constr 2021;123:103526.
[16] Gangi AF. Variation of whole and fractured porous rock permeability with confining pressure. Int J Rock Mech Min Sci Geomech Abstr 1978;15 (5):249–57.
[17] Chen JY, Zhou ML, Huang HW, Zhang DM, Peng ZC. Automated extraction and evaluation of fracture trace maps from rock tunnel face images via deep learning. Int J Rock Mech Min Sci 2021;142:104745.
[18] Chen JY, Huang HW, Zhou ML, Chaiyasarn K. Towards semi-automatic discontinuity characterization in rock tunnel faces using 3D point clouds. Eng Geol 2021;291:106232.
[19] Kemeny J, Post R. Estimating three-dimensional rock discontinuity orientation from digital images of fracture traces. Comput Geosci 2003;29(1):65–77.
[20] Riquelme A, Cano M, Tomás R, Abellán A. Identification of rock slope discontinuity sets from laser scanner and photogrammetric point clouds: a comparative analysis. Procedia Eng 2017;191:838–45.
[21] Reid TR, Harrison JP. A semi-automated methodology for discontinuity trace detection in digital images of rock mass exposures. Int J Rock Mech Min Sci 2000;37(7):1073–89.
[22] Lemy F, Hadjigeorgiou J. Discontinuity trace map construction using photographs of rock exposures. Int J Rock Mech Min Sci 2003;40(6):903–17.
[23] Tsai YC, Kaul V, Mersereau RM. Critical assessment of pavement distress segmentation methods. J Transp Eng 2010;136(1):11–9.
[24] Zalama E, Gómez-García-bermejo J, Medina R, Llamas J. Road crack detection using visual features extracted by Gabor filters. Comput-Aided Civ Infrastruct Eng 2014;29(5):342–58.
[25] Dorafshan S, Azari H. Evaluation of bridge decks with overlays using impact echo, a deep learning approach. Autom Constr 2020;113:103133.
[26] Ester M, Kriegel HP, Sander J, Xu X. Density-based spatial clustering of applications with noise. In: Proceedings of the international conference on knowledge discovery and data mining. p. 226–31.
[27] He KM, Zhang XY, Ren SQ, Sun J. Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on computer vision and pattern recognition (CVPR). Las Vegas, NV; 2016. p. 770–8.
[28] Zhao S, Shadabfar M, Zhang DM, Chen JY, Huang HW. Deep learning-based classification and instance segmentation of leakage-area and scaling images of shield tunnel linings. Struct Control Heal Monit 2021;28(6):e2732.
[29] Roncella R, Forlani G, Remondino F. Photogrammetry for geological applications: automatic retrieval of discontinuity orientation in rock slopes. Proceedings of SPIE – The international society for optical engineering, 2005.
[30] Guo JT, Wu LX, Zhang MM, Liu SJ, Sun XY. Towards automatic discontinuity trace extraction from rock mass point cloud without triangulation. Int J Rock Mech Min Sci 2018;112:226–37.
[31] Li XJ, Chen JQ, Zhu HH. A new method for automated discontinuity trace mapping on rock mass 3D surface model. Comput Geosci 2016;89:118–31.
[32] Umili G, Ferrero A, Einstein HH. A new method for automatic discontinuity traces sampling on rock mass 3D model. Comput Geosci 2013;51:182–92.
[33] Zhang P, Li JH, Yang X, Zhu HH. Semi-automatic extraction of rock discontinuities from point clouds using the ISODATA clustering algorithm and deviation from mean elevation. Int J Rock Mech Min Sci 2018;110:76–87.
[34] Slob S, van Knapen B, Hack R, Turner K, Kemeny J. Method for automated discontinuity analysis of rock slopes with three-dimensional laser scanning. Transp Res Rec 2005;1913(1):187–94.
[35] Lato M, Diederichs MS, Hutchinson DJ, Harrap R. Optimization of LiDAR scanning and processing for automated structural evaluation of discontinuities in rockmasses. Int J Rock Mech Min Sci 2009;46(1):194–9.
[36] García-Luna R, Senent S, Jurado-Piña R, Jimenez R. Structure from Motion photogrammetry to characterize underground rock masses: experiences from two real tunnels. Tunn Undergr Space Technol 2019;83:262–73.
[37] Gigli G, Casagli N. Semi-automatic extraction of rock mass structural data from high resolution LIDAR point clouds. Int J Rock Mech Min Sci 2011;48 (2):187–98.
[38] Gikas V. Three-dimensional laser scanning for geometry documentation and construction management of highway tunnels during excavation. Sensors (Basel) 2012;12(8):11249–70.
[39] Slob S, Hack H, Feng Q, Roshoff K, Turner A. Fracture mapping using 3D laser scanning techniques. Proceedings of the 11th ISRM congress. OnePetro, 2007.
[40] Vöge M, Lato MJ, Diederichs MS. Automated rockmass discontinuity mapping from 3-dimensional surface data. Eng Geol 2013;164:155–62.
[41] Slob S. Automated rock mass characterisation using 3-D terrestrial laser scanning Doctoral dissertation. Enschede: University of Twente; 2010.
[42] Cao T, Xiao AC, Wu L, Mao LG. Automatic fracture detection based on Terrestrial Laser Scanning data: a new method and case study. Comput Geosci 2017;106:209–16.
[43] Thiele ST, Grose L, Samsu A, Micklethwaite S, Vollgger SA, Cruden AR. Rapid, semi-automatic fracture and contact mapping for point clouds, images and geophysical data. Solid Earth 2017;8(6):1241–53.
[44] Girardeau-Montaut D, CloudCompare. EDF R&D Telecom ParisTech; 2016.
[45] Zhang KS, Wu W, Zhu HH, Zhang LY, Li XJ, Zhang H. A modified method of discontinuity trace mapping using three-dimensional point clouds of rock mass surfaces. J Rock Mech Geotech Eng 2020;12(3):571–86.
[46] Han XQ, Yang SM, Zhou FF, Wang J, Zhou DB. An effective approach for rock mass discontinuity extraction based on terrestrial LiDAR scanning 3D point clouds. IEEE Access 2017;5:26734–42.
[47] Wang FY, Zhou ML, Zhang DM, Huang HW, Chapman D. Random evolution of multiple cracks and associated mechanical behaviors of segmental tunnel linings using a multiscale modeling method. Tunn Undergr Space Technol 2019;90:220–30.
[48] Weidner L, Walton G, Kromer R. Generalization considerations and solutions for point cloud hillslope classifiers. Geomorphology 2020;354:107039.
[49] Zhou J, Qiu YG, Armaghani DJ, Zhang WG, Li CQ, Zhu SL, Tarinejad R. Predicting TBM penetration rate in hard rock condition: a comparative study among six XGB-based metaheuristic techniques. Geosci Front 2021;12(3):101091.
[50] Huang ZK, Pitilakis K, Argyroudis S, Tsinidis G, Zhang DM. Selection of optimal intensity measures for fragility assessment of circular tunnels in soft soil deposits. Soil Dyn Earthq Eng 2021;145:106724.
[51] Zhou J, Li XB, Shi XZ. Long-term prediction model of rockburst in underground openings using heuristic algorithms and support vector machines. Saf Sci 2012;50(4):629–44.
[52] Lary DJ, Alavi AH, Gandomi AH, Walker AL. Machine learning in geosciences and remote sensing. Geosci Front 2016;7(1):3–10.
[53] Weidner L, Walton G, Kromer R. Classification methods for point clouds in rock slope monitoring: A novel machine learning approach and comparative analysis. Eng Geol 2019;263:105326.
[54] Xiao SH, Zhang J, Ye JM, Zheng JG. Establishing region-specific N - Vs relationships through hierarchical Bayesian modeling. Eng Geol 2021;287:106105.
[55] Zhang JZ, Huang HW, Zhang DM, Zhou ML, Tang C, Liu DJ. Effect of ground surface surcharge on deformational performance of tunnel in spatially variable soil. Comput Geotech 2021;136:104229.
[56] Wei X, Zhang LL, Yang HQ, Zhang LM, Yao YP. Machine learning for pore-water pressure time-series prediction: Application of recurrent neural networks. Geosci Front 2021;12(1):453–67.
[57] Dietterich TG. Machine learning for sequential data: a review. In: Lecture notes in computer science. Berlin: Springer Berlin Heidelberg; 2002. p. 15–30.
[58] Brodu N, Lague D. 3D terrestrial lidar data classification of complex natural scenes using a multi-scale dimensionality criterion: applications in geomorphology. ISPRS J Photogramm Remote Sens 2012;68:121–34.
[59] Guo J, Liu Y, Wu L, Liu S, Yang T, Zhu W, Zhang Z. A geometry- and texture-based automatic discontinuity trace extraction method for rock mass point cloud. Int J Rock Mech Min Sci 2019;124:104132.
[60] Gallwey J, Eyre M, Coggan J. A machine learning approach for the detection of supporting rock bolts from laser scan data in an underground mine. Tunn Undergr Space Technol 2021;107:103656.
[61] O'Brien R, Ishwaran H. A random forests quantile classifier for class imbalanced data. Pattern Recogn 2019;90:232–49.
[62] Sun YM, Wong AKC, Kamel MS. Classification of imbalanced data: a review. Int J Patt Recogn Artif Intell 2009;23(4):687–719.
[63] Galar M, Fernandez A, Barrenechea E, Bustince H, Herrera F. A review on ensembles for the class imbalance problem: bagging-, boosting-, and hybrid-based approaches. IEEE Trans Syst Man Cybern Part C Appl Rev 2012;42 (4):463–84.
[64] Friedman JH. Greedy function approximation: a gradient boosting machine. Ann Statist 2001;29(5):1189–232.
[65] Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: synthetic minority over-sampling technique. J Artificial Intell Res 2002;16:321–57.
[66] Bergstra J, Yamins D, Cox D. Hyperopt: a python library for optimizing the hyperparameters of machine learning algorithms. In: Proceedings of the 12th python in science conference. Austin, Texas; 2013. p. 20.
[67] Breiman L. Random forests. Mach Learn 2001;45(1):5–32.
[68] Safavian SR, Landgrebe D. A survey of decision tree classifier methodology. IEEE Trans Syst Man Cybern 1991;21(3):660–74.

[69] Seiffert U. Multiple layer perceptron training using genetic algorithms. In: Proceedings of the European symposium on artificial neural networks. Bruges; 2001. p. 159–64.
[70] Westoby MJ, Brasington J, Glasser NF, Hambrey MJ, Reynolds JM. 'Structure-from-Motion' photogrammetry: a low-cost, effective tool for geoscience applications. Geomorphology 2012;179:300–14.
[71] Llc A. Agisoft PhotoScan user manual-professional edition. St. Petersburg: Agisoft LLC; 2016.
[72] Behley J, Steinhage V, Cremers AB. Efficient radius neighbor search in three-dimensional point clouds. In: Proceedings of the 2015 IEEE international conference on robotics and automation (ICRA). Seattle, WA; 2015. p. 3625–30.
[73] Schnabel R, Klein R. Octree-based point-cloud compression. In: Proceedings of the 3rd eurographics/IEEE VGTC conference on Point-Based Graphics. p. 111–20.
[74] Wang XC, Wang XL, Ma YQ, Wilkes DM. A fast MST-inspired kNN-based outlier detection method. Inf Syst 2015;48:89–112.
[75] Arthur D, Vassilvitskii S. k-means++: the advantages of careful seeding. Stanford; 2006.
[76] Rousseeuw PJ. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. J Comput Appl Math 1987;20:53–65.
[77] Khan MMR, Siddique MAB, Arif RB, Oishe MR. ADBSCAN: adaptive density-based spatial clustering of applications with noise for identifying clusters with varying densities. In: Proceedings of the 4th international conference on electrical engineering and information & communication technology (iCEEiCT). Dhaka, Bangladesh; 2018. p. 107–11.
[78] Hastie T, Tibshirani R, Friedman J. The elements of statistical learning: data mining, inference, and prediction. Springer Science & Business Media; 2013.
[79] Gordon AD, Breiman L, Friedman JH, Olshen RA, Stone CJ. Classification and regression trees. Biometrics 1984;40(3):874.
[80] Lewis RJ. An introduction to classification and regression tree (CART) analysis. In: Proceedings of the annual meeting of the society for academic emergency medicine. San Francisco, California; 2000.
[81] James G, Witten D, Hastie T, Tibshirani R. Tree-based methods. Springer Texts in Statistics. New York, NY: Springer New York, 2013. p. 303–35.
[82] Yeh CC, Lin FY, Hsu CY. A hybrid KMV model, random forests and rough set theory approach for credit rating. Knowl – Based Syst 2012;33:166–72.
[83] Panchal G, Ganatra A, Kosta YP, Panchal D. Behaviour analysis of multilayer perceptronswith multiple hidden neurons and hidden layers. Int J Comput Theory Eng 2011:332–7.
[84] Zhu TF, Lin YP, Liu YH. Synthetic minority oversampling technique for multiclass imbalance problems. Pattern Recogn 2017;72:327–40.
[85] Geetha R, Sivasubramanian S, Kaliappan M, Vimal S, Annamalai S. Cervical cancer identification with synthetic minority oversampling technique and PCA analysis using random forest classifier. J Med Syst 2019;43(9):1–19.
[86] Bergstra J, Bengio Y. Random search for hyper-parameter optimization. J Mach Learn Res 2012;13(1):281–305.
[87] Stone M. Cross-validatory choice and assessment of statistical predictions. J Royal Stat Soc: Ser B Methodol 1974;36(2):111–33.
[88] Weinmann M, Jutzi B, Hinz S, Mallet C. Semantic point cloud interpretation based on optimal neighborhoods, relevant features and efficient classifiers. ISPRS J Photogramm Remote Sens 2015;105:286–304.
[89] Hapfelmeier A, Hothorn T, Ulm K, Strobl C. A new variable importance measure for random forests with missing data. Stat Comput 2014;24(1):21–34.
[90] Bond CE, Shipton ZK, Jones R, Butler RWH, Gibbs AD. Knowledge transfer in a digital world: Field data acquisition, uncertainty, visualization, and data management. Geosphere 2007;3(6):568–76.
[91] Becker C, Rosinskaya E, Häni N, d'Angelo E, Strecha C. Classification of aerial photogrammetric 3D point clouds. Photogramm Eng Remote Sensing 2018;84(5):287–95.
[92] Chen JY, Zhang DM, Huang HW, Shadabfar M, Zhou ML, Yang TJ. Image-based segmentation and quantification of weak interlayers in rock tunnel face via deep learning. Autom Constr 2020;120:103371.