Advances in Geo-Energy Research

Vol. 15, No. 2, p. 158-171, 2025

Original article

Integration of image recognition and expert system for real-time wellbore stability analysis

Yongdong Fan[1,2], Huiwen Pang[1,2,3], Yan Jin[1,2]*, Han Meng[1], Yunhu Lu[2], Shiming Wei[2,3], Hanqing Wang[4,5]

[1]*College of Artificial Intelligence, China University of Petroleum, Beijing 102249, P. R. China*

[2]*State Key Laboratory of Petroleum Resources and Prospecting, China University of Petroleum, Beijing 102249, P. R. China*

[3]*College of Science, China University of Petroleum, Beijing 102249, P. R. China*

[4]*Resource Geophysics Academy, Imperial College London, London SW7 2BP, United Kingdom*

[5]*Petroleum Exploration and Production Research Institute, SINOPEC, Beijing 102206, P. R. China*

Keywords:
Image recognition
wellbore stability
expert system
oil and gas extraction
drilling cuttings

Cited as:
Fan, Y., Pang, H., Jin, Y., Meng, H., Lu, Y., Wei, S., Wang, H. Integration of image recognition and expert system for real-time wellbore stability analysis. Advances in Geo-Energy Research, 2025, 15(2): 158-171.
https://doi.org/10.46690/ager.2025.02.07

Abstract:
Wellbore stability is a key factor affecting safe and efficient drilling. At present, it is difficult to conduct real-time and accurate analysis of wellbore stability in related research. To address the current research shortcomings, this study proposes a real-time analysis model of wellbore stability integrating image recognition and an expert system, which mainly includes caving image segmentation and recognition, and a wellbore stability expert system. The caving image recognition proposes a new dynamic threshold segmentation method based on simple linear iterative clustering superpixel segmentation and visual geometry group 19-layer image classification. After completing the segmentation of the caving image, the geometric features of the caving are calculated, and the multi-source feature fusion GoogleNet model is established by integrating the geometric features with the convolution features extracted by GoogleNet to identify the caving types efficiently. After segmentation and recognition of caving images. The wellbore stability expert system uses the caving features to establish an expert system model to determine the mechanism of wellbore instability and provide reasonable solutions. Finally, the wellbore stability integrating image recognition and an expert system model was applied to a well in field production, accurately determining the mechanism of wellbore instability in real time and effectively solving the corresponding wellbore instability problem based on the measures provided by the model.

1. Introduction

With the development of related technologies, the extraction of unconventional oil and gas resources such as shale oil and gas, tight oil and gas, and deep-water oil and gas has received widespread attention. The extraction of unconventional oil and gas resources is challenging and has a low recovery rate. Complex well structures like extended-reach wells and horizontal wells are widely utilized to achieve efficient extraction of these resources (Zou et al., 2015). The risk of wellbore instability in complex wells is often higher than in vertical wells, thus higher requirements for wellbore stability are necessary (Ma et al., 2015). The most common form of wellbore instability is the collapse of wellbore rock, which typically produces cavings that can further lead to incidents such as the sticking of tools (Xu et al., 2024).

Experts and scholars in this field have conducted long-term

Yandy Scientific Press

*Corresponding author.
E-mail address: fanyongdong0912@163.com (Y. Fan); phw9999@126.com (H. Pang); jiny@cup.edu.cn (Y. Jin); h.meng94@outlook.com (H. Meng); luyh@cup.edu.cn (Y. Lu); we_shiming@163.com (S. Wei); h.wang@imperial.ac.uk (H. Wang).

Received November 19, 2024; revised December 13, 2024; accepted January 6, 2025; available online January 12, 2025.

research and established analytical models, empirical models, and numerical simulation methods to quantitatively characterize the risk of wellbore instability (Chen et al., 2002; Yuan et al., 2013; Ayoub et al., 2019; Gong et al., 2022). These existing models mainly focus on constitutive models that describe the stress-strain relationship and failure criteria that describe the ultimate state of rock failure. Constitutive models include linear elastic models, elastoplastic models, poroelastic models, and chemo poroelastic models, which consider the effects of different conditions and environments on rock deformation to establish the mapping relationship between stress and strain (John et al., 2014; Gao et al., 2022; Peng et al., 2023). The main failure criteria include the Mohr-Coulomb criterion, the Mogi-Coulomb criterion, the Hoek-Brown criterion, and the Drucker-Prager criterion. These criteria describe the critical failure state of rocks under different conditions (Al-Ajmi and Zimmerman, 2006; Zoback, 2007; Shen et al., 2019). While these constitutive models and failure criteria can effectively describe the stress-strain relationship and failure strength of rocks to a certain extent, they mostly rely on logging data and experimental results, thus having hysteresis and being unable to perform real-time analysis. For drilling engineering, it is particularly important to analyze the risk of wellbore instability and make real-time adjustments to reduce the risk of instability.

During drilling, cavings are returned to the surface in real-time with the mud. Since cavings are generated due to wellbore instability and the detachment of surrounding rock, the shape and surface changes of the cavings can reflect the type and mechanism of wellbore instability in real-time (Edwards et al., 2004; Skea et al., 2018). Patel et al. (2018) proposed a real-time analysis of formation pore pressure by analyzing the shape and quality of downhole cavings. Purkayastha et al. (2020) analyzed the morphological characteristics of the cavings and analyzed the collapse of the wellbore, ultimately providing corresponding remedial measures. In recent years, some experts and scholars have used machine learning methods to identify the types of cavings, in order to determine the mechanism of wellbore instability and provide corresponding solutions based on the mechanism of wellbore instability (Izurieta et al., 2019; Jin et al., 2022).

Analyzing wellbore stability through caving images is typical image recognition research. Research on image recognition can be divided into two categories from the perspective of image feature extraction. The first category calculates the features of the image regions through image segmentation to identify the type of images (Purswani et al., 2020; Song et al., 2022; Zhai et al., 2024). The second category uses convolutional neural networks to directly identify the type of images (He et al., 2019; Houshmand et al., 2022; Gupta et al., 2024). Cuttings recognition can use convolutional neural network to directly recognize cutting images, such as effectively recognizing cuttings images using the combination of visual geometry group 16-layer (VGG16) and transfer learning (Wang, 2022). For some complex cutting images, direct recognition is difficult. Image segmentation technology can be used to segment the cuttings and calculate their features for final recognition, such as segmenting original cuttings images using the watershed algorithm, calculating the features of the segmented images, and finally identifying the lithology of the cuttings through feature similarity matching (Huo et al., 2021).

The convolutional neural networks combine feature extraction and classification functions and are widely used in image recognition fields, such as visual geometry group 19-layer (VGG19), ResNet50, MobileNet, and GoogleNet (Christian et al., 2014; Simonyan and Zisserman, 2015; Howard et al., 2017; Ren et al., 2023). The VGG19 uses smaller convolutional kernels, making it easier to extract local features from the image. Additionally, VGG19 increases the receptive field by stacking multiple convolutional layers, allowing for the extraction of global features from the image (Simonyan and Zisserman, 2015). Therefore, VGG19 is widely used in image recognition research where local features are prominent and the dataset is large (Favorskaya and Pakhirka, 2019; Rill-García et al., 2022; Mishra et al., 2024). VGG19 is highly sensitive to local features of the image, while GoogleNet, by introducing the Inception module, excels at handling multi-scale features (Christian et al., 2014). In recent years, GoogleNet has shown excellent performance in image recognition tasks with complex backgrounds and multiple scales of image features (Wang et al., 2019; Khan et al., 2021; Bezabh et al., 2024).

The segmentation and recognition of cavings images aim to better analyze wellbore instability. Based on cavings segmentation and recognition, an expert system for wellbore instability can be established by combining other cavings information, enabling real-time analysis of wellbore instability. Current expert systems can be divided into rule-based expert systems, frame-based expert systems, and neural network-based expert systems (Stahl et al., 2015; Paradarami et al., 2017; Shishehchi and Banihashem, 2021; Chen et al., 2024). Rule-based expert systems provide solutions to problems by collecting domain knowledge, designing rule bases, and using inference mechanisms (Yang et al., 2017; Kolodziejczyk et al., 2022). Belief rule-based expert systems are a widely applied type of rule-based expert systems. In belief rule-based expert systems, the conditions and conclusions of rules can be uncertain, and this uncertainty can be represented and handled through probability, fuzzy logic, and other methods (Xu et al., 2007, 2018; Yang et al., 2023).

To achieve real-time analysis of wellbore stability, this study proposes a model of wellbore stability integrating image recognition and an expert system. This model mainly includes the recognition of caving images and a wellbore stability expert system based on caving characteristics. For the caving image recognition, this study introduces a multi-source feature fusion GoogleNet model. The multi-source feature fusion GoogleNet model segments caving images and extracts geometric features using the improved threshold segmentation model proposed in this study. Finally, it fuses the geometric features with convolutional features to recognize the cavings. The expert system combines the cavings recognition results with surface roughness and edge angle changes of the cavings to establish a rule-based wellbore stability expert system model.

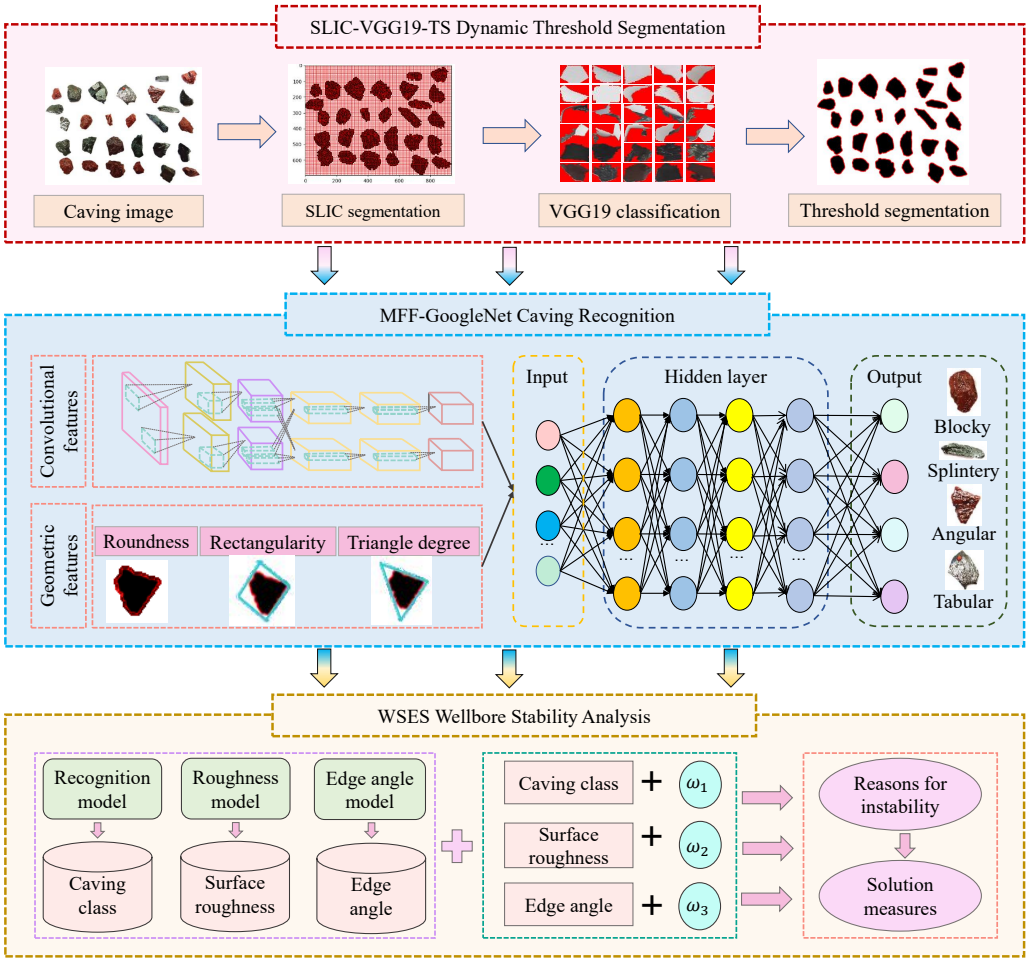

Fig. 1. WSIRES model.

2. Methodology

Imaging logging and caliper logging data can be used to analyze wellbore collapse, but such data is mostly obtained after drilling, making real-time analysis of wellbore collapse through these data impossible. Cavings are downhole data that are returned with the mud in real time during drilling. Therefore, if the mapping relationship between cavings and wellbore collapse can be established, real-time analysis of wellbore collapse can be realized. This research uses images of cavings as the dataset and proposes the real-time analysis model of wellbore stability integrating image recognition and an expert system (WSIRES). The WSIRES model primarily includes an improved threshold segmentation model, multi-source feature fusion GoogleNet image recognition model, and wellbore stability expert system (WSES) (Fig. 1).

2.1 Improved threshold segmentation model

To achieve effective segmentation and extraction of blocky cuttings, this study proposes an improved thresholding segmentation method, named SLIC-VGG19-TS. The SLIC-VGG19-TS model includes simple linear iterative clustering superpixel segmentation (SLIC), The visual geometry group 19-layer (VGG19) superpixel region classification, and threshold segmentation (TS). The SLIC superpixel segmentation segments the original caving images into multiple superpixel regions based on image features. The VGG19 superpixel region classification judges the superpixel regions obtained from SLIC segmentation to identify the superpixel regions containing cavings. The TS threshold segmentation performs threshold segmentation on the superpixel regions containing cavings, and different superpixel regions correspond to different segmentation thresholds, thus making the threshold segmentation dynamic.

2.1.1 SLIC superpixel segmentation

The core of SLIC is to segment the image into tightly adjacent and similar superpixel regions rather than independent pixel points. Such superpixel regions can better represent the structure and texture information of the image in image analysis. The SLIC algorithm segments N pixel points in the original image into k superpixel regions based on a clustering algorithm. During initialization, the number of superpixels k and the compactness parameter m need to be specified. The parameter m controls the shape of the superpixels. The larger the value of m, the more compact and regular the superpixel regions become. The smaller the value of m, the more irregular the superpixel regions become. After specifying

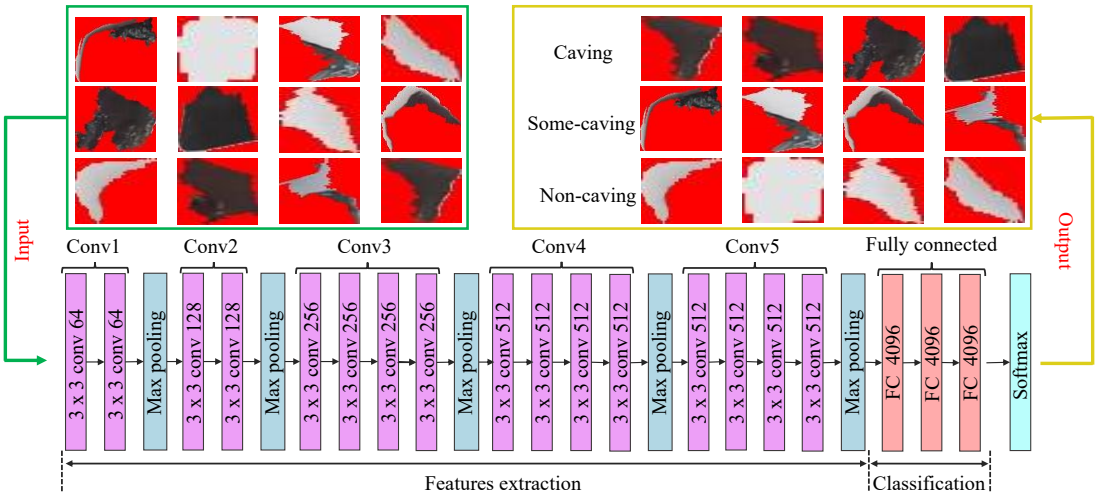

Fig. 2. VGG19 network architecture.

the initial parameters, multiple iterations of optimization are required. During iterative optimization, the distance between each pixel and the surrounding superpixels is calculated. If the calculated distance is less than the current minimum distance, the minimum distance is updated, and the pixel is assigned to the corresponding superpixel. The SLIC distance consists of color distance and spatial distance. When calculating the color distance, the original RGB color image needs to be converted to a Lab color image, and the color distance should be calculated based on the Lab color image:

$$d_c(i,j) = \sqrt{(l_j - l_i)^2 + (a_j - a_i)^2 + (b_j - b_i)^2} \quad (1)$$

where l_j, a_j, and b_j are the Lab color values of the superpixel center; l_i, a_i, and b_i are the Lab color values of pixels; d_c is the color distance.

Spatial distance is the distance between pixels calculated based on the pixel coordinates of an image:

$$d_s(i,j) = \sqrt{(x_j - x_i)^2 + (y_j - y_i)^2} \quad (2)$$

where x_j and y_j are the coordinates of the superpixel center; x_i and y_i are the coordinates of pixels; d_s is the spatial distance.

By integrating color distance and spatial distance, combined with the compactness and number of superpixels in SLIC superpixel segmentation, the comprehensive distance of SLIC can be obtained:

$$d = \sqrt{\left(\frac{d_c}{m}\right)^2 + \left(\frac{d_s}{\sqrt{\frac{N}{k}}}\right)^2} \quad (3)$$

where d is the distance between the pixel point and the center of the surrounding superpixels; m and k are the compactness and the number of superpixels, respectively; N is the number of pixel points.

After calculating the distance in each iteration, the pixels are readjusted, the superpixel center is updated, and the segmentation is completed after reaching the termination condition. After completing the segmentation, continue to calculate the similarity between adjacent superpixels, merge adjacent and similar superpixels, and obtain the final result.

2.1.2 VGG19 classification and threshold segmentation

After SLIC segmentation, the original image is divided into many superpixel regions. These superpixel regions can be categorized into three types: Superpixels containing only cavings (cavings superpixels), superpixels containing only non-cavings areas (non-cavings superpixels), and superpixels containing both cavings and non-cavings areas (some-cavings superpixels). These three types of superpixel regions are processed differently in the final cavings extraction. Cavings superpixels are assigned black, non-cavings superpixels are assigned white, and some-cavings superpixels require threshold segmentation to assign black to the cavings areas and white to the non-cavings areas (Fig. 2). Therefore, it is necessary to classify all superpixel regions accordingly.

Due to its simple yet powerful architecture, VGG19 has been widely used in the field of image classification in recent years. This study employs VGG19 for the classification of superpixel regions. The VGG19 includes 16 convolutional layers, 5 max-pooling layers, and 3 fully connected layers. Each convolutional layer uses 3×3 convolutional kernels to increase the depth of the network while keeping the receptive field unchanged. Each pooling layer uses 2×2 kernels, which can effectively aggregate local features while reducing computational complexity. After inputting the extracted features into the fully connected layer, three fully connected layers can effectively classify the image (Fig. 2). This study uses VGG19 to identify the superpixel regions segmented by SLIC, and finally uses threshold segmentation to segment some-cavings superpixel regions to achieve image segmentation with cavings.

2.2 Multi-source feature fusion GoogleNet model

Convolutional neural networks can classify the original images of cavings, but they often encounter recognition errors with cavings that have similar shapes. To address this issue, a multi-source feature fusion GoogleNet model (MMF-GoogleNet) based on the GoogleNet network is established for

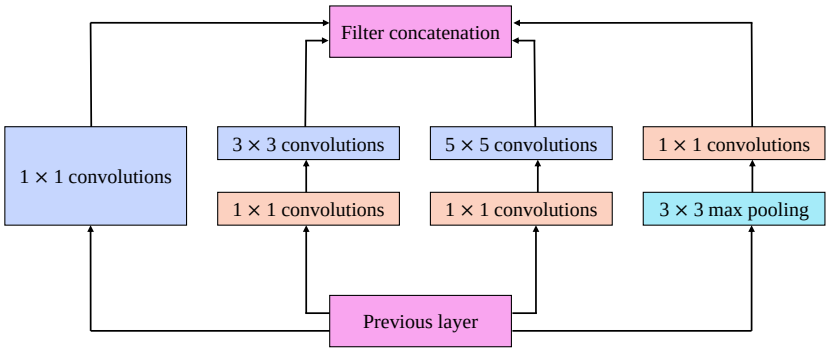

Fig. 3. Inception architecture.

caving type recognition. The multi-source features include the geometric shape features calculated from the cavings images segmented by the SLIC-VGG19-TS model and the convolutional features extracted by GoogleNet. Finally, these geometric features and convolutional features are fused through a fully connected layer to improve the accuracy of caving type recognition (Fig. 1).

2.2.1 Geometric feature extraction of cavings

Based on the mechanism of wellbore instability, the debris can be classified into tabular cavings, angular cavings, blocky cavings, and splintery cavings. The geometric features of cavings are important factors in distinguishing their types. In previous studies, roundness and rectangularity have been effective in describing the extent to which a 2D image approximates a circle and a rectangle, respectively:

$$C = \frac{4\pi A}{p^2} \tag{4}$$

$$R = \frac{A}{S_R} \tag{5}$$

where C is roundness, R is rectangularity, A is the area of the caving, p is the perimeter of the caving, S_R is the minimum bounding rectangle area of the caving.

Rectangularity and roundness can effectively distinguish tabular and blocky cavings from the other two types of cavings. However, these two geometric features cannot describe the extent to which a 2D image approximates a triangle. Triangularity can effectively distinguish angular cavings from the other three types of cavings. Therefore, this study introduces the concept of triangularity, based on the definitions of roundness and rectangularity:

$$T = \frac{A}{S_T} \tag{6}$$

where T is triangularity, S_T is the minimum bounding triangle area of the caving.

2.2.2 GoogleNet model

GoogleNet utilizes the Inception module in its deep neural network, which enhances feature-capturing capabilities by simultaneously using convolutional kernels and pooling at different scales. The Inception module in GoogleNet performs multi-scale feature extraction, where the 1×1 convolution is primarily used for linearly combining input channels, while the 3×3 and 5×5 convolutions capture features over larger receptive fields. The 1×1 convolution in the module not only extracts features but also effectively reduces the number of input channels, thus reducing computational resource consumption (Fig. 3).

2.2.3 Multi-source feature fusion

When merging geometric features with convolutional features extracted by GoogleNet, it is necessary to combine these features through a fully connected layer (Fig. 1). Geometric features are processed through a fully connected layer to obtain a matrix representing geometric features:

$$h_g = f(\omega_g g + b_g) \tag{7}$$

where g is geometric features, ω_g and b_g are the corresponding weights and biases for geometric features, f is the activation function.

Convolutional features are input into another fully connected layer to obtain a matrix representing convolutional features:

$$h_c = f(\omega_c c + b_c) \tag{8}$$

where c is convolutional features, ω_c and b_c are the corresponding weights and biases for convolutional features, h_c is the matrix of convolutional features after the fully connected layer operation.

Finally, different weights are assigned to geometric and convolutional features, and feature fusion is performed based on weighted averaging:

$$h = \alpha h_g + (1 - \alpha) h_c \tag{9}$$

where α is the weight for geometric features.

The weights are determined by the recognition accuracies of geometric and convolutional features:

$$\alpha = \frac{p_g}{p_g + p_c} \tag{10}$$

where p_g is the accuracy when using geometric features alone for recognition, p_c is the accuracy when using convolutional features alone for recognition.

The fused features are then input into another fully connected layer to obtain the cavings recognition results:

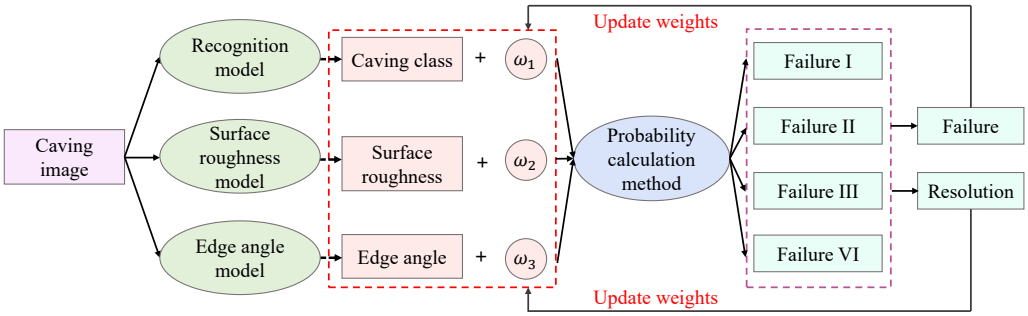

Fig. 4. Wellbore stability expert system.

$$o = f(\omega_m h + b_m) \tag{11}$$

where h is the fused features after merging, ω_m and b_m are the corresponding weights and biases for the merged features, and o is the final output.

2.3 WSES model

This study establishes the WSES model based on caving type, surface roughness, and edge angles, integrating years of domain expertise. Through the computation of these three features, the model determines the mechanisms of wellbore instability and proposes corresponding solutions. Edge angle calculations utilize the cartToPolar module from OpenCV, which first transforms pixel coordinates from Cartesian to polar coordinates, and then computes the magnitude and direction of pixel points. Firstly, it is necessary to calculate the gradient values of pixels in both horizontal and vertical directions:

$$G_x = \boldsymbol{I} \cdot \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \tag{12}$$

$$G_y = \boldsymbol{I} \cdot \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \tag{13}$$

where G_x is the gradient value of pixels in the horizontal direction, G_y is the gradient value of pixels in the vertical direction, $\boldsymbol{I}$ is the matrix of the input image.

After calculating the gradient value of the pixel, calculate the pixel coordinates in polar coordinate system based on the gradient values:

$$r = \sqrt{G_x^2 + G_y^2} \tag{14}$$

$$\theta = \arctan \frac{G_y}{G_x} \tag{15}$$

where θ is the direction of pixels in the polar coordinate system, r is the magnitude of pixels in the polar coordinate system.

Surface roughness calculations rely on the Sobel module in Python, which calculates the roughness of the image based on gradient changes in different directions.

The WSES model is based on existing rule-based expert systems and is improved by introducing feature weights and iterative updates (Fig. 4). It integrates real-time production knowledge to update these weights dynamically, thereby enhancing decision accuracy. During the weight updating process, decisions are made based on three features, and the accuracy of these decisions is evaluated with field feedback. Each feature is used individually to make decisions based on the expert system, and the accuracy of decisions made using each feature is accumulated. If the accuracy of decisions made using a particular feature is higher, the weight corresponding to that feature will be greater. The calculation of specific feature weights can be implemented using:

$$\omega_i = \frac{\frac{n_i}{N_i}}{\sum_{i=1}^{3} \frac{n_i}{N_i}} \tag{16}$$

where N_i is the total number of decisions made based on feature i, n_i is the total number of correct decisions made based on feature i, and ω_i is the weight assigned to feature i.

3. Multi-source feature fusion for caving recognition

The MFF-GoogleNet model proposed in this study effectively recognizes caving types through multi-feature fusion. Caving type recognition serves as a crucial basis for analyzing the causes of wellbore instability in the WSES model. This section primarily includes the description and analysis of the dataset, caving extraction using SLIC-VGG19-TS, and caving recognition using MFF-GoogleNet. SLIC-VGG19-TS can effectively segment the cavings areas, making it easier to calculate three geometric features: Roundness, rectangularity, and triangularity. Integrating these geometric features with the convolutional features extracted by GoogleNet can partially compensate for the issue of insufficient samples in cavings recognition by GoogleNet, thereby improving the accuracy of cavings recognition.

3.1 Data set description

This study collected a total of 517 samples of cavings, which were gathered through photographic means. Typically, each photograph contains multiple cavings (Fig. 5(a)). The im-

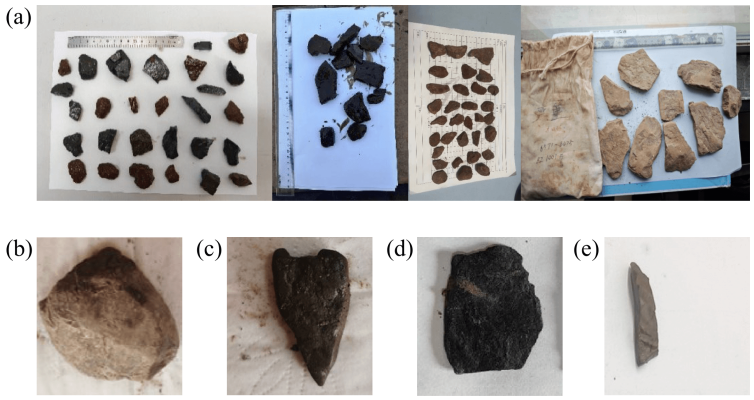

Fig. 5. Caving dataset. (a) Caving image, (b) blocky caving, (c) angular caving, (d) tabular caving and (e) splintery caving.

Table 1. Caving image dataset.

Dataset	Splintery	Angular	Tabular	Blocky
Origin	91	107	186	133
Transformation	273	321	558	399

Table 2. The main hyperparameters of VGG19 and GoogleNet.

Model	Learning rate	Kernel size	Pooling size
VGG19	0.001	(3,3)	(2,2)
GoogleNet	0.0005	(1,1) (3,3) (5,5)	(3,3)

ages of cavings taken from different angles may show some variations. To more accurately determine the type of cuttings, the studychose to position the flattest and larger side of the cavings facing downward during the photography. Based on the shape characteristics, cavings can be categorized into four main types: Blocky caving, angular caving, tabular caving, and splintery caving (Fig. 5). Each type corresponds to different collapse mechanisms. Blocky cavings appear as rough-surfaced blocks, angular cavings present as rough triangular or arrow-shaped forms, tabular cavings show smooth-surfaced plates and splintery cavings appear as concave elongated pieces. Among these types, angular cavings and splintery cavings differ significantly from others in shape, while tabular cavings and blocky cavings often exhibit smaller differences.

The 517 cavings images are segmented into 68,932 superpixel regions using SLIC superpixel segmentation. These superpixel regions are then classified using VGG19. After classification, thresholding segmentation is applied to different types of superpixel regions to obtain the final cavings areas. Therefore, during VGG19 image classification, 68,932 segmented images are used.

The partitioning of the dataset is particularly important for convolutional neural networks. Considering that the GoogleNet model requires only 517 images, one of the following operations-rotation, mirroring, or adding noise-was applied to each image. As a result, the sample dataset was expanded to 1,551 images (Table 1). Of these, 80% were used as the training data and 20% as the test data. For VGG19, the images used are superpixels generated by SLIC segmentation, and since the sample size is sufficient, 80% of the images were directly used as the training dataset, with the remaining 20% used as the test dataset. The hyperparameters of the models were primarily determined through a grid search optimization algorithm. The main hyperparameters of the models obtained after training VGG19 and GoogleNet are shown in Table 2.

3.2 Caving extraction based on image segmentation

Cavings extraction is crucial for calculating geometric features and obtaining surface roughness and edge angles. This study proposes the SLIC-VGG19-TS image segmentation model, comprising SLIC superpixel segmentation, VGG19 superpixel classification, and dynamic threshold segmentation. In SLIC superpixel segmentation, the number of superpixels k and compactness m significantly affect segmentation effectiveness (Eq. (3)). The parameter m primarily determines the regularity of superpixel regions; smaller m values result in more irregular superpixel shapes. Within a suitable range, variations in m have less impact on the final segmentation results. Given the irregular nature of cavings, smaller values of m are preferable. In contrast, the choice of k varies widely, with different k values exerting a greater influence on segmentation outcomes. As depicted in Fig. 6 with 15 cavings examples, varying m and k settings illustrate the segmentation effects of SLIC. Compactness m and the number of superpixels k directly influence the segmentation outcome, thereby affecting the final classification results of VGG19. Through multiple trials, it has been found that when m and k are set to 8 and 2,000 respectively, the VGG19 achieves the highest classification accuracy (Fig. 6(d)). Therefore, it is finally decided that $m = 8$ and $k = 2,000$ (Fig. 6).

For the three types of superpixels, different processing methods are required, necessitating the classification of all superpixels obtained from SLIC segmentation (Fig. 2). The superpixel regions obtained from SLIC are irregular two-dimensional images, whereas VGG19 image classification

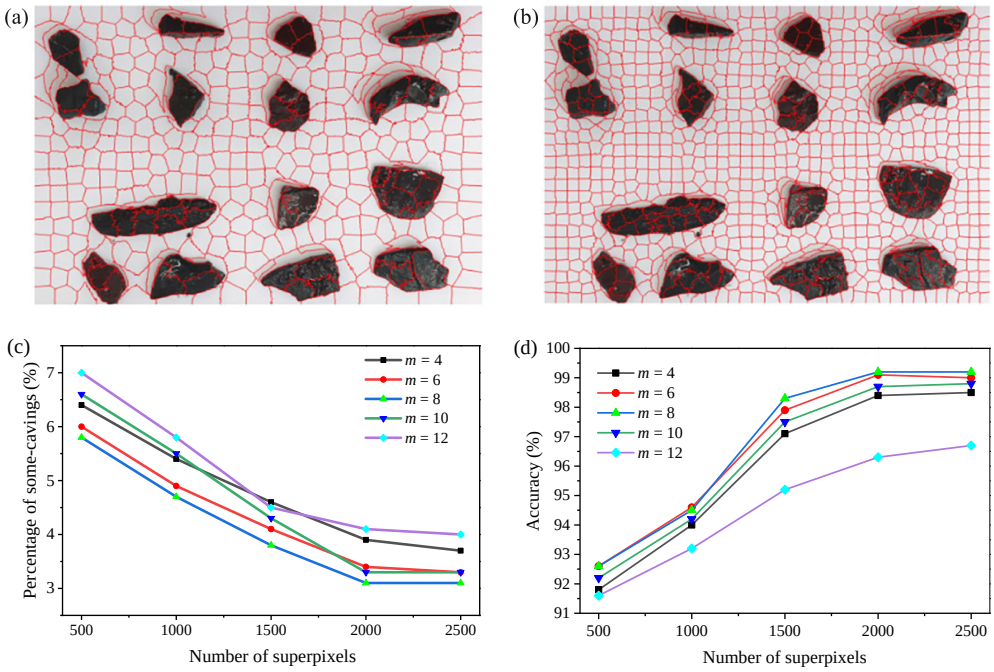

Fig. 6. Determining key parameters for SLIC superpixel segmentation. (a) $m = 8$ and $k = 500$, (b) $m = 8$ and $k = 1,000$, (c) percentage of some-savings superpixels after SLIC segmentation and (d) VGG19 recognition accuracy.

requires input in standard rectangular format. Therefore, in this study, the non-superpixel regions of the SLIC segmented superpixel images are filled with red color (Fig. 2). Red color was chosen for filling non-superpixel regions because the research found that pure red is not present in any of the non-caving areas, which helps reduce background interference after filling.

The extracted cavings closely resemble the original caving shapes, maintaining their fundamental characteristics (Fig. 7). Subsequent calculations of geometric features are primarily based on these extracted caving images. This study demonstrates that the SLIC-VGG19-TS segmentation and extraction of cavings meet the research requirements effectively.

3.3 Cavings recognition

In the recognition of caving shapes, this study proposes the MFF-GoogleNet model, which integrates two main types of features: Convolutional features and geometric features. Convolutional features are extracted by GoogleNet through convolution and pooling operations on the original caving images. Geometric features are extracted based on binary images of cavings segmented by SLIC-VGG19-TS. Geometric features mainly include roundness, rectangularity, and triangularity of cavings. Analysis of the relationship between caving types and geometric features in the sample dataset reveals that triangularity tends to be higher for angular cavings, rectangularity is generally higher for tabular cavings, and all three geometric features are smaller for splintery cavings (Fig. 8). This indicates a direct correlation between geometric features and caving types.

Geometric features are closely related to caving types, but alone they are insufficient to determine the type of cavings. Therefore, in MMF-GoogleNet, convolutional features extracted by GoogleNet are also integrated. GoogleNet extracts convolutional features from the input raw caving images through convolution and pooling operations, resulting in convolutional feature maps (Fig. 9). The effectiveness of feature extraction through convolution significantly impacts the final classification performance of the model. Taking the first convolutional layer of GoogleNet as an example, which has 64 output channels, it can be viewed as producing 64 feature maps (Fig. 9). Each feature map from every convolutional layer captures partial features of the image, and no single feature map can describe all features of the image. These feature maps collectively represent different aspects of the caving image. After multiple convolution and pooling operations, the final pooling layer converts the feature maps into feature matrices of a fixed size. In GoogleNet, this results in 1,024 feature matrices, each with spatial dimensions of 7×7 (Fig. 9). These feature matrices are then fused with the feature matrices obtained from geometric feature extraction to form new feature matrices (Eqs. (7)-(11)). These new feature matrices are input into fully connected layers for the final recognition of caving types.

MMF-GoogleNet achieved an overall recognition accuracy of 94.97% on the test dataset. Among the different types of cavings, angular cavings had the highest recognition accuracy, followed by splintery cavings. However, recognition accuracy was lower for blocky and tabular cavings (Fig. 10(a)). The lower accuracy for blocky and tabular cavings is primarily due to their similar appearance in images, leading to potential misidentification. Specifically, 80.5% of misidentified blocky

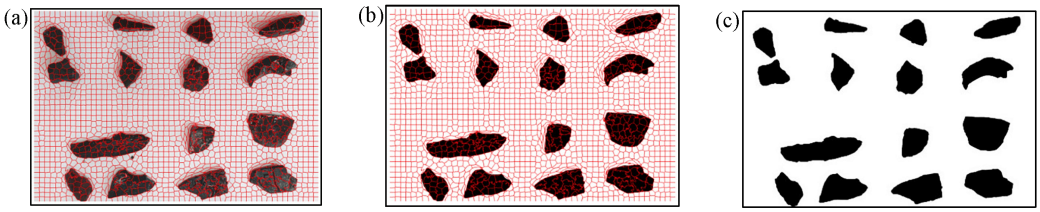

Fig. 7. SLIC-VGG19-TS caving extraction. (a) SLIC, (b) SLIC-VGG19-TS and (c) extraction results.

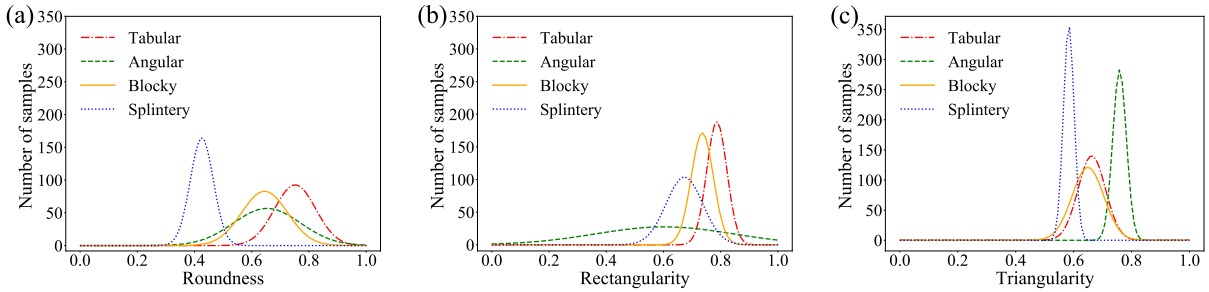

Fig. 8. Types of cavings and Gaussian distribution of geometric features. (a) Roundness of cavings, (b) rectangularity of cavings and (c) triangularity of cavings.

cavings were mistakenly classified as tabular cavings, while 62.5% of misidentified tabular cavings were incorrectly identified as blocky cavings (Fig. 10(b)). Tabular cavings and blocky cavings are prone to mutual recognition errors due to their similar appearances. Additionally, after drilling fluid infiltrates formations with well-developed fractures and bedding, and it is likely to form tabular and blocky cavings. The mechanisms of wellbore instability for both types of cavings are quite similar. The lower recognition accuracy of blocky cavings compared to tabular cavings is because there are more tabular cavings in the dataset (Table 1). During model training, it is easier for the model to learn the features of tabular cavings, which is why the recognition accuracy of blocky cavings is lower than that of tabular cavings. Angular cavings and splintery cavings, characterized by their unique structural features, had higher recognition accuracy, with a lower likelihood of being misidentified than other types of cavings.

To validate the performance of the MMF-GoogleNet caving recognition, it was compared against several other models: The threshold segmentation-based geometric feature model (TS-GF), the SLIC-VGG19-TS segmentation recognition model based on geometric features (SLIC-VGG19-GF), the GoogleNet model based on convolutional features alone, and the threshold segmentation improved GoogleNet model based on multiple-sources features (TS-GoogleNet). The results showed that both TS-GF and SLIC-VGG19-GF, based solely on geometric features, had significantly lower recognition accuracy compared to the other three models. SLIC-VGG19-GF had slightly higher accuracy than TS-GF, indicating that SLIC-VGG19-TS segmentation performs better than threshold segmentation for caving segmentation. Both TS-GoogleNet and MMF-GoogleNet, which integrate both geometric and convolutional features, exhibited higher recognition accuracy than the other three models. This demonstrates that the fusion of geometric and convolutional features effectively enhances the model's recognition accuracy. Among the five models evaluated, MMF-GoogleNet consistently achieved the highest recognition accuracy for all types of cavings, indicating that the proposed model improvement approach in this study is of practical value (Fig. 11).

4. Analysis of wellbore instability based on expert systems

4.1 The relationship between cavings and wellbore failure

Cavings are a type of drilling cuttings, where conventional cuttings are generated by the drill bit damaging the rock at the bottom of the wellbore, whereas cavings are larger cuttings that fall due to damage to the wellbore, often larger in size compared to normal cuttings. According to the mechanism of wellbore instability and the shapes of the cavings, cavings can be classified into four types: Blocky caving, tabular caving, angular caving, and splintery caving (Skea et al., 2018; Jin et al., 2022). The type of cavings can indicate the mechanism of their formation, thereby revealing the mechanism of wellbore damage. During construction, if real-time analysis of the type of cuttings and analysis of the wellbore instability mechanism can be performed based on the cavings returned from the bottom of the well, appropriate solutions can be implemented accordingly. This approach helps in reducing the risk of wellbore instability and drilling costs. While the type of cavings can indicate the mechanism of wellbore instability, this determination is not always absolutely accurate. To enhance the accuracy of these determinations, this study introduces a WSES model. The WSES model mainly focuses on the type

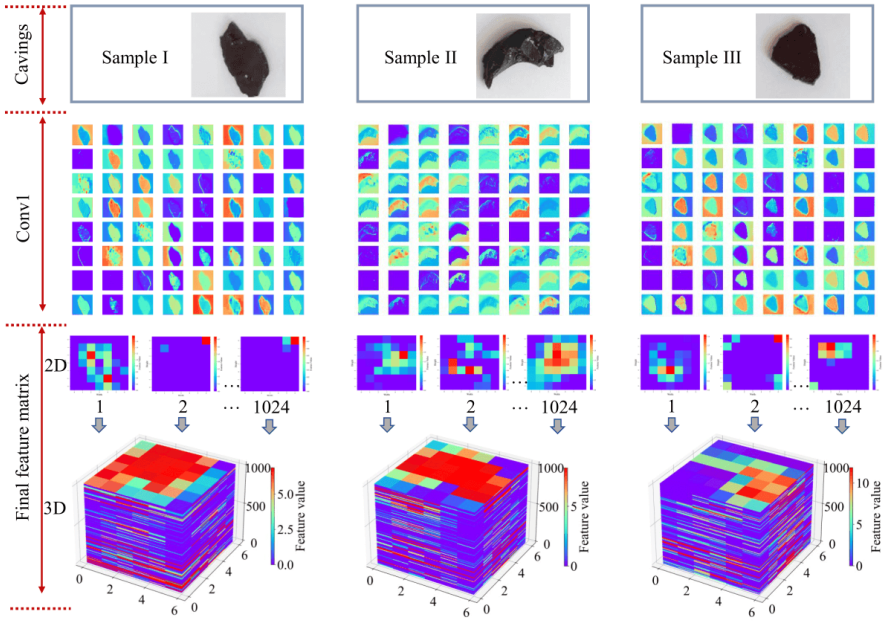

Fig. 9. GoogleNet feature extraction process.

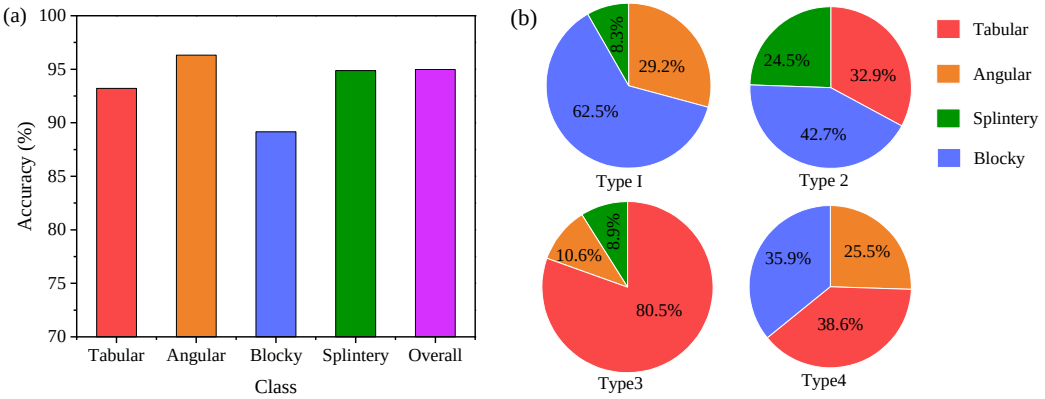

Fig. 10. Analysis of identification results. (a) Overall recognition accuracy and (b) after misrecognition, different types of cavings are recognized as the result of other cavings.

Table 3. Concrete measure of wellbore instability (Skea et al., 2018; Jin et al., 2022).

Failure	Measure	Concrete measure
I	I	Increase drilling fluid density
II	II	Increase flow rate and drilling fluid density
III	III	Reduce the rate of penetration and adjust drilling fluid density
IV	IV	Increase drilling fluid density and sealing ability

of cavings, and combines the surface roughness and edge angle changes of cavings to establish an expert system for wellbore stability analysis (Fig. 4). Based on the results, corresponding solutions are provided. According to current research and field construction feedback, instability types can be categorized into four classes (Table 4), each corresponding to specific

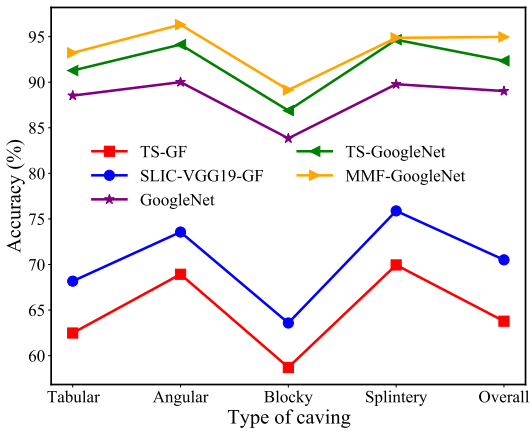

Fig. 11. Recognition accuracy of different models and types of caving.

mitigation measures (Table 3).

Table 4. Classification of wellbore instability mechanisms (Skea et al., 2018; Izurieta et al., 2019; Jin et al., 2022).

Failure	Caving	Formation	Instability mechanism
I	Splintery	High-pressure and high-rock strength	Drilling fluid density below pore pressure, tensile failure
II	Angular	High-stress difference, low rock strength	Drilling fluid density below collapse pressure, shear failure
III	Tabular	Well-developed bedding and weak planes	Low adhesion of weak surfaces, weak slip failure
IV	Blocky	Well-developed fractures and faults	Drilling fluid into natural fractures

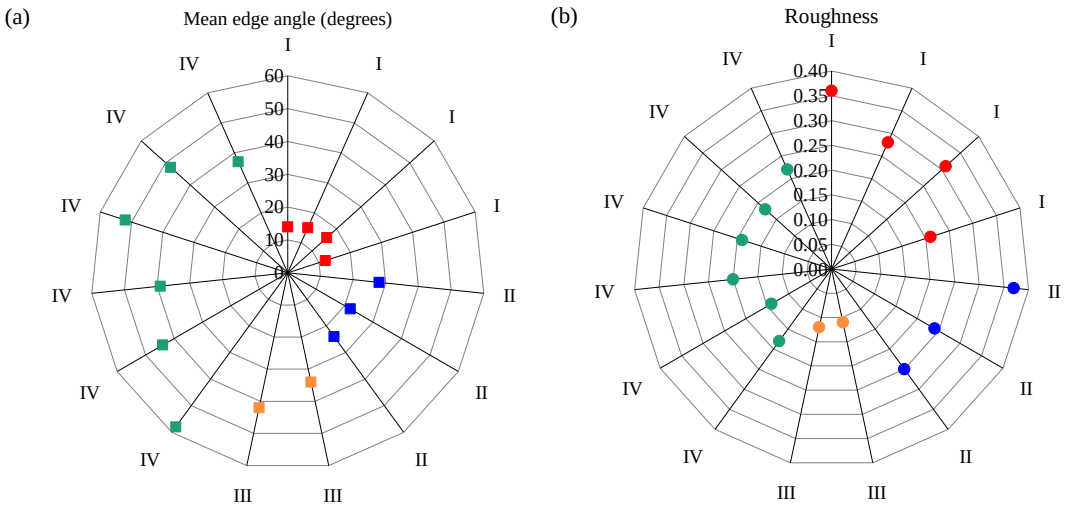

Fig. 12. The relationship between the edge angle, surface roughness of some cavings and the type of instability at a depth of 5,783 meters. (a) Mean edge angle of cavings and (b) roughness of cavings.

4.2 The influence of edge angle and roughness on wellbore instability

Based on research, the type of cavings can indicate the majority of reasons for wellbore instability. However, some instability causes may still be misjudged. Therefore, in developing an expert system for wellbore instability, the primary consideration is the shape characteristics of cavings, with secondary considerations given to the edge angles and surface roughness of the cavings. Research conducted on cavings collected from two sets of production wells revealed specific findings: when instability occurs due to slip along weak planes like bedding planes, the surface roughness of cavings is minimal, but the edge angles are large. When the drilling fluid density is insufficient and causes shear failure or tensile failure of the wellbore walls, cavings exhibit larger surface roughness but smaller edge angles (Fig. 12).

In the context of the four types of wellbore instability, if bedding planes are present, most of the instability occurs along these planes. Consequently, the surface roughness of the cavings is minimal. However, cavings produced under these conditions tend to have larger roundness, leading to significant variations in edge angles (Fig. 12). Instability caused by drilling fluid intrusion into natural fractures, results in cavings with surface roughness and edge angles that are similar to cavings occurring on bedding planes (Fig. 12). Therefore, distinguishing between damage along bedding planes and instability due to drilling fluid intrusion into natural fractures based solely on edge angle variations and surface roughness is challenging. These two types of instability are primarily differentiated based on caving types, but edge angles and surface roughness can help distinguish them from the other two types of instability. For shear and tensile failures causing instability, which results from stress imbalances in the formation, cavings typically exhibit larger surface roughness and smaller edge angles (Fig. 12). These patterns generally apply to most cases of instability, but not every collapse-induced instability scenario fits this analysis. Therefore, for a more accurate analysis of instability mechanisms, combining caving types, surface roughness, and edge angles is crucial. Utilizing an expert system model can provide the final analysis results based on these factors.

4.3 Engineering application analysis

To validate the effectiveness of the WSIRES model proposed in this study, we applied this method to a development well under construction in Xinjiang, China. The studied reservoir is situated in Cretaceous formations predominantly composed of mudstone and sandstone. We collected four batches of cavings from this formation and conducted a detailed analysis.

The first batch of cavings returned to the surface at a depth of 5,498 meters, with slight drilling sticking observed during the process. The largest length of the cavings at this depth was

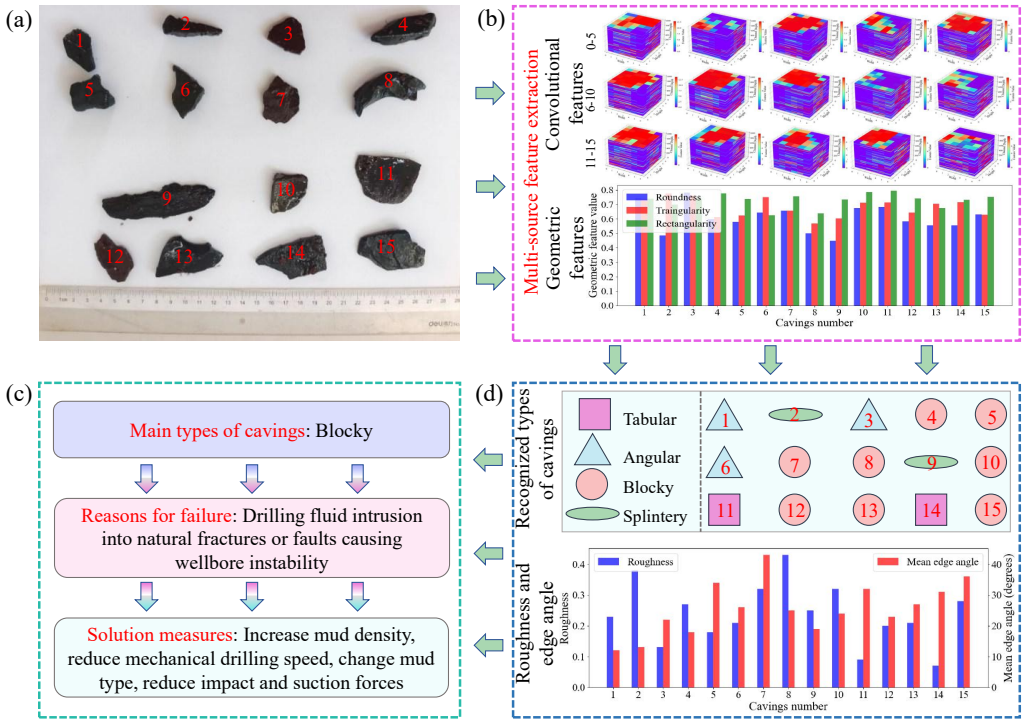

Fig. 13. At a depth of 5,498 meters, the WSIRES model was used for cuttings identification and wellbore stability analysis. (a) Cavings, (b) multi-source feature extraction, (c) reasons for failure and solutions and (d) the recognition results, edge angle, and roughness.

Table 5. Comprehensive analysis of cavings identification results and wellbore stability.

No.	Depth (m)	Types	Failure	Measure
1	5,498	Blocky	IV	IV
2	5,662	Tabular	III	III
3	5,783	Blocky	IV	IV
4	5,901	Angular	II	II

approximately 7 cm. Initially, convolutional features and geometric features including triangularity, rectangularity, and circularity were extracted separately using GoogleNet and SLIC-VGG19-TS methods. The convolutional features resulted in a 3D matrix of dimensions $(7, 7, 1024)$, while geometric features were also extracted (Fig. 13(b)). After feature extraction, the types of cavings were identified. A total of 15 samples were collected, out of which 8 were identified as block-shaped cavings (Fig. 13(d)). Therefore, the predominant type of cavings at this depth was identified as blocky cavings. Using the caving type as the primary criterion and considering variations in edge angles and surface roughness, an analysis of the wellbore instability mechanism was conducted. The analysis concluded that the instability of the wellbore at this depth was primarily caused by drilling fluid intrusion into natural fractures, resulting in blocky cavings (Fig. 13(c)). During the construction process, the instability issue at this depth can be resolved by increasing mud density, reducing mechanical drilling speed, and changing mud types.

We collected four batches of cavings from different depths of the same well. The identification and analysis process for each batch of cavings was similar to that of the cavings at 5,498 meters depth (Fig. 13). The final analysis revealed that the instability mechanism of the third batch of cavings was similar to that of the first batch. The second batch of cavings experienced instability due to slip failure along weak planes or reduced adhesion of weak planes under stimulation (Table 5). For such instability, it is recommended to drill slowly, make minor adjustments to mud density, and minimize agitation. The fourth batch of cavings experienced shear failure of the wellbore walls due to drilling fluid density being lower than the collapse pressure. For this type of instability, measures should include controlling drilling speed, increasing circulation rate, and increasing mud density within the safe density window (Table 5).

4.4 Limitations and future work

The WSIRES model proposed in this study can monitor wellbore instability during drilling and determine the instability mechanism, providing corresponding solutions based on the identified mechanism. This study is not limited to any specific formation, and as long as relevant image datasets are collected, wellbore instability analysis can be performed for any formation.

Due to limitations in the research conditions, this study has some shortcomings; however, they do not affect the current stage of its application. The study assumes that the cavings

formed by wellbore collapse do not undergo shape changes during the cuttings transportation process. However, for some cavings from rocks with weaker strength, a small portion of them may change shape during transportation. This is an aspect that needs to be further considered in future research.

5. Conclusions

This study proposes the WSIRES model for real-time analysis of wellbore stability based on image recognition and expert systems. Firstly, this study improves the threshold segmentation based on SLIC superpixel segmentation and VGG19 image classification to calculate the important geometric features of cavings. In addition, an MMF-GoogleNet model for caving type recognition was established by fusing geometric features with convolutional features. Finally, a WSES model was established to analyze the mechanism of wellbore instability by combining the caving type, surface roughness, and edge angle. Through modeling, testing, and practical application, this study draws the following conclusions:

1) The improved threshold segmentation based on SLIC superpixel segmentation and VGG19 image classification effectively segments the caving regions in the images and extracts their edges.
2) The MFF-GoogleNet model, by integrating geometric features of cavings with convolutional features extracted by GoogleNet, provides more accurate recognition of dropped block types.
3) There is a good mapping relationship between caving types and wellbore instability mechanisms, which allows us to infer the mechanisms of wellbore instability to a certain extent based on the caving types.
4) Using caving types as the primary dataset, the expert system for wellbore stability analysis, which integrates surface roughness and changes in edge angles of cavings, can analyze the causes of wellbore instability in real-time and effectively. It can also provide reasonable solutions accordingly.

Acknowledgements

The authors are grateful to the support provided by the National Natural Science Foundation of China (No. 52334001), the Science Foundation of China University of Petroleum, Beijing (No. 2462024YJRC021), and the Top Talent Program of China University of Petroleum, Beijing (No. ZX20230042).

Conflict of interest

The authors declare no competing interest.

References

Al-Ajmi, A. M., Zimmerman, R. W. Stability analysis of vertical boreholes using the Mogi-Coulomb failure criterion. International Journal of Rock Mechanics and Mining Sciences, 2006, 43(8): 1200-1211.

Ayoub, D., Masoud, C. S., Anthony, W. D., et al. Wellbore stability analysis to determine the safe mud weight window for sandstone layers. Petroleum Exploration and Development, 2019, 46(5): 1031-1038.

Bezabh, Y. A., Ayalew, A. M., Abuhayi, B. M., et al. Classification of mango disease using ensemble convolutional neural network. Smart Agricultural Technology, 2024, 8: 100476.

Chen, Q., Jiang, F., Guo, X., et al. Combine temporal information in session-based recommendation with graph neural networks. Expert Systems with Applications, 2024, 238: 121969.

Chen, X., Tan, C. P., Haberfield, C. M. A Comprehensive, Practical approach for wellbore instability management. SPE Drilling & Completion, 2002, 17: 224-236.

Christian, S., Liu, W., Jia, Y., et al. Going deeper with convolutions. Paper CVPR 4842 Presented at IEEE Conference on Computer Vision and Pattern Recognition, Boston, USA, 7-12 June, 2014.

Edwards, S., Matsutsuyu, B., Willson, S. Image unstable wellbores while drilling. SPE Drilling & Completion, 2004, 19(4): 236-243.

Favorskaya, M., Pakhirka, A. Animal species recognition in the wildlife based on muzzle and shape features using joint CNN. Procedia Computer Science, 2019, 159: 933-942.

Gao, J., Lin, H., Sun, J., et al. Double-porosity poromechanical models for wellbore stability of inclined borehole drilled through the naturally fractured porous rocks. SPE Journal, 2022, 27(4): 2491-2509.

Gong, F., Zhang, P., Xu, L. Damage constitutive model of brittle rock under uniaxial compression based on linear energy dissipation law. International Journal of Rock Mechanics and Mining Sciences, 2022, 160: 105273.

Gupta, A. K., Mathur, P., Sheth, F., et al. Advancing geological image segmentation: Deep learning approaches for rock type identification and classification. Applied Computing and Geosciences, 2024, 23: 100192.

He, M., Zhang, Z., Ren, J., et al. Deep convolutional neural network for fast determination of the rock strength parameters using drilling data. International Journal of Rock Mechanics and Mining Sciences, 2019, 123: 104084.

Houshmand, N., Goodfellow, S., Esmaeili, K., et al. Rock type classification based on petrophysical, geochemical, and core imaging data using machine and deep learning techniques. Applied Computing and Geosciences, 2022, 16: 100104.

Howard, A. G., Zhu, M., Chen, B., et al. MobileNets: Efficient convolutional neural networks for mobile vision applications. Paper CVPR 04861 Presented at IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, USA, 21-26 July, 2017.

Huo, F., Li, A., Zhao, X., et al. Novel lithology identification method for drilling cuttings under PDC bit condition. Journal of Petroleum Science and Engineering, 2021, 5: 108898.

Izurieta, C. A., Rocha, L. A., Sui, D. An approach in caving

recognition by an integrated model of computer vision and machine learning for any drilling environment. Paper AAPG 42374 Presented at Southwest Section Annual Convention, Texas, USA, 6-9 April, 2019.

Jin, J., Jin, Y., Lu, Y., et al. Image processing and machine learning based cavings characterization and classification. Journal of Petroleum Science and Engineering, 2022, 208: 109525.

John, E. U., Bernt, S. A., Kjell, K. F. Uncertainty evaluation of wellbore stability model predictions. Journal of Petroleum Science and Engineering, 2014, 124: 254-263.

Khan, A., Chefranov, A., Demirel, H. Image scene geometry recognition using low-level features fusion at multi-layer deep CNN. Neurocomputing, 2021, 400: 111-126.

Kolodziejczyk, J., Grzegorczyk-Dluciak, N., Kuliga, E. Rule-based expert system supporting individual education-and-therapeutic program composition in SYSABA. Procedia Computer Science, 2022, 207: 4535-4544.

Ma, T., Chen, P., Yang, C., et al. Wellbore stability analysis and well path optimization based on the breakout width model and Mogi-Coulomb criterion. Journal of Petroleum Science and Engineering, 2015, 135: 678-701.

Mishra, G., Gupta, P., Tanwar, R. Target recognition using pre-trained convolutional neural networks and transfer learning. Procedia Computer Science, 2024, 235: 1445-1454.

Paradarami, T. K., Bastian, N. D., Wightman, J. L. A hybrid recommender system using artificial neural networks. Expert Systems with Applications, 2017, 83: 300-313.

Patel, N., Penkar, S., Blyth, M. Managing drilling risk using an integrated approach to real-time pore pressure prediction. Paper SPE 192692 Presented at the Abu Dhabi International Petroleum Exhibition & Conference, Dhabi, UAE, 12-15 November, 2018.

Purkayastha, A. D., Rana, R., Kumar, R. R., et al. Cavings morphology analysis: A critical geomechanical tool in optimizing drillability. Paper IPTC 19981 Presented at the International Petroleum Technology Conference, Dhahran, Kingdom of Saudi Arabia, 13-15 January, 2020.

Purswani, P., Karpyn, Z. T., Enab, K., et al. Evaluation of image segmentation techniques for image-based rock property estimation. Journal of Petroleum Science and Engineering, 2020, 195: 107890.

Peng, H., Yue, Y., Luo, X., et al. Double-porosity poromechanical models for wellbore stability of inclined borehole drilled through the naturally fractured porous rocks. Geoenergy Science and Engineering, 2023, 228: 211756.

Ren, F., Zhu, C., Yuan, Z., et al. Recognition of shear and tension signals based on acoustic emission parameters and waveform using machine learning methods. International Journal of Rock Mechanics and Mining Sciences, 2023, 171: 105578.

Rill-García, R., Dokladalova, E., Dokládal, P. Pixel-accurate road crack detection in presence of inaccurate annotations. Neurocomputing, 2022, 480: 1-13.

Shen, P., Tang, H., Ning, Y., et al. A damage mechanics based on the constitutive model for strain-softening rocks. Engineering Fracture Mechanics, 2019, 216: 106521.

Shishehchi, S., Banihashem, S. Y. A rule based expert system based on ontology for diagnosis of ITP disease. Smart Health, 2021, 21: 100192.

Simonyan, K., Zisserman, A. Very deep convolutional networks for large-scale image recognition. Paper ICLR 1556 Presented at International Conference on Learning Representations, Banff, Canada, 14-16 April, 2015.

Skea, C., Rezagholilou, A., Far, P. B., et al. An approach for wellbore failure analysis using rock cavings and image processing. Journal of Rock Mechanics and Geotechnical Engineering, 2018, 10, 865-878.

Song, J., Jiao, W., Lancowicz, K., et al. A two-stage adaptive thresholding segmentation for noisy low-contrast images. Ecological Informatics, 2022, 69: 101632.

Stahl, B., Zhong, Z., Plehn, C., et al. Fuzzy expert system based evaluation framework for management procedure models. IFAC-PapersOnLine, 2015, 48: 1173-1178.

Wang, H. Intelligent identification of logging cuttings based on deep learning. Paper Presented at 2022 International Conference on the Energy Internet and Energy Interactive Technology, Wuhan, China, 25-27 March, 2022.

Wang, R., Li, W., Zhang, L. Blur image identification with ensemble convolution neural networks. Signal Processing, 2019, 155: 73-82.

Xu, D., Liu, J., Yang J., et al. Inference and learning methodology of belief-rule-based expert system for pipeline leak detection. Expert Systems with Applications, 2007, 32: 103-113.

Xu, K., Zouwei, L., Chen, Q., et al. Application of machine learning in wellbore stability prediction: A review. Geoenergy Science and Engineering, 2024, 232: 212409.

Xu, X., Xu, H., Wen, C., et al. A belief rule-based evidence updating method for industrial alarm system design. Control Engineering Practice, 2018, 81: 73-84.

Yang, L., Wang, Y., Chang, L., et al. A disjunctive belief rule-based expert system for bridge risk assessment with dynamic parameter optimization model. Computers & Industrial Engineering, 2017, 113: 459-474.

Yang, L., Ye, F., Liu, J., et al. Belief rule-base expert system with multilayer tree structure for complex problems modeling. Expert Systems with Applications, 2023, 217: 119567.

Yuan, J., Deng, J., Tan, Q., et al. Borehole stability analysis of horizontal drilling in shale gas reservoirs. Rock Mechanics and Rock Engineering, 2013, 46: 1157-1164.

Zhai, Y., Ng, A. H., Luo, Z., et al. Dynamic image segmentation and recognition measurement of axial compression experiment based on image clustering and semantic segmentation in RC column with FRP tubes. Measurement, 2024, 227: 114207.

Zoback, M. D. Reservoir Geomechanics. New York, USA, Cambridge University Press, 2007.

Zou, C., Zhai, G., Zhang, G., et al. Formation, distribution, potential and prediction of global conventional and unconventional hydrocarbon resources. Petroleum Exploration and Development, 2015, 42: 14-28.