Advances in Geo-Energy Research

Vol. 14, No. 2, p. 106-118, 2024

Original article

Artificial intelligence-based investigation of fault slip induced by stress unloading during geo-energy extraction

Zhenlong Song[1,2], Yunyi Qian[1]*, Yuxing Mao[1], Xiaofei Chen[2], P. G. Ranjith[3], Qinglin Deng[1]

[1]*State Key Laboratory of Coal Mine Disaster Dynamics and Control, School of Resources and Safety Engineering, Chongqing University, Chongqing 400044, P. R. China*

[2]*Department of Earth and Space Sciences, Southern University of Science and Technology, Shenzhen 518055, P. R. China*

[3]*Deep Earth Energy Laboratory, Department of Civil Engineering, Monash University, Melbourne 3800, Australia*

Keywords:
Unsupervised learning
acoustic emission signals
rock fracture identification
induced seismicity

Cited as:
Song, Z., Qian, Y., Mao, Y., Chen, X., Ranjith, P. G., Deng, Q. Artificial intelligence-based investigation of fault slip induced by stress unloading during geo-energy extraction. Advances in Geo-Energy Research, 2024, 14(2): 106-118.
https://doi.org/10.46690/ager.2024.11.04

Abstract:
Seismic events triggered by stress unloading during geo-energy extraction activities have become a key focus in both seismological research and engineering safety. This study presents a novel application of waveform neural networks, combining unsupervised and supervised learning techniques to classify and characterize fractures in laboratory-induced seismic events. Initially, A neural network model was initially developed that is capable of extracting time-frequency features from waveforms through unsupervised training on 1.2 million Acoustic Emission waveforms. Subsequently, this model was fine-tuned using a labeled dataset obtained from Brazilian split and uniaxial compression tests. The final result was a highly accurate model, achieving an accuracy rate of 97.6%. By applying this refined model, insights have been gained into the complex fault slip behaviors induced by geo-energy extraction activities. Our findings reveal that fluid infiltration at the onset triggers low-energy, shear-induced fractures in low-stress fault regions, which then escalate into tensile fractures during critical sliding in high-stress areas. Key precursors to fluid-induced seismicity have been identified, providing a major advance in early seismic hazard detection. These insights are essential for monitoring and early warning of induced seismicity during geo-energy extraction activities. Our work contributes significantly to improving the safety and efficiency of geo-energy extraction, including geothermal, shale gas, and conventional hydrocarbon production.

1. Introduction

Activities such as unconventional hydrocarbon extraction, geothermal energy production, and reservoir filling have been associated with minor to moderate seismic events (Elsworth et al., 2016; Grigoli et al., 2018; Kim et al., 2018; Zhuang et al., 2019; Lengliné et al., 2023), with some instances exceeding a magnitude of M_w 4.0 (Ellsworth, 2013; Keranen et al., 2014). These human-induced seismic events pose risks to infrastructure, safety, and industrial operations (Zöller and Hainzl, 2023). Consequently, comprehensive monitoring systems and stringent regulations are essential to anticipate, manage, and mitigate the seismic hazards associated with geo-energy extraction activities (Zoback and Gorelick, 2012; McGarr et al., 2015).

Currently, scholars widely focus on induced seismicity primarily caused by fault slip due to fluid intrusion during energy extraction. In open-pit mining, as well as natural oil and gas extraction, changes in in-situ stress within and around reservoirs can lead to fault stress unloading, which disrupts fault stress balances and triggers seismic activity (Wu, 2021). For instance, seismicity related to stress unloading

Yandy Scientific Press

*Corresponding author.
E-mail address: songzl@sustech.edu.cn (Z. Song); qianyy@cqu.edu.cn (Y. Qian); myx@stu.cqu.edu.cn (Y. Mao); chenxf@sustech.edu.cn (X. Chen); ranjith.pg@monash.edu (P. G. Ranjith); ethandeng07@gmail.com (Q. Deng).

Received August 11, 2024; revised September 5, 2024; accepted September 24, 2024; available online September 29, 2024.

is frequently observed in open-pit mining contexts (Lynch and Malovichko, 2006). Emanov et al. (2014) investigated induced seismicity at open-pit mines in Kuzbass, highlighting the significant seismic event known as the Bachatsky earthquake on June 18, 2013, which underscores the impact of mining activities on seismicity in the region. Similarly, at the Lacq gas field in France, more than 2,000 seismic events with magnitudes less than 4.2, attributed to natural gas extraction, were recorded between 1974 and 1997 (Bardainne et al., 2008). These unloading-induced earthquakes often involve subsurface fluid interactions. However, the role of groundwater and stress unloading in these events still requires further exploration. Qian et al. (2019) observed a shallow M_w 4.1 earthquake in the historically low seismic Sichuan Basin, China. This study emphasized that construction activities may have induced up to 0.11 MPa of Coulomb stress unloading and facilitated fluid ingress into faults, leading to fluid-induced seismicity under unloading conditions. However, the fluid-solid coupling processes within the fault during unloading-induced seismicity have received relatively little attention. Seismicity may arise as unloading processes create pathways for fluids to penetrate faults, initiating seismic events. The coupled fluid-mechanical unloading process is highly complex (Cai et al., 2024). While stress unloading reduces the load on the fractured rock masses, decreasing the risk of rock instability, the unloading and stress redistribution caused by fluid injection lead to crack propagation and coalescence, enhancing permeability and increasing the risk of fault instability. These mechanisms interact, complicating the assessment of fault stability during unloading processes. Conversely, stress unloading may also trigger fault instability. This complexity poses significant challenges in assessing the stability of excavations involving the coupled hydro-mechanical unloading process in geo-energy extraction (Zhou et al., 2021). Therefore, a comprehensive understanding of the mechanical deformation and permeability evolution in faults under coupled hydro-mechanical unloading conditions is crucial. This understanding is essential for interpreting and predicting induced seismicity.

To quantitatively investigate the fluid-solid coupling process within faults, accurately detecting fluid distribution within the fault zone is essential. Scholars have utilized numerical simulations (Wang et al., 2021a), such as the Discrete Element Method (Kou et al., 2019; Song et al., 2019; Zhang et al., 2020) and Peridynamics (Wang et al., 2016; Wang et al., 2019; Bazazzadeh et al., 2020; Zhou and Wang, 2021), to study fluid-solid coupling and hydro-mechanical behavior in rocks. However, these methods often rely on assumptions that may limit their accuracy in simulating real-world processes. Acoustic Emission (AE), characterized by its high temporal resolution and relatively high spatial resolution, has emerged as an effective tool for analyzing fluid-solid interaction dynamics (Zhan et al., 2019; Wang et al., 2021b). Kwiatek et al. (2014) conducted moment tensor analysis on stick-slip events in Westerly granite, revealing that faults exhibited distinct double-couple components under varying sliding conditions. The rise time/amplitude-average frequency (RA-AF) method, which utilizes waveform parameter information, has been extensively applied to classify the types of fractures (Aggelis, 2011). Cheng et al. (2022) conducted AE laboratory experiments to simulate induced seismic activity. Their findings indicate that high injection flow rates and elevated confining pressures increase the maximum magnitude of induced seismic events, thus heightening the associated seismic risks. Wang et al. (2024) utilized high-resolution AE records to perform full moment tensor inversion on all located AE sources, investigating the changes in AE characteristics related to induced fault slip and their relationship to fault roughness. While these studies have employed AE for qualitative research, they have yet to fully explore the rich information contained in AE waveforms. The sequential characteristics of AE waveforms hold critical data about tensile and shear fractures during fault slip. Employing neural networks to analyze this tensile-shear fracture data could provide valuable insights into the dynamics of fluid-solid coupled fracturing processes. In recent years, artificial intelligence has taken center stage in research and development across diverse domains, including impressive strides within the field of geological science (Jia et al., 2024). Yang et al. (2024) developed a multi-scale residual neural network that incorporates prior knowledge to predict dissolution characteristics, effectively identifying key dissolution zones and potential formation of significant channels. However, much of the current research remains rooted in supervised learning paradigms that require extensive, often costly, labeled datasets (Huang et al., 2021; Song et al., 2022). Therefore, this study is primarily focused on developing an innovative neural network capable of accurately identifying rock fracture types with minimal labeled data. Building on this model, this study analyzes the characteristics of tensile-shear fractures during fault slip in a fluid-solid coupling context, thereby elucidating its dynamic evolution mechanisms.

In this paper, unsupervised learning technique is used to develop a model capable of extracting features from time-frequency representations of waveforms, amassed from over 1.2 million rock fracturing experimental outputs. Subsequently, a small subset of labeled data is employed to fine-tune this pre-trained model. This approach allows the neural network to generalize effectively, enabling it to accurately distinguish tensile from shear-induced fractures. Using this model, we investigate the complex interplay and trigger relationships between tensile and shear-induced fractures in the context of fault slip events. The study provides a detailed examination of the initial triggering mechanisms associated with stress unloading during geo-energy extraction activates.

2. Data preparation and neural network methods

2.1 Data fundamentals preparation

To acquire labeled AE data for fracture induced by tensile and shear forces, we conducted laboratory experiments on both artificial and natural rock samples. The artificial samples included pure concrete (PC), concrete mixed with sand in a 1:1 ratio (CS1:1), and concrete mixed with sand in a 1:2 ratio (CS2:1). The natural rock samples comprised sandstone (SS), dolomite (DL), granite (GN) and shale (SH) (see Fig. 1(c)). It is widely recognized that the Brazilian split test predominantly

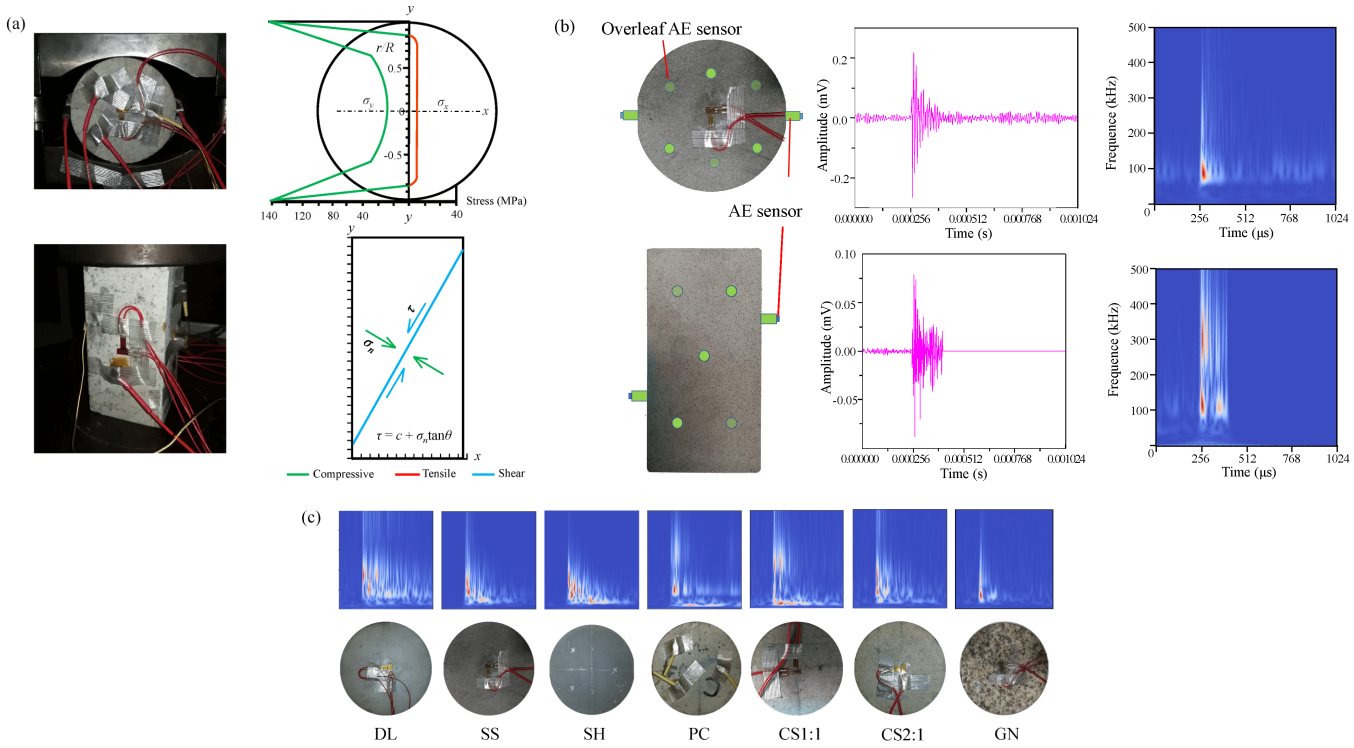

Fig. 1. Stress analysis and time-frequency data for Brazilian splitting and uniaxial compression. (a) Stress Analysis, (b) waves transformed time-frequency diagrams and (c) seven different materials and their corresponding time-frequency diagrams.

generates fractures under tensile stress, whereas uniaxial compression tests primarily induce fractures due to shear stress (Li et al., 2011; Chen et al., 2014; Garcia-Fernandez et al., 2018) (as shown in Fig. 1(a)). Given the intricacy of rock interior structures, even in regions dominated by shear stress, localized tensile fractures may still occur. Furthermore, the existence of microscopic-scale shear fractures remains a subject of ongoing debate (Petit and Barquins, 1988). Accordingly, in this study, fracture signal datasets were classified by the loading modes applied, such as Brazilian split (tensile-induced fracture) and uniaxial compression (shear-induced fracture), rather than by fracture types. It is also important to highlight that underground rock is typically subjected to compressive stress, which results in shear forces, while tensile stress can arise locally from block rotations or fluid intrusions. Thus, focusing on the stress conditions leading to fractures proves to be of greater practical significance. In this study, we used an AE probe with a resonance frequency of 300 kHz. The signals recorded by this probe typically contain critical information regarding the fracture mechanisms and rupture dynamics. To ensure accurate acquisition of all relevant details of the fracture process, we employed a sampling frequency of 1 MHz.

To further analyze the information contained within the AE waveforms, we applied a wavelet transformation to convert the waveforms into spectrograms. Time-frequency representations significantly improve feature extraction efficiency, making features that are not apparent in the time domain more prominent in the frequency or time-frequency domain (Bahmaninezhad et al., 2019). This allows the model to effectively capture relevant features without the need for increased complexity or extensive training data. The utility of time-frequency representations has been demonstrated across various fields, such as speech recognition (Arias-Vergara et al., 2021). Experimental results indicate that using spectrograms is superior to directly processing waveforms, as they more accurately reflect the energy distribution of signals and facilitate the model's learning of key features (Bahmaninezhad et al., 2019; Ma et al., 2019). This is especially important for our study, as the limited amount of waveform data constrains the training of large models. In this context, spectrograms allow us to fully extract features using simpler models with limited data, thereby enhancing both the model's performance and its reliability.

The resulting time-frequency diagrams are presented in Fig. 1(b), where the horizontal axis represents time (spanning 1,024 μs) and the vertical axis represents frequency (0 to 500 kHz). Spectrograms facilitate the observation of distinct features. For instance, AE signals from rock fractures under tensile stress primarily exhibit banded patterns, whereas those under compressive stress display network-like and partially banded patterns (as shown in Fig. 1(b)). This observation is consistent with previous studies (Song et al., 2022).

In this experiment, a total of 223,380 labeled AE data points were collected from Brazilian splitting and uniaxial compression tests conducted on seven types of rocks. These data were then transformed into time-frequency spectrograms. Subsequently, 80% of these spectrograms were randomly selected to form the training dataset, while the remaining 20% were set aside. These labeled datasets are essential for fine-tuning the model following unsupervised training. This approach allows the model to further refine its performance

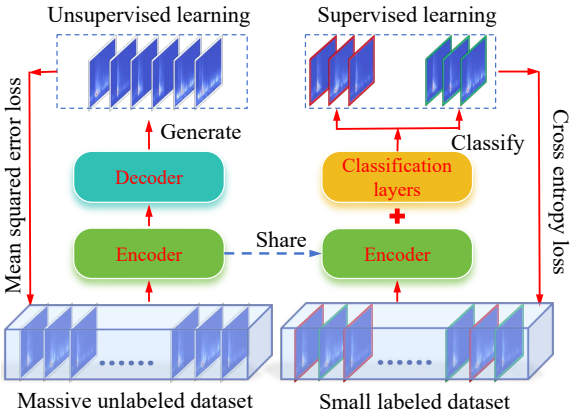

Fig. 2. Introduction of unsupervised-finetuning approach.

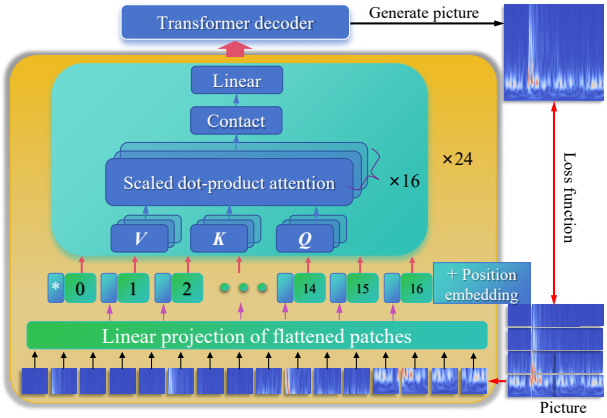

Fig. 3. Introduction of unsupervised model.

using the labeled data, thereby enhancing its ability to accurately interpret and classify AE patterns.

2.2 Introduction to neural network training methods

Compared to convolutional neural network models, the self-attention mechanism utilized in the Transformer models exhibits enhanced advantages on a global scale (Vaswani et al., 2017; He et al., 2022). This global attention mechanism effectively captures relationships between distant features in waveform sequences, whereas convolutional neural networks are more adept at focusing on local interactions between adjacent sequence features. This enables Transformers to better understand and integrate context across entire sequences, providing a deeper and more comprehensive analysis of complex patterns (Dosovitskiy et al., 2021). Consequently, this paper employs the Transformer model (including encoder and decoder) for training and recognition tasks. However, due to the substantial number of parameters inherent in the Transformer model, a large volume of labeled AE data is required to mitigate the risk of overfitting during training. But collecting a large amount of labeled data always costs time and effort. To address this challenge, this study adopts a strategy that combines unsupervised learning with fine-tuning. Initially, the Transformer model is trained in an unsupervised manner using a large-scale unlabeled dataset to learn the distribution of the waveform sample space. This foundational training allows the model to capture the general features of the data without labels, learning generalized feature representations that serve as a solid foundation for downstream tasks. The pretrained encoder functions as a feature extractor by transforming input data into representations that encapsulate important information. These representations provide a rich and informative starting point for supervised learning, enhancing the model's performance even with limited labeled data. Subsequently, supervised training is conducted with a small labeled dataset, utilizing the pre-trained encoder and a classifier. This process enables the model to effectively map labeled data features onto the learned sample space distribution, improving its ability to generalize to new or unseen data. This approach facilitates the training of large models with limited labeled data, as illustrated in Fig. 2. Accordingly, this paper will first pre-train the Transformer model using triaxial test data of sandstone from earlier studies, and then employ the AE waveforms from Brazilian splitting and crack propagation tests as labeled data to segregate the features extracted by the model.

3. Unsupervised and supervised models and training introduction

3.1 Introduction to unsupervised models and training

After undergoing a time-frequency transformation, the one-dimensional waveform data is converted into two-dimensional image representations, which better highlight the sequential characteristics of AE. To capture the features of these spectrograms, this study utilized 1.2 million AE spectrograms generated from our prior sandstone fracturing experiments (Huang et al., 2022). These AE events were obtained from experiments conducted under seven distinct true triaxial stress states, encompassing various Lode angles (-30°, -20°, -10°, 0°, 10°, 20°, 30°), representing different states of stress that may be encountered in underground environments. This diversity ensures that our model can effectively learn and make predictions across a broad range of geological conditions. The Transformer model was trained on this dataset to extract useful features from the spectrograms. Fig. 3 illustrates the schematic of the unsupervised training process, which comprises modules for embedding, an encoder, and a decoder. Initially, the images were partitioned into multiple segments; in this study, each image was divided into 14×14 sections. This segmentation was inspired by the Vision Transformer architecture (Dosovitskiy et al., 2021), where images are split into fixed-size patches to capture local and global features effectively. Dividing the 224×224 pixel images into 14×14 patches results in patch sizes of 16×16 pixels, balancing detail preservation and computational efficiency. These segments are combined with positional encoding and input into the encoder component of the Transformer network for feature extraction. The encoder in this study consists of 24 layers of attention and linear structures. The 24-layer configuration was selected to ensure sufficient model capacity for learning intricate features from the spectrograms while maintaining

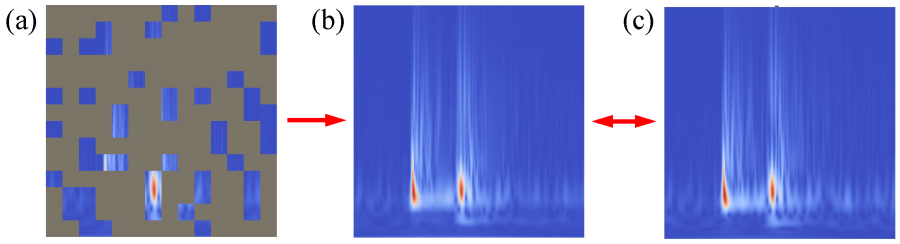

Fig. 4. Time-frequency images generated based on the pretrained model. (a) Input image, (b) reconstructed image and (c) original image.

computational feasibility. The attention mechanism enables the model to assign different weights to input data at diverse positions, thereby concentrating on parts most relevant to the current context. The subsequent Linear layer transforms the output from the attention mechanism into a fixed-size vector, enabling further processing by subsequent layers or other modules. The combination of Attention and Linear architectures aims to capture the dependencies between various positions within the input sequence (Azad et al., 2022; Wang et al., 2022). Notably, each Attention layer in this model comprises 16 Attention heads, which are capable of capturing different aspects or features of the time-frequency diagrams in parallel. This parallel processing enhances the capacity of the model to discern more intricate dependencies (Ma et al., 2019; Reza et al., 2022). After the images are encoded by the Encoder, they are passed through a simple decoder layer to reconstruct the original image. The decoder is relatively lightweight, comprising fewer Transformer layers. These layers are responsible for reconstructing the masked sections of the image, and in this study, the decoder typically includes 8 layers each with 8 attention heads. The difference between the generated and original images serves as an evaluation metric during model training (the loss function L_{re} is shown in Eq. (1)), facilitating improved feature extraction by the model:

$$L_{re} = -\frac{1}{n}\sum_{p\in P}\frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1}\frac{2\sigma_{xy} + C_2}{\sigma_x^2 + \sigma_y^2 + C_2} \tag{1}$$

where n is the number of images in a training batch (i.e., batch size), P is the image set of this training batch, p an individual image from that set P, x and y represent the processed and real images respectively, μ denotes the mean of the image, σ indicates the variance of the image, σ_{xy} denotes the covariance between images, and C_1 and C_2 are constants (to maintain stability and prevent division by zero in the denominator).

To assess the learning effectiveness of the unsupervised model in extracting time-frequency features, we employed a missing information restoration approach to evaluate the Encoder's ability to extract features from spectrograms. The model can simultaneously consider information from both the left and right sides of a patch. This approach facilitates a comprehensive representation of an image within its context, enabling the capture of intricate patterns and relationships. The final model can learn the feature information of all waveform spectrograms even without labeled data. When applied to specific downstream tasks, this methodology exhibits commendable performance even with a limited amount of labeled data (Devlin et al., 2019). This flexibility showcases the model's robustness and adaptability, making it highly effective for practical applications in geosciences where labeled data are often scarce. In this paper, partial time-frequency information was provided, and the model was tasked with using prior knowledge to reconstruct the complete time-frequency data. The closer the restoration matches the real image, the better the model's learning performance. The results of the unsupervised training are depicted in Fig. 4.

From Fig. 4, it is evident that when the model is presented with randomly masked images (with 75% of the image covered), the pretrained model reconstructs the time-frequency image by utilizing the provided prior information (the unmasked portion) based on the time-frequency features learned by the Encoder module. A comparison between the original and reconstructed images shows that the model can effectively restore the entire spectrogram. To quantitatively assess the similarity between the original and reconstructed images, we used histogram correlation analysis. Histograms represent the distribution of pixel intensities within an image, and their comparison provides a metric for assessing the similarity of intensity distributions between the two images. The histogram correlation coefficient between the original and reconstructed images was found to be 0.9880, indicating a high level of similarity. This result confirms that the pretrained model effectively captures the underlying patterns of AEs and can infer missing information from the unmasked portion.

3.2 Classification network training based on unsupervised encoder

The pretrained encoder module obtained through unsupervised learning is integrated with a two-layer fully connected network and a Sigmoid activation function to construct the classification network. In this model, we used a batch size of 64 and applied the Adam optimizer with a learning rate of 0.0001, along with a binary cross-entropy loss function. These parameters were selected to ensure the stability and effectiveness of the model during training. The network was trained using labeled data of tensile-induced and shear-induced fractures from seven types of rocks. The primary objective of this model is to classify the input AE spectrograms into two categories: rock fracture due to tensile stress and those caused by shear stress.

Fig. 5 presents the recognition accuracy and loss variation of the model throughout the training process. The model was

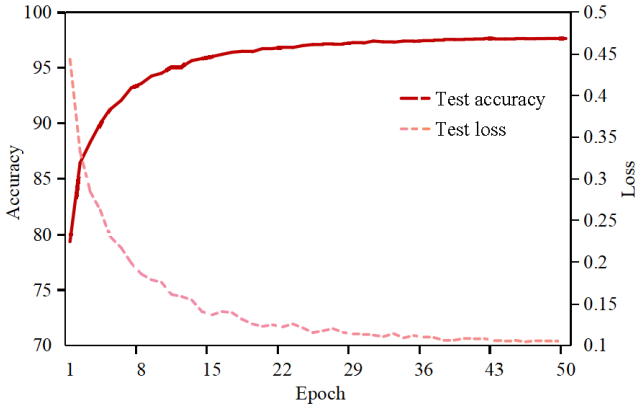

Fig. 5. Accuracy and loss variation on the test set during model training.

trained for 50 epochs. During the initial 15 epochs, the model's accuracy increased rapidly, indicating continuous adjustment of feature weights to improve accuracy. After 30 epochs, the accuracy on the test dataset stabilized at around 97% without further improvement. The best-performing training run, with an accuracy of 97.6%, was selected as the final model. This model was then used to classify and recognize fractures generated during true triaxial fracturing experiments.

4. Applications of the model to the study of induced seismicity mechanisms

4.1 Methods for fault induction experiments

To thoroughly understand the process of fault activation induced by stress unloading and fluid infusion, this study uses granite as the research subject, simulating earthquake induction through depressurization in a laboratory setting. Initially, triaxial pressures were applied incrementally to 5, 10 and 15 MPa. While maintaining these confining pressures constant, the water pressure was gradually increased to 3.5 MPa. Once the water pressure stabilized, the maximum principal stress was reduced, allowing fluids to permeate the fault thoroughly and attempting to induce slippage. Throughout the loading process, AE events during fault slippage were recorded. The specific distribution of the granite specimens, AE sensors and spectrograms is illustrated in Fig. 6.

4.2 Fault rupture process analysis

Fig. 7 illustrates the stress-strain response of specimens at various stages of a pressure relief experiment and the corresponding distribution of fault ruptures. In Fig. 7(b), the blue line denotes the changes in displacement along the axis of minimum principal stress. This displacement decreases progressively as fluid pressure increases, indicating expansion in that direction due to the reduction in effective stress by fluid injection. Concurrently, a decrease in the maximum principal stress results in a reduction in this expansion, suggesting a partial elastic rebound. Notably, a rapid expansion in the direction of minimum principal stress occurs when the maximum principal stress aligns with the intermediate principal stress, signaling a swift fault slip and potential earthquake initiation. This phenomenon can be attributed to the initial constriction of the fault by the maximum principal stress, which hinders fluid penetration. However, as the stress decreases, fluid gradually infiltrates the fault's locked segment, altering the frictional properties and triggering unstable slip. To delve deeper into the fault slip dynamics, this study employs AE monitoring to analyze the fault activation process, facilitated by depressurization and fluid injection.

Figs. 7(a), 7(d) and 7(e) show the dynamic evolution of fracture mechanisms within the fault, where the size of the shapes indicates the energy of AE events, and the color representing timing. The fault activation process can be categorized into three distinct phases based on the stress conditions and fracture distribution. Initially, low-energy AE events are predominantly detected in the fault's center with some high-energy events at the fault edge (Fig. 7(a)). These events are generated during the initial loading process. During the water injection phase, AE events spread from the center to the edges, illustrating the outward expansion of the primary friction zone within the fault due to water infusion. The third phase, characterized by induced fault slipping, shows a continuation of this expansion at the fault edges. At this stage, AE events have higher energy due to a sudden energy release, and they're more focused, creating a narrow pattern rather than the wider distribution in earlier stages. This phenomenon may be due to fluid diffusion into the fault locking area under unloading conditions, altering the stress states and the fault's physical properties, thus facilitating slip.

An interesting deviation of AE events is observed in the fault's lower section, attributed to rock fracturing during the preloading phase, which causes a delayed AE signal detection. Utilizing the time-difference localization algorithm (Li et al., 2016), these delays lead to a perceived displacement of the localization point, resulting in a deflection greater than the actual fracture (Figs. 7(a) and 7(b)). Consequently, a refined localization algorithm is required to correct this discrepancy. However, this does not compromise the analysis of the AE timing features, nor the assessment of fluid dynamics within the fault. For a more nuanced examination of fluid-induced fault slipping, the AE localizations have been segmentally analyzed at different stages.

4.2.1 Loading stage

Fig. 8 presents the spatiotemporal dynamics of fault rupture throughout the stress loading stage. In this figure, circles signify fractures induced by tension, while squares denote fractures caused by shear. In the initial phase of loading, with fluid pressure at 0.5 MPa, the hydrostatic stress across the system is relatively low. This facilitates fluid permeation from the fault's core to its edge, leading to an array of fracturing events. High-energy AE events are recorded in the fault's locking zone at 257 seconds. As stress escalates, the locking zone compacts, redirecting fluids towards the central area. From Fig. 8(a), it can be seen that fractures caused by tensile stress predominantly occur after 200 seconds, during the phase of stress loading. While fractures produced in the early stages of fluid injection are primarily caused by shear. These fractures, induced by the applied stress, display a rup-

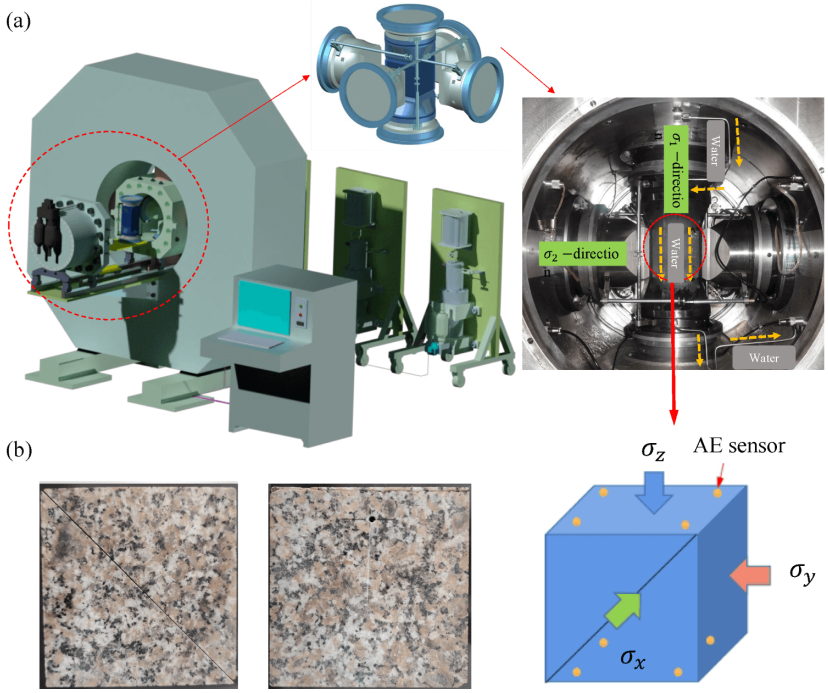

Fig. 6. Experimental Overview. (a) Experimental equipment and (b) specimen diagram.

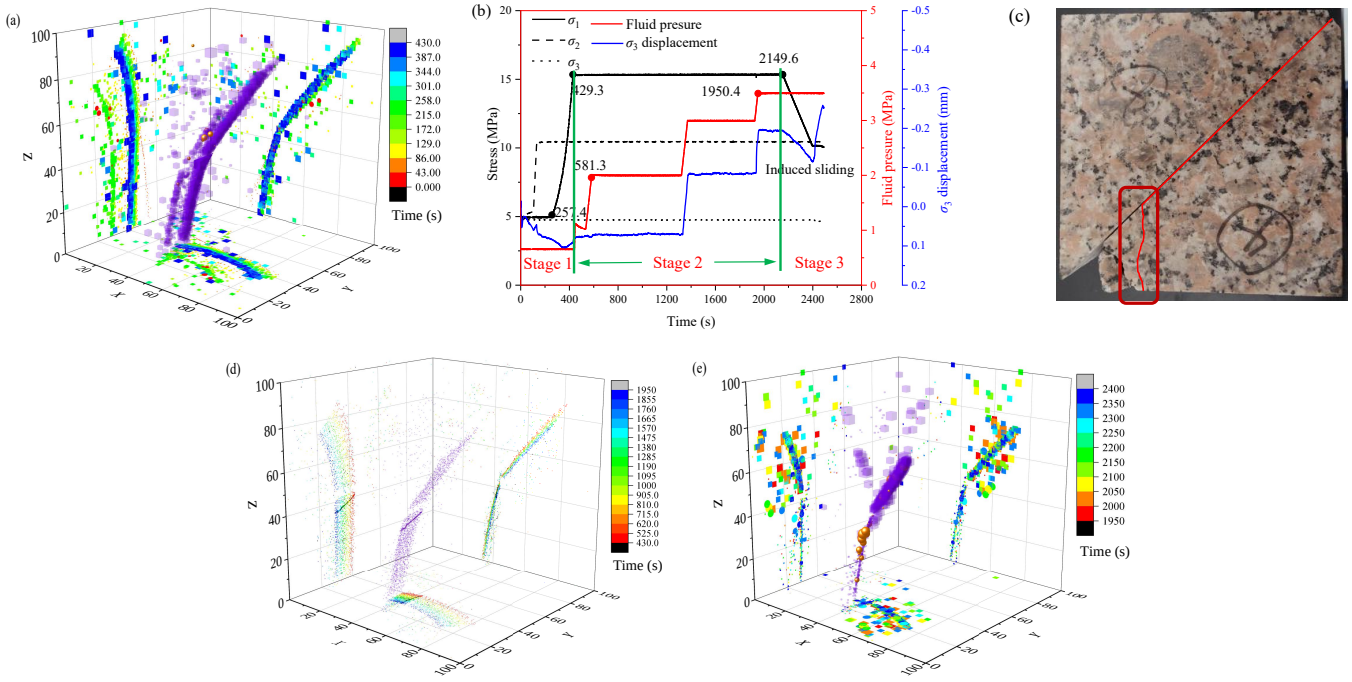

Fig. 7. The internal evolution process of faults and their corresponding loading states. (a) The distribution of fault activity during the initial loading stage (stage 1), (b) stress, water pressure, and fault adjustment during the induced fracturing process, (c) the fracture distribution of the granite fault specimen after the experiment, (d) the distribution of fault rupture during the fluid injection stage (stage 2) and (e) the distribution of fault rupture during the stress unloading-induced sliding stage (stage 3).

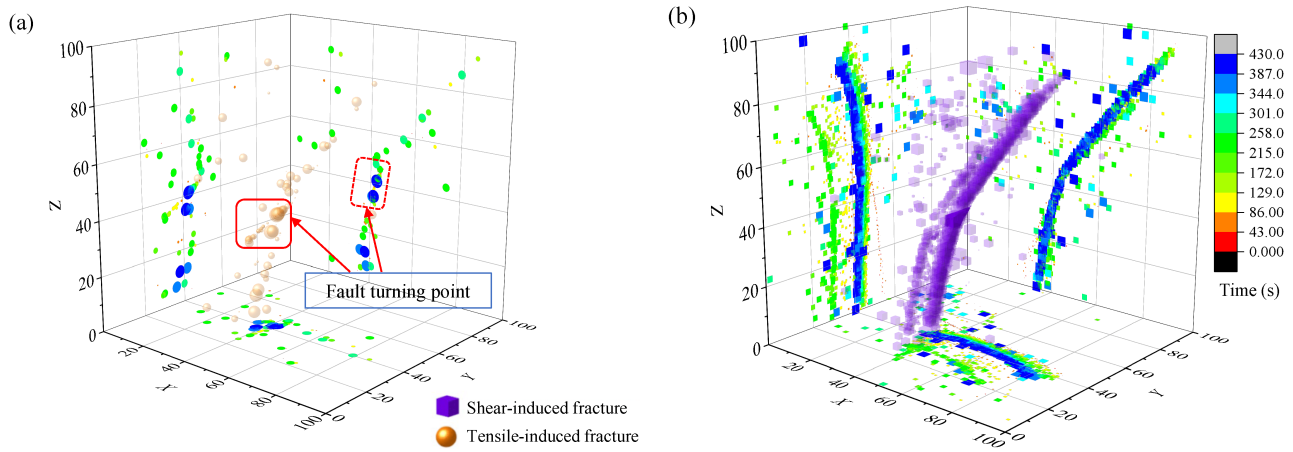

Fig. 8. Spatiotemporal evolution patterns of different fractures during the stress loading phase. (a) Distribution of tensile induced fractures and (b) distribution of shear induced fractures.

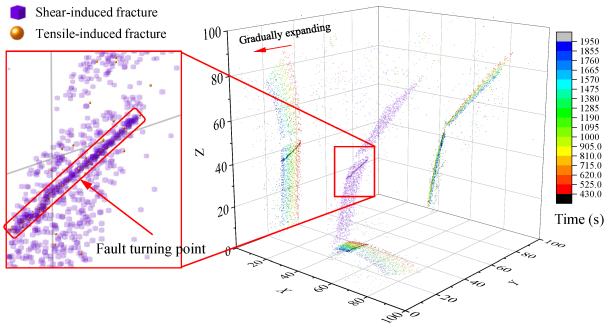

Fig. 9. Evolution law of AE during fluid pressure rise.

ture mechanism that is distinct from those initiated by fluid penetration. During the phase of stress loading, a significant number of fractures are caused by tensile forces. Additionally, we note that at the fault's turning point, where stress concentrates, a series of fractures induced by tension are generated (as shown in Fig. 8(a)). This phenomenon indicates that areas of stress concentration are prone to develop fractures triggered by tension, whereas fractures due to fluid diffusion are typically caused by shear stress. This pattern also repeatedly appears in subsequent loading.

4.2.2 Injection stage

After the triaxial stress reached its predetermined threshold and stabilized, the fluid pressure began to increase gradually. Consequently, ruptures within the fault initiated at the center and propagated outwards, casting a rainbow-like distribution on the Y-Z plane (Fig. 9). Initially, as the fluid pressure rose, the ruptures predominated near the fault's central axis. With continued increase in fluid pressure, the ruptures gradually extended towards the edge of the fault. During the third increase in fluid pressure, the ruptures continued to propagate towards the edge of the fault; however, the diffusion rate of the fluid decreased, indicating the presence of a high-stress zone within the fault that can resist further fluid penetration. A noteworthy observation from Fig. 9 is that ruptures caused by tensile stress are predominantly concentrated in the high-stress regions at fault bends, with other areas experiencing fewer tensile-induced ruptures. The fluid-induced ruptures exhibit a rainbow-like distribution primarily resulting from shear stress. This pattern is similar to the AE phenomena observed in stage 1 before stress loading, where ruptures induced by fluids are exclusively shear-caused, while those due to tensile stress typically result from increased stress levels. Given that the stress conditions were constant during this phase, it becomes evident that the ruptures were exclusively induced by fluid permeation. Evidently, the fractures caused by fluid intrusion contrast markedly from those under standard rock loading conditions. This unique pattern could serve as a crucial framework for elucidating fluid-induced fault sliding and provide essential insights for earthquake prediction systems.

4.2.3 Unloading and induced sliding stage

After the third fluid injection, we maintained a constant fluid pressure while simultaneously decreasing the maximum principal stress to induce fault sliding. The unloading phase, as shown in Fig. 10, resulted in the fault generating a series of high-energy ruptures that propagated along the edge of the previously formed rupture zone. Notably, a few tensile ruptures were observed throughout this process, mainly in the lower half of the fault. This pattern suggests that during unloading, some stress-induced ruptures occurred, with shear ruptures dominating in the upper part of the fault, showing that these were mostly due to fluid entering the fault. The data show that reducing stress helped open pathways for fluids, allowing them to penetrate the fault's locked area. This movement reduced the shear strength in the region, ultimately leading to seismic events. During the sliding stage, more tensile fractures appeared in the upper part of the fault, indicating that high stress from fault movement can also cause ruptures.

Through our analysis, we have elucidated the flow dynamics of fault fluids influenced by the combined actions of the stress and seepage fields. An increase in fluid pressure initially channels the flow longitudinally along the fault traces, which is subsequently impeded by transverse barriers. This process

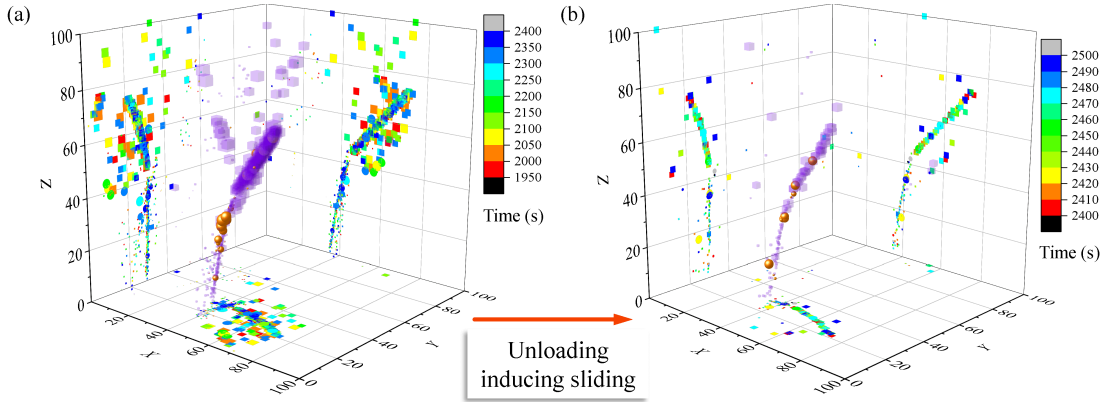

Fig. 10. Spatiotemporal evolution of AEs during fluid induced sliding process. (a) Unloading stage and (b) sliding stage.

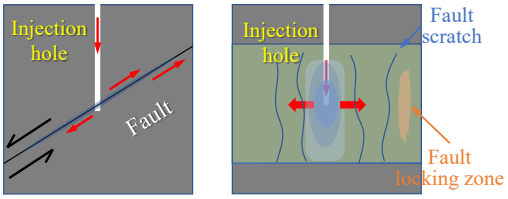

Fig. 11. Schematic diagram of fluid diffusion under the influence of fault scratches.

is similar to previous studies, where Hui et al. (2021) identified fault structures comprising an impermeable core and high-permeability damage zones. Specifically, in the fault zone architecture section, the study describes a typical fault zone containing a low-permeability fault core surrounded by highly fractured, high-permeability damage zones. These characteristics are crucial for understanding fluid flow dynamics during hydraulic fracturing and fault reactivation. The damage zones act as hydraulic conduits for pressure diffusion along faults, while the fault core serves as an impermeable barrier preventing cross-fault flow.

During the unloading stage, rapid fluid diffusion overcomes the barriers of the fault trace, leading to fault sliding (Fig. 11). This observation aligns with prior research by Anderson and Bakker, which investigated the role of faults in directing and obstructing fluid movement (Anderson and Bakker, 2008). Our findings reveal the fundamental mechanisms driving earthquakes induced during unloading. Specifically, our analysis indicates that such earthquakes are triggered by the infiltration and lubrication of the fault by fluids, rather than solely by a reduction in effective stress due to fluid entry. To avoid induced seismicity, it is essential to maintain the integrity of the fault blocking zone. The onset of earthquakes is due to the weakening of shear strength across the fault caused by fluid intrusion, which can be initiated by either heightened fluid pressure or by unloading processes that facilitate fluid entry. This mechanism may also account for delayed induced seismicity, where seismic activity occurs after the cessation of fluid injection due to the ongoing fluid diffusion within the fault, eventually leading to fault slip and earthquakes after a certain period.

4.3 Discussion on induced earthquake early warning based on fracture distribution

To further analyze the relationship between the stress states and the proportion of tensile-induced fractures on the fault surface, and based on this, to propose parameters for earthquake early warning, we divided the fault on the Z-Y plane into 15 regions. The fault was segmented according to the rupture sequence from the center to the edges into five areas, and vertically from top to bottom into three sections, resulting in a total of fifteen distinct regions, as illustrated in Fig. 12(a). To more precisely assess the stress distribution and fracture conditions in each region, we calculated and averaged the improved b-value (*ib* values) (Grosse et al., 2021) from AE events for each area. The formula for calculating *ib* values is as follows:

$$ib = \frac{\log_{10} N(w_1) - \log_{10} N(w_2)}{(\alpha_1 + \alpha_2)\sigma} \tag{2}$$

where $N(w_1)$ and $N(w_2)$ denote the amount of amplitude exceeding $\mu - \alpha_1\sigma$ and $\mu + \alpha_2\sigma$, respectively; α_1 and α_2 are taken as 0 and 1 (Shiotani et al., 2001; Colombo et al., 2003; Watanabe et al., 2007). We test this method with the experimental data in this paper.

According to previous research, lower *ib* values indicate a concentration of stress in those regions, whereas high *ib* values suggest a low stress state. Additionally, we have quantified the proportion of tensile-induced fractures relative to the total fractures in each area. In Figs. 12(b) and 12(c), the distribution of *ib* values and the proportion of tensile-induced fractures at different locations on the fault are displayed. It is observed that the *ib* values decrease from the center to the edge of the fault, while the distribution of tensile-induced fractures shows the opposite trend, being lower in the central region and higher towards the edges. This distribution pattern reveals the stress changes caused by fluid invasion, where the fluid spreads from the center to the edges of the fault and encounters high-stress areas that impede further expansion. Figs. 12(d) and 12(e) further analyze the distribution of *ib* values and the proportion of tensile-induced fractures before and after unloading. The results show that unloading leads to a significant increase in *ib* values and a notable reduction in the proportion of tensile induced fractures at the fault edges. This phenomenon

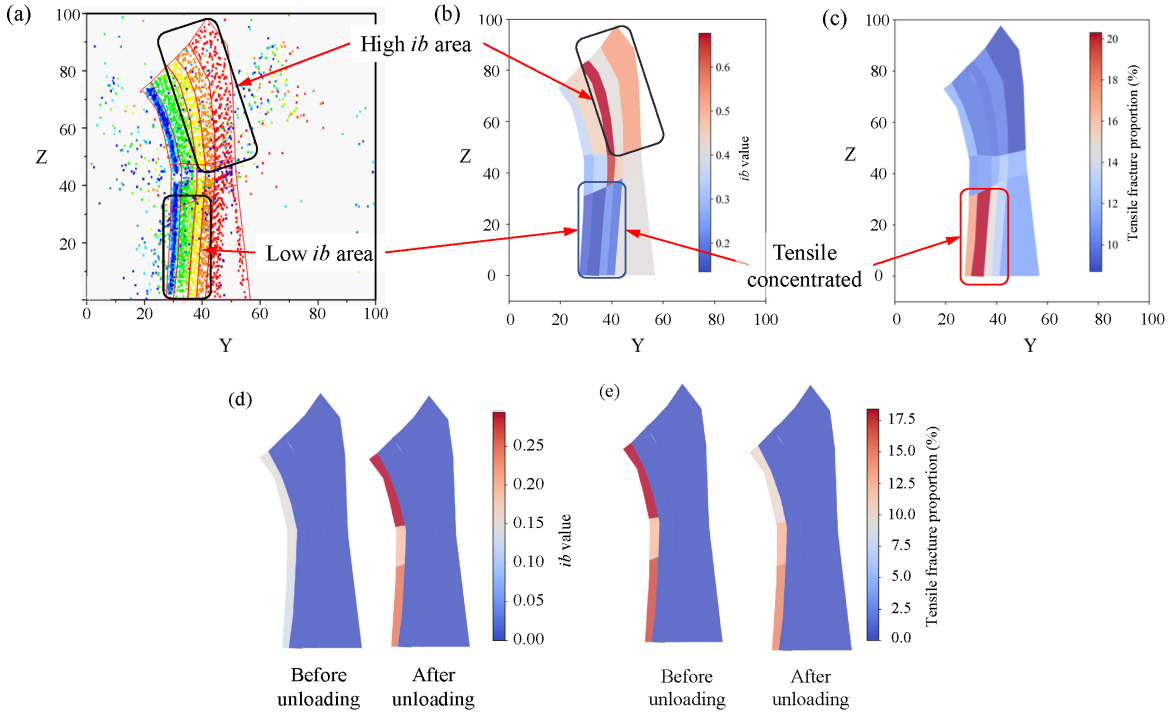

Fig. 12. Temporal and spatial distribution map of *ib* values and tensile-induced fractures on the fault plane. (a) Regional division based on fluid diffusion processes along the fault plane, (b) distribution of *ib* values in different regions, (c) proportion of tensile-induced fractures in different regions, (d) changes in *ib* values on the fault plane before and after unloading and (e) changes in the proportion of tensile-induced fractures on the fault plane before and after unloading.

is very similar to the findings in other studies (Lei, 2024). This indicates that the stress is released during the depressurization process, allowing further fluid penetration and inducing fault sliding.

Additionally, an increase in loading stress in these areas usually corresponds with an increase in the proportion of tensile-induced fractures. For example, in Fig. 9, the stress-concentrated areas at the fault's turn exhibit a large number of tensile-induced fractures. This suggests that tensile fractures primarily occur in areas of stress concentration, indicating that these areas may be more prone to ruptures and potentially unstable slips. In contrast, during the stage of fault stable slip induced by fluid, the observed fractures are predominantly caused by shear, which differs significantly from the mechanism of tensile-induced fracture formation.

Fluid invasion into low-stress areas of a fault typically leads to shear-induced fractures, accompanied by slow fault slip, during which the *ib* values are significantly higher. Low-pressure areas do not exhibit fault locking. In these regions, fluid injection decreases the fault's shear strength, leading to ruptures that often manifest as shear slippage along the fault plane with minimal new rock fracturing, which is the primary source of tensile-induced fractures. This results in slow fault slip, Which generally does not produce strong vibrations and thus has a lesser impact on the surface. However, over time, this slow slip may cause fault creep deformation, potentially affecting the fault's stability and necessitating prolonged monitoring to evaluate its impact. In contrast, when fluids infiltrate the high-stress boundaries of the fault, the resulting *ib* values from fluid invasion are very low due to the high stress in this area, along with a significant increase in tensile-induced fractures. In high-pressure regions, the fault is locked, meaning that new fractures need to be generated during rupture, leading to a substantial number of tensile fractures. Once these locked, high-stress areas fail, rapid fault slip occurs, which may trigger earthquakes. This rapid slip and the associated seismic activity pose significant risks to nearby environment and infrastructure.

When fluid enters a low stress region, the *ib* value is high, and the proportion of tensile-induced fractures is relatively low. As the *ib* value increases and the proportion of tensile induced fractures rises, it indicates that the fluid has entered a high-stress area, often corresponding to the fault core (Hui et al., 2021). This suggests an elevated risk of seismic events, necessitating continuous monitoring of the *ib* value and the proportion of tensile-induced fractures. When the *ib* value begins to increase while the proportion of tensile-induced fractures decreases, it insinuates that the fault in this region is becoming unstable and may be approaching failure. Simultaneous consideration of these two parameters can provide a more reliable early warning for fault slip. In this experiment, the average *ib* value increased by 66%, while the proportion of stretch-induced fractures decreased by 22% before and after fault depressurization. For other regions, specific early warning indicators need to be determined based on local conditions. These findings underscore the importance of continuous monitoring of stress parameters and fracture types for more accurate seismic risk assessment. Such insights provide valuable geological and mechanical data that can

improve the effectiveness of seismic activity monitoring and early warning systems, especially in the context of energy extraction-induced seismicity.

5. Conclusions

An advanced neural network model that employs unsupervised learning coupled with fine-tuning strategies have been engineered. Requiring only a small labeled dataset, this model excels in its precision, achieving an accuracy rate of 97.6% in distinguishing between rock fractures caused by tensile and shear forces.

In low-stress areas of a fault, the intrusion of fluids can trigger fault rupture through shearing and allow it to spread outward. This spreading process is obstructed upon encountering high-stress regions, where stress unloading in these high stress areas further promotes fluid diffusion and may lead to unstable fault sliding. Additionally, the observation of reduced *ib* values and an increase in tensile-induced fractures within the fault zone could indicate a risk of fault instability, thereby increasing the likelihood of induced seismicity. Therefore, monitoring *ib* values and tensile-induced fractures is crucial for assessing the stability of faults and the risk they pose in triggering earthquakes. In the context of energy extraction, continuous monitoring of these parameters is essential for effective seismic hazard assessment and the development of early warning systems, aiding in the mitigation of induced earthquakes and ensuring safer extraction practices.

Acknowledgements

This study was funded by Shenzhen Science and Technology Program (No. JCYJ20210324104804013), the Fundamental Research Funds for the Central Universities (No. 2024CDJXY025) and the National Natural Science Foundation of China (Nos. 42274070 and 42004036). The work was carried out at National Supercomputer Center in Tianjin, and the calculations were performed on Tianhe new generation supercomputer. This study was also partly supported by Center for Computational Science and Engineering at Southern University of Science and Technology.

Conflict of interest

The authors declare no competing interest.

References

Anderson, E. I., Bakker, M. Groundwater flow through anisotropic fault zones in multiaquifer systems. Water Resources Research, 2008, 44(11): W11433.

Aggelis, D. G. Classification of cracking mode in concrete by acoustic emission parameters. Mechanics Research Communications, 2011, 38: 153-157.

Arias-Vergara, T., Klumpp, P., Vasquez-Correa, J. C., et al. Multi-channel spectrograms for speech processing applications using deep learning methods. Pattern Analysis and Applications, 2021, 24: 423-431.

Azad, R., Al-Antary, M. T., Heidari, M., et al. TransNorm: Transformer provides a strong spatial normalization mechanism for a deep segmentation model. IEEE Access, 2022, 10: 108205-108215.

Bahmaninezhad, F., Wu, J., Gu, R., et al. A comprehensive study of speech separation: Spectrogram vs waveform separation. Paper Presented at Interspeech 2019, Graz, Austria, 15-19 September, 2019.

Bardainne, T., Dubos-Sallée, N., Sénéchal, G., et al. Analysis of the induced seismicity of the Lacq gas field (Southwestern France) and model of deformation. Geophysical Journal International, 2008, 172(3): 1151-1162.

Bazazzadeh, S., Mossaiby, F., Shojaei, A. An adaptive thermo-mechanical peridynamic model for fracture analysis in ceramics. Engineering Fracture Mechanics, 2020, 223: 106708.

Cai, J., Jiao, X., Wang, H., et al. Multiphase fluid-rock interactions and flow behaviors in shale nanopores: A comprehensive review. Earth-Science Reviews, 2024, 257: 104884.

Chen, J., Guo, Q., Liu, B., et al. Dynamic Brazilian test of brittle materials using the split Hopkinson pressure bar and digital image correlation. Strain, 2014, 50(6): 563-570.

Cheng, Y., Xu, T., Zhang, Y., et al. An laboratorial investigation of induced seismicity characteristics in EGS hydraulic fracturing. Geothermics, 2022, 105: 102482.

Colombo, S., Main, I. G., Forde, M. C. Assessing damage of reinforced concrete beam using "*b*-value" analysis of acoustic emission signals. Journal of Materials in Civil Engineering, 2003, 15(3): 280-286.

Devlin, J., Chang, M. W., Lee, K., et al. BERT: Pre-training of deep bidirectional transformers for language understanding. Paper Presented at the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Minnesota, USA, 2-7 June, 2019 .

Dosovitskiy, A., Beyer, L., Kolesnikov, A., et al. An image is worth 16×16 words: Transformers for image recognition at scale. Paper Presented at The Ninth International Conference on Learning Representations, Appleton, Wisconsin, USA, 4-5 May, 2021.

Ellsworth, W. L. Injection-induced earthquakes. Science, 2013, 341(6142): 1225942.

Elsworth, D., Spiers, C. J., Niemeijer, A. R. Understanding induced seismicity. Science, 2016, 6318(354): 1380-1381.

Emanov, A. F., Emanov, A. A., Fateev, A. V., et al. Mining-induced seismicity at open pit mines in Kuzbass (Bachatsky earthquake on June 18, 2013). Journal of Mining Science, 2014, 50(2): 224-228.

Garcia-Fernandez, C. C., Gonzalez-Nicieza, C., Alvarez-Fernandez, M. I., et al. Analytical and experimental study of failure onset during a Brazilian test. International Journal of Rock Mechanics and Mining Sciences, 2018, 103: 254-265.

Grigoli, F., Cesca, S., Rinaldi, A. P., et al. The November 2017 M_w 5.5 Pohang earthquake: A possible case of induced

seismicity in South Korea. Science, 2018, 360(6392): 1003-1006.

Grosse, C. U., Ohtsu, M., Aggelis, D. G., et al. Acoustic Emission Testing: Basics for Research-Applications in Engineering. Berlin, Germany, Springer International Publishing, 2021.

He, K., Chen, X., Xie, S., et al. Masked autoencoders are scalable vision learners. Paper Presented at IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), New Orleans, LA, USA, 18-24 June, 2022.

Huang, J., Hu, Q., Qin, C., et al. Pre-peak acoustic emission characteristics of tight sandstone failure under true triaxial stress. Journal of Natural Gas Science and Engineering, 2022, 102: 104576.

Huang, J., Hu, Q., Song, Z., et al. Classification of cracking sources of different engineering media via machine learning. Fatigue & Fracture of Engineering Materials & Structures, 2021, 44: 2475-2488.

Hui, G., Chen, S., Chen, Z., et al. Investigation on two M_w 3.6 and M_w 4.1 earthquakes triggered by poroelastic effects of hydraulic fracturing operations near Crooked Lake, Alberta. Journal of Geophysical Research: Solid Earth, 2021, 126(5): e2020JB020308.

Jia, J., Li, D., Wang, L., et al. Novel transformer-based deep neural network for the prediction of post-refracturing production from oil wells. Advances in Geo-Energy Research, 2024, 13(2): 119-131.

Keranen, K. M., Weingarten, M., Abers, G. A., et al. Sharp increase in central Oklahoma seismicity since 2008 induced by massive wastewater injection. Science, 2014, 345(6195): 448-451.

Kim, K. H., Ree, J. H., Kim, Y., et al. Assessing whether the 2017 M_w 5.4 Pohang earthquake in South Korea was an induced event. Science, 2018, 360(6392): 1007-1009.

Kou, M., Han, D., Xiao, C., et al. Dynamic fracture instability in brittle materials: Insights from DEM simulations. Structural Engineering and Mechanics, 2019, 71(1): 65-75.

Kwiatek, G., Goebel, T. H. W., Dresen, G. Seismic moment tensor and *b* value variations over successive seismic cycles in laboratory stick-slip experiments. Geophysical Research Letters, 2014, 41(16): 5838-5846.

Lei, X. Fluid-driven fault nucleation, rupture processes, and permeability evolution in oshima granite-Preliminary results and acoustic emission datasets. Geohazard Mechanics, 2024, https://doi.org/10.1016/j.ghm.2024.04.003. (in Press)

Lengliné, O., Schmittbuhl, J., Drif, K., et al. The largest induced earthquakes during the GEOVEN deep geothermal project, Strasbourg, 2018-2022: From source parameters to intensity maps. Geophysical Journal International, 2023, 234(3): 2445-2457.

Li, B., Wang, W., Dong, J. Locating single-point sources from arrival times containing large picking errors (LPEs): the virtual field optimization method (VFOM). Scientific Reports, 2016, 6(1): 19205.

Li, H., Shi, P., Tang, M. Discussions on fracture factors of brittle materials under the shear-compression condition. Advanced Materials Research, 2011, 250-253: 90-94.

Lynch, R., Malovichko, D. Seismology and slope stability in open pit mines. Paper Presented at International Symposium on Stability of Rock Slopes in Open Pit Mining and Civil Engineering, Cape Town, South Africa, January, 2006.

Ma, X., Zhang P., Zhang, S., et al. A tensorized transformer for language modeling. Paper Presented at the 33rd International Conference on Neural Information Processing Systems, Vancouver, Canada, 8-14 December, 2019.

McGarr, A., Bekins, B., Burkardt, N., et al. Coping with earthquakes induced by fluid injection. Science, 2015, 347(6224): 830-831.

Petit, J. P., Barquins, M. Can natural faults propagate under Mode II conditions? Tectonics, 1988, 7(6): 1243-1256.

Qian, Y., Chen, X., Luo, H., et al. An extremely shallow M_w 4.1 thrust earthquake in the eastern Sichuan Basin (China) likely triggered by unloading during infrastructure construction. Geophysical Research Letters, 2019, 46(23): 13775-13784.

Reza, S., Ferreira, M. C., Machado, J. M., et al. A multi-head attention-based transformer model for traffic flow forecasting with a comparative analysis to recurrent neural networks. Expert Systems with Applications, 2022, 202: 117275.

Shiotani, T., Ohtsu, M., Ikeda, K. Detection and evaluation of AE waves due to rock deformation. Construction and Building Materials, 2001, 15(5): 235-246.

Song, Z., Konietzky, H., Herbst, M. Bonded-particle model-based simulation of artificial rock subjected to cyclic loading. Acta Geotechnica, 2019, 14(4): 955-971.

Song, Z., Zhang, Z., Zhang, G., et al. Identifying the types of loading mode for rock fracture via convolutional neural Networks. Journal of Geophysical Research: Solid Earth, 2022, 127(2): e2021JB022532.

Vaswani, A., Shazeer, N., Parmar, N., et al. Attention is all you need. Paper Presented at the 31st International Conference on Neural Information Processing Systems, Red Hook, NY, USA, 4-9 December, 2017.

Wang, G., Shen, J., Liu, S., et al. Three-dimensional modeling and analysis of macro-pore structure of coal using combined X-ray CT imaging and fractal theory. International Journal of Rock Mechanics and Mining Sciences, 2019, 123: 104082.

Wang, L., Kwiatek, G., Bohnhoff, M., et al. Injection-induced fault slip and associated seismicity in the lab: Insights from source mechanisms, local stress states and fault geometry. Earth and Planetary Science Letters, 2024, 626: 118515.

Wang, Q., Hu, X., Zheng, W., et al. Mechanical properties and permeability evolution of red sandstone subjected to hydro-mechanical coupling: Experiment and discrete element modelling. Rock Mechanics and Rock Engineering, 2021a, 54(5): 2405-2423.

Wang, R., Ao, J., Zhou, L., et al. Multi-view self-attention based transformer for speaker recognition. Paper Presented at IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Singapore, 22-

27 May, 2022.

Wang, X., Wang, E., Liu, X., et al. Failure mechanism of fractured rock and associated acoustic behaviors under different loading rates. Engineering Fracture Mechanics, 2021b, 247: 107674.

Wang, Y., Zhou, X., Xu, X. Numerical simulation of propagation and coalescence of flaws in rock materials under compressive loads using the extended non-ordinary state-based peridynamics. Engineering Fracture Mechanics, 2016, 163: 248-273.

Watanabe, T., Nishibata, S., Hashimoto, C., et al. Compressive failure in concrete of recycled aggregate by acoustic emission. Construction and Building Materials, 2007, 21(3): 470-476.

Wu, W. A review of unloading-induced fault instability. Underground Space, 2021, 6(5): 528-538.

Yang, Y., Liang, C., Liu, F., et al. Predicting carbonate rock dissolution using multi-scale residual neural networks with prior knowledge. Gas Science and Engineering, 2024, 124: 205268.

Zhang, Y., Liu, S., Kou, M., et al. Mechanical and failure characteristics of fissured marble specimens under true triaxial compression: Insights from 3-D numerical simulations. Computers and Geotechnics, 2020, 127: 103785.

Zhan, T., Chen, h., Li, L. Numerical study of in situ acoustic emission monitoring for small-scale hydraulic fracturing, Journal of Geophysics and Engineering, 2022, 19(4): 615-629.

Zhou, W., Liao, S., Men, Y. Effect of localized water pressure on mountain tunnels crossing fracture zone. Transportation Geotechnics, 2021, 28: 100530.

Zhou, X., Wang, Y. State-of-the-art review on the progressive failure characteristics of geomaterials in peridynamic theory. Journal of Engineering Mechanics, 2021, 147(1): 03120001.

Zhuang, L., Kim, K., Jung, S., et al. Cyclic hydraulic fracturing of pocheon granite cores and its impact on breakdown pressure, acoustic emission amplitudes and injectivity. International Journal of Rock Mechanics and Mining Sciences, 2019, 122: 104065.

Zoback, M. D., Gorelick, S. M. Earthquake triggering and large-scale geologic storage of carbon dioxide. Proceedings of the National Academy of Sciences of the United States of America, 2012, 109(26): 10164-10168.

Zöller, G., Hainzl, S. Seismicity scenarios for the remaining operating period of the gas field in Groningen, Netherlands. Seismological Research Letters, 2023, 94(2A): 805-812.