

ARTICLE

Computational efficiency of deep learning-based seismic fault interpretation considering context window and input resolution

Bowen Deng¹, Guangui Zou^{1,2*}, Suping Peng^{1,2}, Chengyang Han¹, and Jingwen Xue¹¹College of Geoscience and Surveying Engineering, China University of Mining and Technology-Beijing, Beijing, China²State Key Laboratory for Fine Exploration and Intelligent Development of Coal Resources, China University of Mining and Technology-Beijing, Beijing, China

Abstract

The application of deep learning to seismic fault interpretation is often constrained by computational costs. To address this, we propose a novel strategy that decouples the context window size (the spatial range of seismic observations) from the input resolution (the actual matrix dimensions fed into the model), systematically investigating their combined impact on computational efficiency and interpretation accuracy. Using field-acquired seismic data from two distinct coal mines, we trained a lightweight two-dimensional convolutional neural network (CNN) on samples extracted with varying context windows (8×8 to 64×64 pixels), which are then uniformly resized to a fixed low resolution of 8×8 pixels. Our results demonstrate that enlarging the context window consistently improved model performance, with the 64×64 window achieving the highest precision (99.46%) and fault continuity, even after downscaling. In contrast, a combined multi-scale training set did not outperform the best single-window model, indicating that effective multi-scale fusion requires more advanced architectural designs. Our workflow highlights that contextual information remains crucial for feature learning despite input standardization, and offers an efficient paradigm: large context window + small fixed input + lightweight network, that maintains high accuracy while significantly reducing computational costs. This approach provides a practical pathway for deploying deep learning models in resource-limited geophysical applications.

***Corresponding author:**Guangui Zou
(zgg@cumtb.edu.cn)

Citation: Deng B, Zou G, Peng S, Han C, Xue J. Computational efficiency of deep learning-based seismic fault interpretation considering context window and input resolution. *J Seismic Explor.* 2026;35(2):026030008. doi: 10.36922/JSE026030008

Received: January 14, 2026**Revised:** March 2, 2026**Accepted:** March 9, 2026**Published online:** April 22, 2026

Copyright: © 2026 Author(s). This is an Open-Access article distributed under the terms of the Creative Commons Attribution License, permitting distribution, and reproduction in any medium, provided the original work is properly cited.

Publisher's Note: AccScience Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Keywords: Seismic fault interpretation; Deep learning; Convolutional neural network; Computational efficiency; Lightweight modeling

1. Introduction

The discipline of geosciences has undergone a remarkable transformation driven by the development of computer science and technology. From initial mathematical treatments of seismic data to today's deep learning algorithms, the computational influence on geophysical science has progressively intensified, creating a paradigm shift in how we understand and visualize geological structures.¹⁻³

Geosciences community commenced adopting machine learning technology in analytical applications around the late 1980s. This era witnessed the emergence of various shallow machine learning methods, including support vector machines,⁴⁻⁷ decision trees,⁸ random forests,⁹ Gaussian processes,¹⁰ and artificial neural networks,¹¹ which would become essential tools in the geophysicists' toolkit. These traditional machine learning algorithms excelled primarily in solving classification and regression problems prevalent in geophysical exploration. The back propagation algorithm for training neural networks became particularly influential, enabling the development of models that could identify complex nonlinear relationships in geophysical data. Despite these advances, the machine learning applications of this era remained largely feature-driven rather than physics-driven, focusing on statistical pattern recognition without deeply incorporating the underlying physical laws governing wave propagation, rock mechanics, or fluid dynamics.

While the conceptual foundations of neural networks, including convolutional neural networks (CNN), were laid as early as the 1980s,¹² their widespread application in geophysics remained limited for decades due to constraints in computational power and data availability. The true transformative period for deep learning arrived in the 2010s, marked by a confluence of factors: the rise of big data, powerful parallel computing hardware, and algorithmic refinements. The catalyst for this transformation was the groundbreaking success of AlexNet¹³ in the 2012 ImageNet competition, which dramatically demonstrated the capability of deep CNNs for complex image recognition tasks, which resonated powerfully with geophysicists, who recognized the structural analogies between seismic data and natural photographs, sparking an explosion of interest in applying deep CNN architectures to geophysical problems.¹⁴

According to a review by An *et al.*,¹⁵ deep learning-based methods for seismic fault identification entered a phase of rapid development in 2018 and have yielded a series of notable achievements. Deep learning eliminates the complex feature engineering and manual parameter tuning of traditional machine learning methods, achieving superior generalization that automatically learn feature representations from raw data. Current implementations primarily adopt two strategies:

- (1) **Image classification:** Dividing images into two classes based on the presence or absence of faults, labeled as positive (fault) and negative (non-fault) for binary classification training.¹⁶
- (2) **Image segmentation:** Repurposing the U-net model originally designed for medical image segmentation,¹⁷

for seismic data profiles, instead of generating probability maps by computing probabilities for individual points, these methods are primarily built on synthetic seismic data. However, they face limitations in transferring to real-world data with more complex geological conditions.¹⁸⁻²¹

Although deep learning offers significant benefits for fault interpretation, the approach confronts a number of challenges. As further investigated in the review by An *et al.*, these challenges can be broadly grouped into four categories:

- (1) **Data-related challenges:** including insufficiency of annotated datasets, the poor quality of seismic datasets, uncertainties of seismic fault interpretation, and the paradox between adopting synthetic seismic datasets and reflecting real-field seismic characteristics.
- (2) **Deep learning-related challenges:** deep learning models spend high computational costs, black-box models lack of interpretability, unbalanced ratio between positive and negative training samples.
- (3) **Evaluation-related challenges:** lack of evaluation criteria for fault interpretation algorithms, lack of open access for data, code, or workflows.
- (4) **Practical use-related challenges:** exclusive to deep learning experts and not friendly to geologists; overly emphasizes seismic data and omit known geological constraints; fault probability maps are not equal to fault interpretations.¹⁵

Given that this research employs actual-acquired seismic data, which lacks the perfect fault contours of synthetically generated laboratory data, the image classification approach was adopted over image segmentation. Based on above considerations, this study focuses on computational efficiency. By optimizing actual-acquired seismic data from mining areas and leveraging the characteristics of black-box models, we aim to reduce the computational demands of both training and prediction procedure.

This study is basically based on seismic sections and focuses on pixel-based fault identification around coal seams. The foundational methodology has been introduced in our prior work.²² In our previous study, we developed a two-dimensional (2D) CNN-based approach for seismic fault interpretation, formulating the fault identification task around coal seams as an image classification problem. Linear annotations were transformed into representative pixel points, and surrounding 2D matrix samples were extracted to construct the training dataset. By incorporating data augmentation strategies developed in the previous study and optimizing the sample selection process, the model demonstrated enhanced generalization capability and high recognition accuracy when applied to real-world

seismic data, validating the feasibility and efficiency of this pixel-based sampling methodology for seismic fault detection. Building upon this algorithm, this study further investigates remaining challenges which affect CNN-based seismic interpretation, such as computational efficiency and the impact of high-resolution seismic image noise.^{23,24}

2. Overview of the study area

The study area of this research comes from two distinct coalfields: the Renjiazhuang Coal Mine located in Lingwu, Ningxia in Northwestern China, and the Yanjiazhuang Coal Mine located near the border of Shouyang and Yangquan, Shanxi Province in North China.

The Renjiazhuang Coal Mine is situated on the western edge of the Ordos Plateau's fold-and-thrust belt (Figure 1), specifically in the Majiatan–Tianshuibu section, which is characterized by low, gentle semi-desert hills. The terrain is higher in the west and lower in the east, with complex topography in the western part and flat terrain in the east. The geological structure is complex, with the development of the Sandaogou anticline, Huangcaogou anticline, Huangcaogou syncline, and Shagou syncline. The folds predominantly exhibit NNW or SN direction. The coal-bearing strata belong to the Taiyuan Formation and Shanxi Formation, consisting of mudstones, siltstones, limestones, sandstones, coal seams, and coal layers from the transitional marine-continental facies. There are more than 20 coal seams, with thicknesses ranging from 8.03 to 33.85 meters, averaging 20.94 meters. The #9 coal seam is stably mineable across the entire area, with thickness ranging from 0.73 to 7.84 meters, averaging 5.12 meters. The coal seams contain one to two interlayers of impurities. The coal body structure is primarily cataclastic coal and granular coal.

The Yanjiazhuang Coal Mine is situated in the northern Qinshui Basin, near the Shouyang–Yangquan area (Figure 2). The region is characterized by an annular slope with superimposed secondary folds and faults of various orientations and ages. The main coal-bearing strata are the Upper Carboniferous Taiyuan Formation and the Lower Permian Shanxi Formation, with more than 10 coal seams, including the #3 and #15 coal seams as the main coal seams.

The coal seams included in this study are the #3, #5, and #9 coal seams from the Renjiazhuang Coal Mine, and the #3, #9, and #15 coal seams from the Yanjiazhuang Coal Mine. The 3D seismic data comprehensively cover both mining areas, with the horizontal grid corrected to an equidistant interval and the vertical scale set at equal time intervals, with a horizontal resolution of 5 meters, and a vertical time sampling interval of 1 millisecond. The seismic data from the Renjiazhuang area consist of 506,541

seismic traces, covering an area of 12.66 square kilometers, with 2,001 data points per trace (ranging from 0 to 2000 ms). The seismic data from the Yanjiazhuang area consist of 577,171 seismic traces, covering an area of 14.43 square kilometers, with 1,501 data points per trace (ranging from 0 to 1500 ms).

The two coalfield regions encompass distinct geological types, each exhibiting high representativeness. The samples derived from these areas capture fault characteristics that are relatively general.

Fault locations are first annotated on multiple sections across both coalfields, covering inline and crossline directions. Partial examples of section labeling are illustrated in Figure 3. (Note that these figures do not display complete sections from the original dataset).

3. Methods

3.1. Overview of prior research

The methodology of this study builds upon our prior work,²² which established a 2D convolutional neural network (CNN)-based approach for seismic fault interpretation. In that previous study, we formulated the task of fault identification around coal seams as an image classification problem. The core innovation was a 'pixel thinking' strategy, where manually drawn linear fault annotations on 2D seismic sections were converted into discrete, representative pixel points ('spot samples'). Two-dimensional matrix samples were then extracted from the seismic data centered on these points, constructing a binary classification dataset (fault/non-fault) for CNN training.

The key workflow of the prior method consisted of:

- (1) Manual linear annotation of faults and target coal seams on seismic sections;
- (2) Pixelation of these linear labels into discrete point sets;
- (3) Extraction of 2D matrix samples around each point based on defined vertical/horizontal extraction ranges and step parameters;
- (4) Optimization of sample representativeness through techniques like applying an enhancement radius (to broaden the effective area of labels) and establishing a buffer zone (to mitigate ambiguity near label boundaries);
- (5) Control of the fault/non-fault sample of one-to-three ratio to reflect geological reality and balance the dataset;
- (6) Design and training of a 2D CNN model comprising convolutional, pooling, and fully connected layers for final classification.

This methodology was empirically validated using

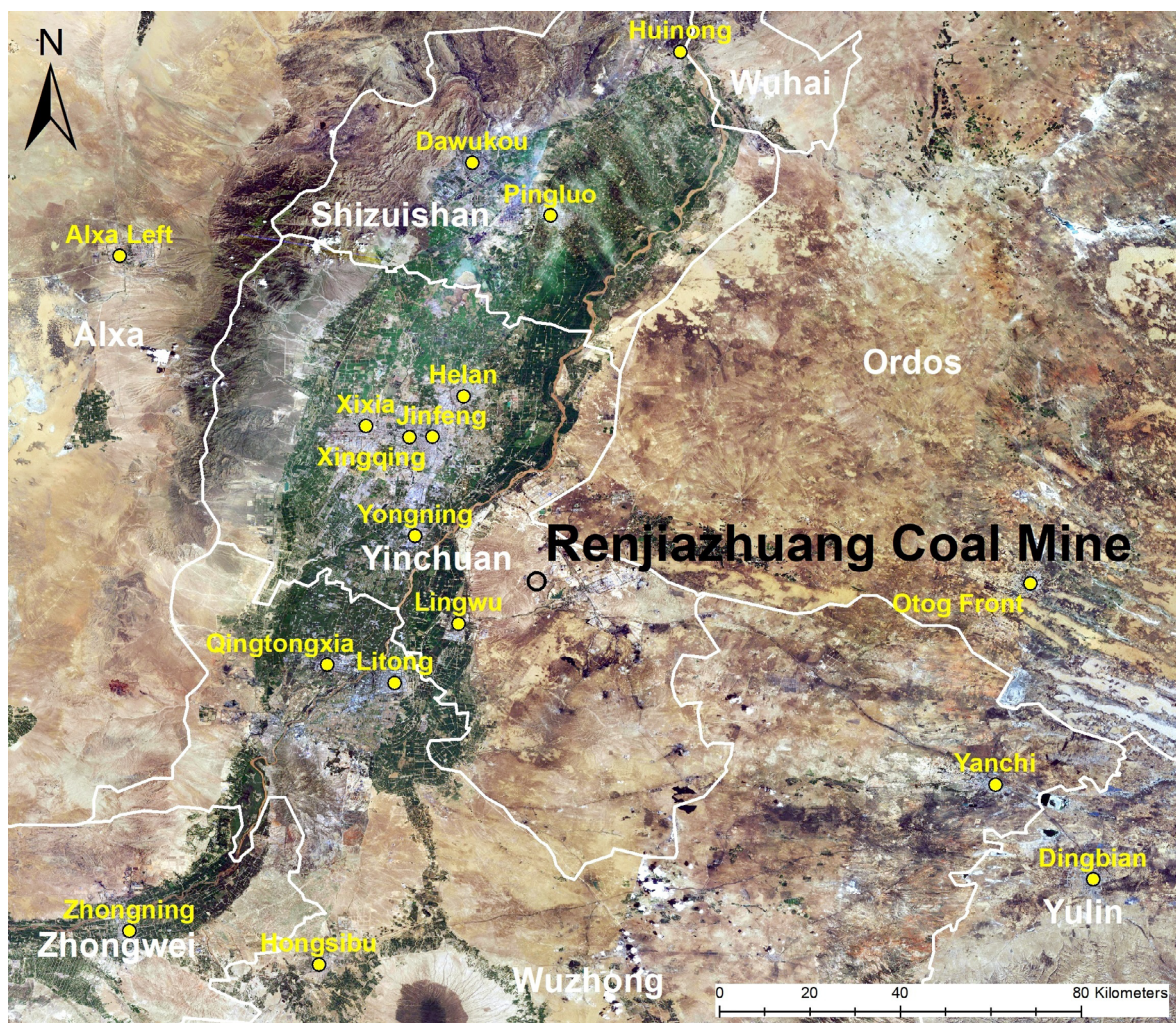


Figure 1. Location of Renjiazhuang Coal Mine

real-field seismic data. The model demonstrated high recognition accuracy and robust generalization on independent seismic sections, confirming the feasibility and effectiveness of this linear-annotation and pixel-based 2D CNN approach for seismic fault detection. The complete algorithmic workflow, sample construction rules, and model architecture are detailed in the prior publication.

3.2. Focus of the present study

The methodology established in our prior research requires the extraction of a substantial number of 2D samples along the orientation of coal seams to construct the training dataset. Although these samples represent different central sampling locations, significant overlap exists among them, resulting in a large volume of training data and considerable computational cost. Furthermore,

training the model with larger input matrix dimensions necessitates correspondingly more complex neural network architectures, which typically involve greater numbers of parameters and increased demand for computational resources. Similar challenges arise during the prediction phase.

This presents a practical trade-off: enlarging the window size for extracting 2D arrays implies a heavier computational burden, while reducing computational costs would require compressing the sample dimensions. Consequently, a key question emerges: Is it possible to minimize the dimensions of the extracted 2D matrix samples while maintaining flexibility in the observational context? Specifically, we propose to first extract data using a variable-sized observational window (e.g., 8×8 , 16×16 , 32×32 pixels) and then uniformly resize these windows to a smaller, fixed dimension (e.g., 8×8) during

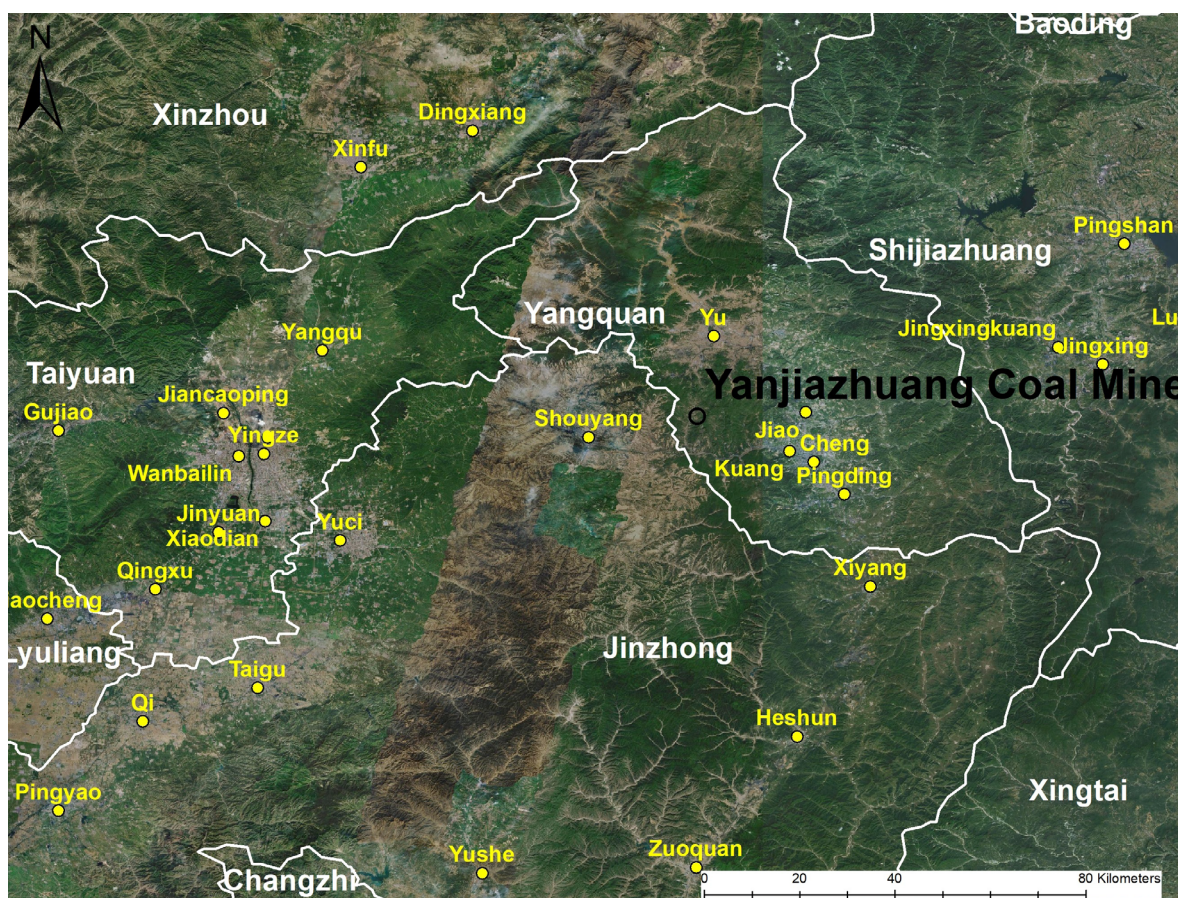


Figure 2. Location of Yanjiazhuang Coal Mine

data preprocessing before putting them into the model for training. This resizing procedure is illustrated in Figure 4.

Owing to the inherent lack of interpretability in deep neural networks, the internal mechanisms by which the model learns discriminative features from positive and negative samples remain opaque. If models trained on samples from different original window sizes but resized to the same small input dimension yield comparable performance, it would suggest that the input matrix size and the inherent characteristics of the sample data have reached the upper limit of the network's learning capacity under the given representation. Conversely, if significant performance variations are observed, it would indicate that even with a fixed small input size, adjusting the spatial context (via the original window size) can meaningfully influence the learning outcome, thereby offering a pathway to optimize model efficacy through careful design of the observational context.

Thus, this study aims to systematically investigate the interplay between context window size (pre-resizing)

and input resolution (post-resizing) and in influencing computational efficiency and model performance, providing insights into the optimal balance between data representation and resource consumption for deep learning-based seismic fault interpretation.

The proposed downscaling strategy, while essential for achieving a fixed low-dimensional input and high computational efficiency, inevitably alters the frequency content of the original seismic data. In principle, resizing to 8×8 pixels may attenuate certain high-frequency components, including both fine fault-related details and high-frequency noise. However, from a data-driven perspective, we posit that a CNN can learn robust discriminative features from such low-resolution inputs if provided with sufficiently rich contextual information from an expanded observation window. The effectiveness of this trade-off and robustness of model, compensating for potential local detail loss with broader spatial context, will be empirically validated by the model performance achieved with varying context window sizes.

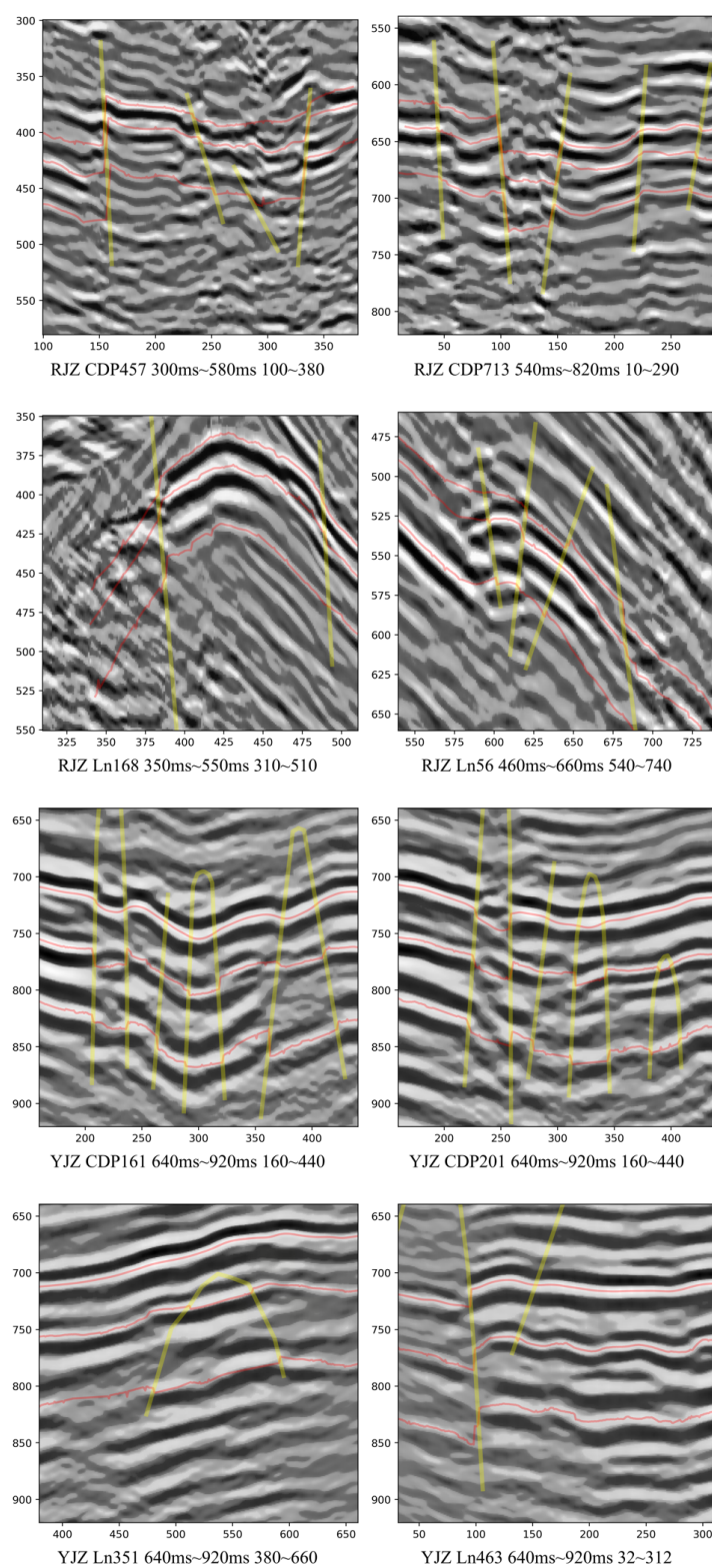


Figure 3. Example of seismic sections and labeling from the two coal mine areas
Abbreviations: RJZ: Renjiazhuang; YJZ: Yanjiazhuang.

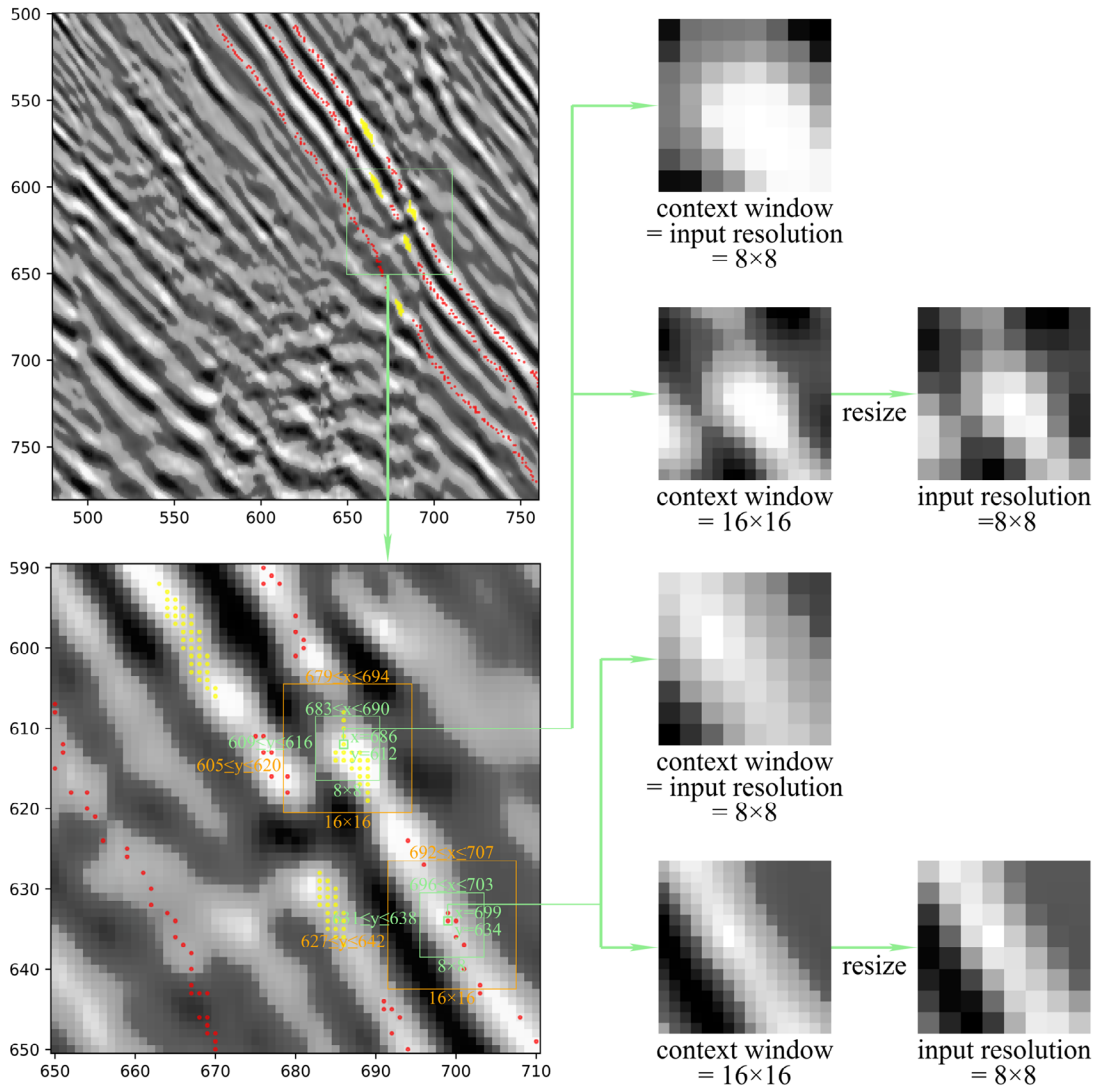


Figure 4. Sample extraction from a seismic section. Matrices centered on fault (positive, yellow dots) and non-fault (negative, red dots) points are clipped from different context window sizes (8×8 and 16×16) and then resized into a uniform size of 8×8 as input resolution.

3.3. Resize algorithm: bicubic interpolation

To resize a 2D matrix into a different size, an interpolation method must be employed. Common interpolation techniques include nearest neighbor, bilinear interpolation, and bicubic interpolation.²⁵⁻²⁷ Their basic principles and mathematical formulations are described below.

Nearest neighbor interpolation is a simple method that assigns the value of the closest pixel in the original image to the target pixel. It is computationally efficient but often leads to jagged edges and noticeable loss of detail. The operation can be expressed as:

$$I'(x', y') = I(\text{round}(x), \text{round}(y)) \quad (1)$$

where (x, y) are the original image coordinates, (x', y') are the target image coordinates, and $\text{round}(\cdot)$ denotes rounding to the nearest integer.

Bilinear interpolation performs a weighted average of the four nearest neighboring pixels. It first interpolates linearly in the horizontal direction, then in the vertical direction, resulting in smoother output compared to nearest neighbor. The formula is given by:

$$I'(x', y') = (1 - \alpha)(1 - \beta)I(x_0, y_0) + \alpha(1 - \beta)I(x_1, y_0) + (1 - \alpha)\beta I(x_0, y_1) + \alpha\beta I(x_1, y_1) \quad (2)$$

where $\alpha = x' - x_0$, $\beta = y' - y_0$, and (x_0, y_0) , (x_1, y_0) , (x_0, y_1) , (x_1, y_1) are the four surrounding pixel coordinates.

However, it may still blur high-frequency details.

Bicubic interpolation extends the support region to 16 neighboring pixels. It uses a cubic polynomial in two dimensions, taking into account not only the pixel values but also the local gradients, which helps preserve edges and fine textures more faithfully than lower-order methods. The general form is:

$$I'(x', y') = \sum_{i=-1}^2 \sum_{j=-1}^2 I(x_i, y_j) \cdot R(\alpha - i) \cdot R(\beta - j) \quad (3)$$

where $R(t)$ is the cubic basis function, defined as:

$$R(t) = \begin{cases} 1.5|t|^3 - 2.5|t|^2 + 1, & 0 \leq |t| < 1 \\ -0.5|t|^3 + 2.5|t|^2 - 4|t| + 2, & 1 \leq |t| < 2 \\ 0, & 2 \leq |t| \end{cases} \quad (4)$$

To facilitate an intuitive understanding, the principles and operation of the three interpolation methods is illustrated in both one and two-dimensional forms in Figure 5.

In this study, bicubic interpolation was employed to resize the context windows to the fixed low resolution. This method was selected because it preserves more detail and provides smoother transitions compared to nearest-neighbor or bilinear interpolation, thereby better maintaining the original numerical characteristics of the seismic data. This helps ensure that the resized samples retain discriminative features essential for effective model learning.

3.4. Deep learning architecture: A lightweight structure

To align with the reduced 8×8 input size, a lightweight CNN is designed for efficient binary classification. The architecture is compact, containing only several thousand parameters, far fewer than typical CNNs used in standard image recognition.

This lightweight design is a direct consequence of the input standardization strategy (resizing all context windows to 8×8), and is based on the premise that effective fault discrimination does not require deep, complex networks when discriminative information is concentrated into a low-dimensional representation. The simple architecture is therefore intended to ensure high computational efficiency in both training and prediction, aligning with the core objective of this study.

The architecture is shown in Figure 6. Training was configured with the following hyperparameter settings: optimizer = Adam (Adaptive Moment Estimation), learning rate = 0.001, loss function = binary crossentropy loss, batch size = 4096, epochs = 10,000, no batch normalization was used in the network architecture.

3.5. Proposed workflow

Building upon the aforementioned rationale, we designed the following experimental workflow. First, we specified a series of context window sizes (*i.e.*, extraction side lengths of 8, 16, 32, and 64 pixels). For each window size, samples were extracted accordingly and then resized to a uniform dimension of 8×8 pixels to serve as training data.

Additionally, we experimented with combining samples from various context windows into a single training dataset, to examine whether the fractal concept of integrating multi-scale information from micro to macro perspectives, could be effectively applied to seismic data feature extraction.

This means that five models will be trained: the first four corresponding to individual context windows, and the last one trained on the combined dataset from all window sizes, namely, Model 8, Model 16, Model 32, Model 64, and Model 8 + 16 + 32 + 64.

During the prediction phase, data were similarly extracted using the predefined context window size, resized to the same 8×8 format, and put into the corresponding model to compute classification probabilities. These probabilities were subsequently aggregated to generate fault distribution maps. Finally, we evaluated the performance of each model to assess whether even under an identical input resolution, different context window sizes can yield varying effects on interpretation outcomes.

The proposed workflow is shown in Figure 7.

4. Results

4.1. Model evaluation by metrics

The training and evaluation metrics for the five models are presented in Table 1. These metrics were computed using datasets aligned with each model's training configuration. From Model 8 (8×8 observation range) to Model 64 (64×64 range), a progressive improvement in accuracy is evident, demonstrating that larger observation ranges enhance model performance. For instance, Model 64 achieved a precision of 99.46%, significantly outperforming smaller-scale models. In contrast, the combined multi-scale model (8 + 16 + 32 + 64) exhibited the lowest precision (89.31%).

4.2. Model evaluation by visualization

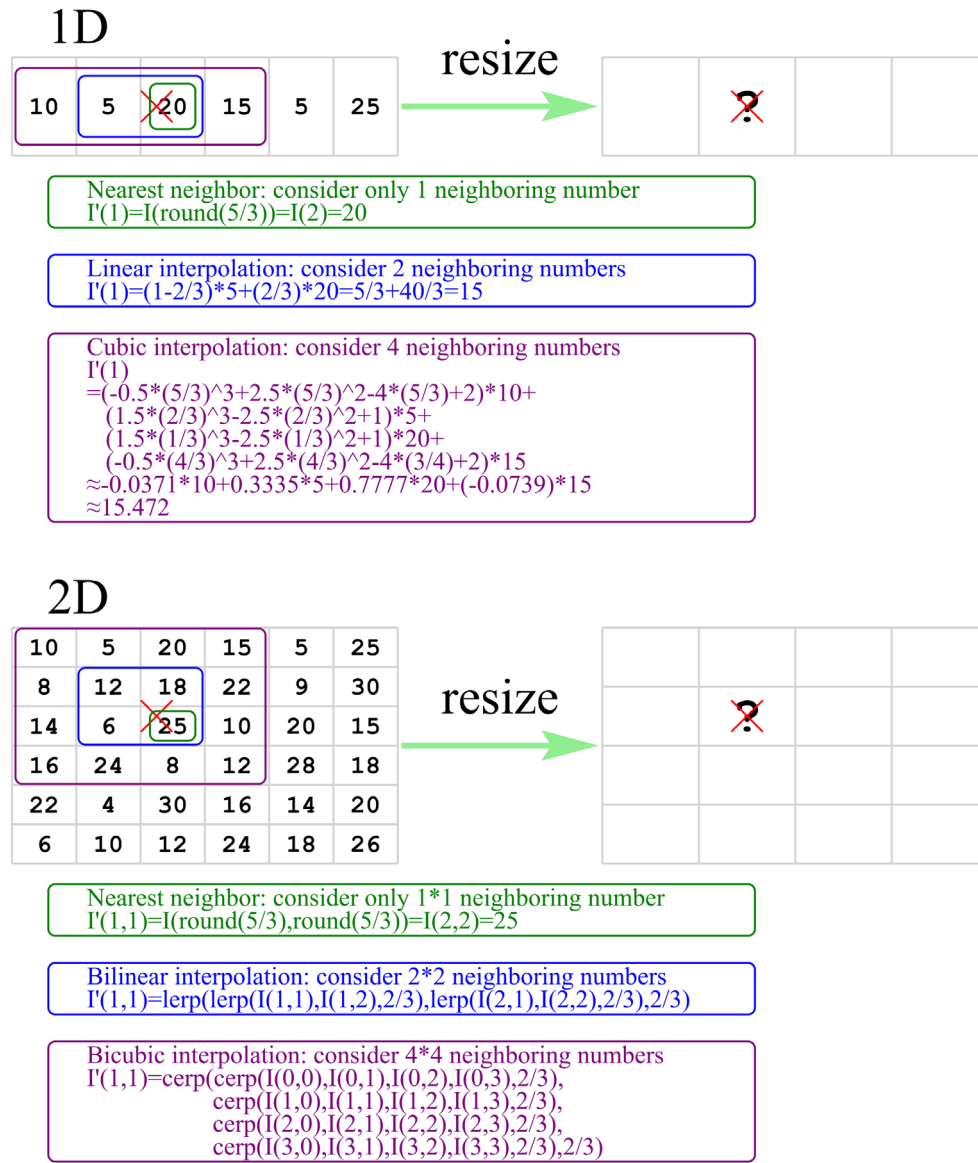


Figure 5. Demonstration of interpolation methods for one- and two-dimensional array resizing

Table 1. Training and evaluation metrics of five models

Model	8	16	32	64	8 + 16 + 32 + 64
Context window	8 × 8	16 × 16	32 × 32	64 × 64	Combined
Training loss	0.201	0.0936	0.0335	0.018	0.2415
Training accuracy	0.9198	0.9636	0.9877	0.9935	0.8977
Validation loss	0.3101	0.1308	0.0588	0.0359	0.2208
Validation accuracy	0.8774	0.9531	0.9844	0.9921	0.9086
Precision	0.9429	0.9811	0.9904	0.9946	0.8931
Recall	0.8436	0.9741	0.9963	0.9986	0.8015
F1 Score	0.8905	0.9776	0.9933	0.9966	0.8448

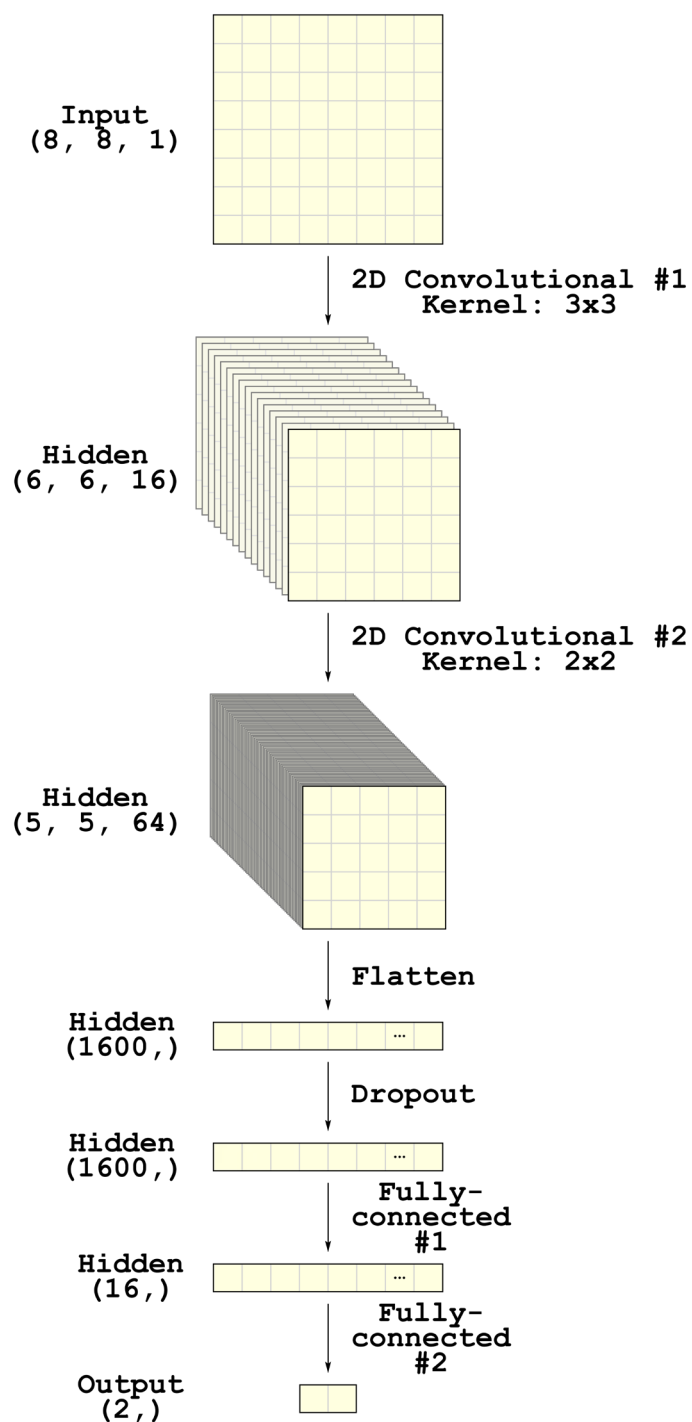


Figure 6. Lightweight deep convolutional architecture designed for this study. The model accepts an $8 \times 8 \times 1$ input matrix and outputs a binary classification result.

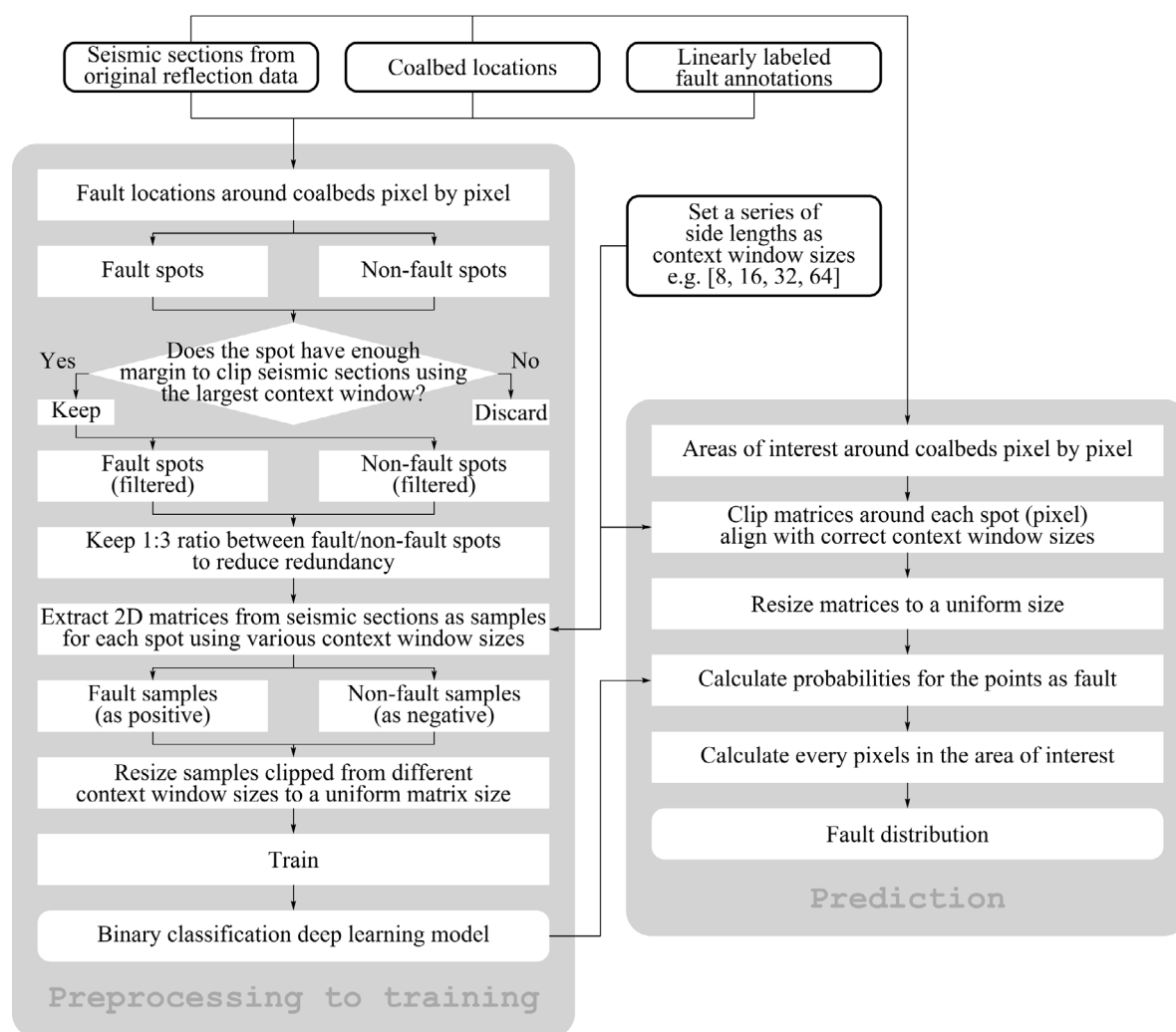


Figure 7. Proposed workflow, covering data preprocessing, model training, and prediction phases

Besides quantitative metrics, we also applied the trained models directly to seismic sections to compute fault/non-fault probabilities and generate corresponding distribution maps around the target coal seams. This visual assessment provides a practical evaluation of each model's performance. The resulting probability maps from applying different models to several test sections are shown in each subplot of Figure 8. Among them, the prediction result of the combined model (8+16+32+64) represents the averaged probability distribution computed from samples of all four different context window sizes fed simultaneously.

Visual assessment of the predicted fault distributions on seismic sections (Figure 8) indicates that Model 64 achieves the best continuity and closest agreement with the manual annotations. This observation aligns with the

pattern of accuracy reflected in the quantitative metrics (Table 1).

4.3 Efficiency comparison of different input resolution

The central objective of this study is to improve computational efficiency in deep learning-based seismic fault interpretation. In principle, reducing the input resolution (*i.e.*, the size of the matrix fed into the model) can significantly decrease the number of network parameters, accelerate training, and reduce inference time during prediction. To quantitatively assess this efficiency gain, we compare the computational costs associated with different input resolutions. Specifically, we consider the scenario where models are designed natively for input sizes of 8×8 , 16×16 , 32×32 , and 64×64 pixels, without the context-

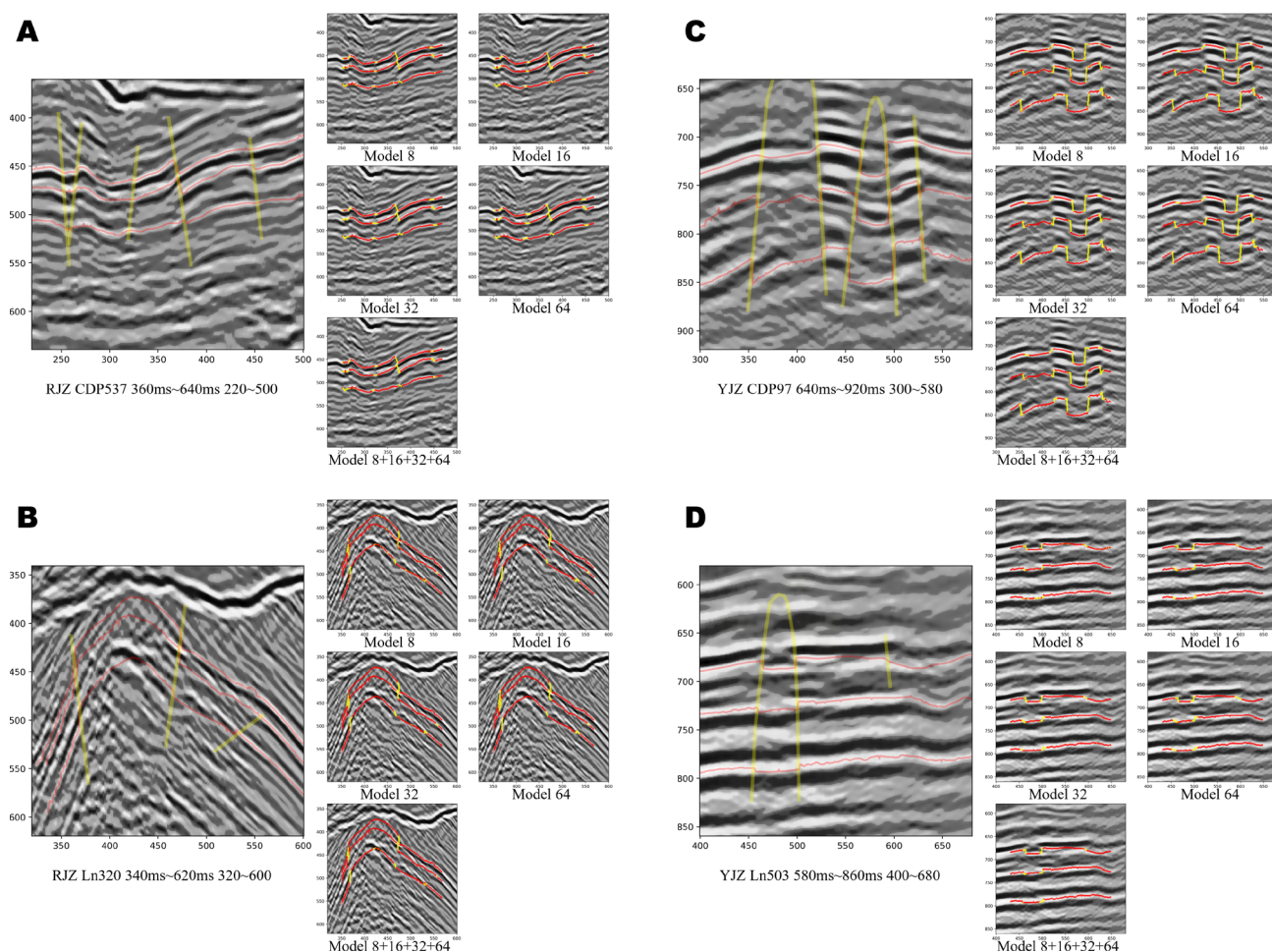


Figure 8. Four seismic sections from the two coal mine areas. (A) Renjiazhuang CDP537 360ms~640ms 220~500. (B) Renjiazhuang Ln320 340ms~620ms 320~600. (C) Yanjiazhuang CDP97 640ms~920ms 300~580. (D) Yanjiazhuang Ln503 580ms~860ms 400~680. Each panel includes original manual annotations (left) and fault distribution maps from the prediction results of various models (right).

window resizing strategy. The comparison focuses on three key metrics: the total number of trainable parameters in the CNN, the time required to complete 10,000 training epochs (the latter three did not complete the entire process, estimated based on the speeds of the previous epochs), and the average inference time per sample. The results are summarized in Table 2.

The 8×8 model requires only ~3k parameters, far fewer than the 64×64 model (~200k), leading to substantially faster training and inference. However, as shown in sections 4.1 and 4.2, using only a small native input window reduces accuracy. Our proposed strategy, extracting large context windows (e.g., 64×64) and resizing them to a small fixed input (8×8), retains rich contextual information while operating within the highly efficient computational regime of the 8×8 model. This approach achieves high interpretation performance with

significantly lower computational cost.

5. Discussion

The experimental results of this study demonstrate that the context window size is a critical factor influencing model performance. From 8×8 to 64×64 , as the original spatial range used for sample extraction increases, the model's performance metrics show systematic improvement in both training and validation, with the largest window model ultimately achieving the highest accuracy and continuity in fault identification. This trend clearly indicates that even when all samples are uniformly resampled to the same low resolution (8×8) before being fed into the network, the scale information of the original geological structures they represent is still effectively preserved and conveyed to the model. A larger context window provides the model with richer spatial structure and contextual relationships,

Table 2. Computational efficiency comparison for models with different native input resolutions

Input resolution	Model parameters	Training time (10,000 epochs)	Inference time per sample (ms)
8 × 8	~3 k	~15.6 h	~0.05
16 × 16	~12.5 k	~62 h	~0.18
32 × 32	~50 k	~240 h	~0.65
64 × 64	~200 k	~936 h	~2.50

Note: Training times are estimated based on a single NVIDIA GTX 1650Ti GPU. Inference times are averaged over 100,000 samples.

enabling it to capture the overall morphology and seismic response characteristics of faults over a broader area, rather than isolated local anomalies. This confirms our core hypothesis: under the constraint of a fixed input resolution, adjusting the original observation window size for sample extraction allows for flexible control over the spatial context information of the input data, thereby significantly influencing the learning outcome and generalization capability of the deep learning model.

However, the combined model trained on samples from multiple window sizes (8 + 16 + 32 + 64) did not outperform the single best-window model, since its accuracy was notably lower. This suggests that simply mixing data from different scales did not automatically induce the anticipated “multi-scale feature fusion” advantage within the lightweight network architecture used here. A potential reason is that samples from different scales, after resizing, may possess inherently different feature distributions, and a simple classifier might struggle to learn a unified and robust cross-scale representation from them unbiasedly. More sophisticated network designs, such as incorporating feature pyramids or dedicated fusion modules, or employing staged, scale-specific training strategies, could be key to effectively leveraging multi-scale information.

Regarding computational efficiency, the strategy of “fixed small input resolution + variable large context window” proposed in this study presents a promising trade-off. All models were trained and predicted using a lightweight network designed for 8 × 8 input, which fundamentally constrains computational complexity. The results demonstrate that while maintaining a lightweight computational core, extracting information from larger original windows can yield performance comparable to or even surpassing that of traditional models with large inputs. This offers a new perspective for deploying deep learning models in practical resource-constrained environments.

The validation of this method on seismic data from

two coal mines with distinct geological backgrounds indicates that the strategy based on pixelated samples and context window adjustment possesses a degree of geological generalizability. The models learned fault response patterns that are not confined to local features of a single area but are more universal. Compared to segmentation methods like U-Net, which require pixel-perfect annotations, this approach reduces the demand for annotation completeness, relying only on linear annotations. Compared to traditional machine learning methods based on handcrafted features, it avoids complex feature engineering. However, as an image classification approach, its output is discrete probability points. The continuity and morphological integrity of fault boundaries need to be reconstructed through post-processing of the probability maps, which remains a challenge when dealing with extremely complex or subtle fault systems.

Although deep learning models are often considered “black boxes”, this study, through the controlled experiment of varying context windows, provides an indirect perspective for understanding model decision-making. The phenomenon of improving performance with larger windows suggests that model decisions heavily rely on regional context beyond local pixels. This insight guides future interpretability research to focus on analyzing the model’s attention mechanisms toward features at different scales, thereby building a bridge from data scale to geological significance.

Therefore, the size of the context window is an important design parameter that connects the physical meaning of the data with the model’s learning capacity. In the deployment of lightweight models that prioritize computational efficiency, the deliberate design and utilization of large context windows is an effective and practical strategy for enhancing model performance. Future work could focus on developing more intelligent architectures for multi-scale information fusion and further exploring the applicability

of this method to more complex geological structures and 3D seismic volume interpretation.

Furthermore, the lightweight modeling philosophy explored in this study, balancing computational efficiency with performance through context window adjustment and fixed low-resolution input, holds promising implications beyond seismic fault interpretation. Similar challenges of resource constraints and the need for efficient, scalable models are prevalent in other data-intensive geoscience domains, such as remote sensing image analysis and cloud detection. Recent advances have demonstrated the effectiveness of lightweight architectures in these areas. For instance, lightweight networks integrating multi-scale observation ranges and adaptive channel-spatial attention mechanisms have achieved high-precision cloud detection in remote sensing imagery, demonstrating feasibility in resource-constrained environments.²⁸ DecoupleNet introduced multi-branch feature decoupling and hybrid downsampling strategies to preserve small-scale objects in remote sensing imagery, achieving state-of-the-art performance with reduced computational overhead.²⁹ Similarly, Xiao *et al.*³⁰ proposed an enhanced inter-layer correlation framework for drone-based small object detection, demonstrating that adaptive resizing and feature separation significantly improve efficiency without compromising accuracy. The strategy proposed here, leveraging large contextual windows while standardizing inputs to a minimal dimension, could be adapted to other fields where data volume and model complexity are limiting factors. This cross-domain applicability highlights the broader relevance of our approach and encourages further exploration of efficient deep learning frameworks in geoscience and beyond.

In this study, we show that computational efficiency can be significantly improved with little sacrifice in training accuracy. However, this study also has some limitations:

- (1) It lacks necessity to trade effectiveness for efficiency under the current dataset; however, for future applications with much larger datasets or more complex tasks, the proposed decoupling strategy offers a practical pathway to design more efficient models, where a controlled balance between efficiency and effectiveness may be meaningfully explored.
- (2) The context window size was explored up to 64×64 pixels. While results show consistent improvement up to this size, whether further enlarging the window continues to enhance performance remains an open question. The choice of 64×64 was based on practical computational considerations and the scope of this work, which focused on validating the decoupling concept rather than pursuing maximum possible

window sizes.

- (3) The method was validated on seismic data from two coal mines with relatively clear fault structures. Its performance in highly complex geological settings (*e.g.*, salt domes, thrust belts, regions with structurally complex faults) has not been tested. Generalizability to such environments would require further investigation, possibly involving architectural adjustments or incorporation of domain-specific constraints.

6. Conclusion

This study systematically investigated the impact of context window size and input resolution on deep learning-based seismic fault interpretation, with a focus on computational efficiency. The core findings demonstrate that context window size is a crucial design parameter that effectively bridges the scale of geological data with model learning capacity, even when input dimensions are standardized. The key conclusions are:

- (1) Under a fixed small input resolution (8×8 pixels), expanding the spatial context window (from 8×8 to 64×64) consistently and significantly improves model accuracy and the continuity of predicted faults. This confirms that richer contextual information, preserved despite downscaling, is critical for the model to learn robust, geologically-relevant features.
- (2) Simply combining multi-scale samples into a single training set did not yield superior performance within our lightweight architecture, indicating that effective multi-scale fusion requires more sophisticated network designs or training strategies beyond simple data aggregation.
- (3) The proposed strategy of “large context window + fixed small input resolution + lightweight network” establishes an efficient paradigm that maximizes computational performance per resource unit. By standardizing inputs to a minimal dimension (8×8), the network remains extremely compact, enabling significantly faster training and prediction cycles. Crucially, this architectural efficiency does not compromise accuracy, the model achieves superior interpretation quality by extracting richer geological context from larger original windows. This approach provides a practical and effective solution for deploying high-performance deep learning models in resource-constrained or time-sensitive scenarios.

Acknowledgments

None.

Funding

This work was supported by the National Key Research and Development Program of China (Grant number 2023YFB3211002) and the National Natural Science Foundation of China (Grant number 42274165).

Conflict of interest

Suping Peng is one of the Editor-in-Chiefs of this journal, but was not in any way involved in the editorial and peer-review process conducted for this paper, directly or indirectly. The authors declare they have no competing interests.

Author contributions

Conceptualization: Bowen Deng

Data curation: Guangui Zou, Suping Peng

Methodology: Bowen Deng

Software: Bowen Deng

Supervision: Suping Peng

Writing–original draft: Bowen Deng

Writing–review & editing: Guangui Zou, Suping Peng, Chengyang Han, Jingwen Xue

Availability of data

Seismic data, codes and neural networks associated with this research are available from the corresponding author upon reasonable request.

References

- Karpatne A, Ebert-Uphoff I, Ravela S, Babaie HA, Kumar V. Machine Learning for the geosciences: Challenges and opportunities. *IEEE Trans Knowl Data Eng.* 2019;31(8):1544-1554.
doi: 10.1109/TKDE.2018.2861006
- Lary DJ, Alavi AH, Gandomi AH, Walker AL. Machine learning in geosciences and remote sensing. *Geosci Front.* 2016;7(1):3-10.
doi: 10.1016/j.gsf.2015.07.003
- Yu S, Ma J. Deep learning for geophysics: Current and future trends. *Rev Geophys.* 2021;59(3):e2021RG000742.
doi: 10.1029/2021RG000742
- Di H, Shafiq MA, AlRegib G. Seismic-fault detection based on multiattribute support vector machine analysis. In: *SEG Technical Program Expanded Abstracts 2017*. Proceedings of the 87th SEG Annual International Meeting; September 24-29, 2017; Houston, Texas, USA. Society of Exploration Geophysicists; 2017:2039-2044.
doi: 10.1190/segam2017-17748277.1
- Ren K, Zou G, Zhang S, Peng S, Gong F, Liu Y. Fault identification and reliability evaluation using an SVM model based on 3-D seismic data volume. *Geophys J Int.* 2023;234(1):755-768.
doi: 10.1093/gji/ggad095
- Han C, Zou G, Yeh HG, Gong F, Shi S, Chen H. Intelligent fault prediction with wavelet-SVM fusion in coal mine. *Comput Geosci.* 2025;194:105744.
doi: 10.1016/j.cageo.2024.105744
- Ren K, Zou G, Peng S, Yeh HG, Deng B, Ji Y. Fault identification based on the kernel principal component analysis-genetic particle swarm optimization-support vector machine algorithm for seismic attributes in the Sihe Coal Mine, Qinshui Basin, China. *Interpretation.* 2022;11(1):T59-T73.
doi: 10.1190/int-2022-0039.1
- Li W, Yang C, Sun D. Mining geophysical parameters through decision-tree analysis to determine correlation with tropical cyclone development. *Comput Geosci.* 2009;35(2):309-316.
doi: 10.1016/j.cageo.2008.02.032
- Li D, Peng S, Lu Y, Guo Y, Cui X. Seismic structure interpretation based on machine learning: A case study in coal mining. *Interpretation.* 2019;7(3):SE69-SE79.
doi: 10.1190/INT-2018-0208.1
- Ray A, Myer D. Bayesian geophysical inversion with trans-dimensional Gaussian process machine learning. *Geophys J Int.* 2019;217(3):1706-1726.
doi: 10.1093/gji/ggz111
- Brown WM, Gedeon TD, Groves DI, Barnes RG. Artificial neural networks: A new method for mineral prospectivity mapping. *Aust J Earth Sci.* 2000;47(4):757-770.
doi: 10.1046/j.1440-0952.2000.00807.x
- LeCun Y, Boser B, Denker JS, et al. Backpropagation Applied to Handwritten Zip Code Recognition. *Neural Comput.* 1989;1(4):541-551.
doi: 10.1162/neco.1989.1.4.541
- Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. *Commun ACM.* 2017;60(6):84-90.
doi: 10.1145/3065386
- Bergen KJ, Johnson PA, de Hoop MV, Beroza GC. Machine learning for data-driven discovery in solid Earth geoscience. *Science.* 2019;363(6433):eaau0323.
doi: 10.1126/science.aau0323
- An Y, Du H, Ma S, et al. Current state and future directions for deep learning based automatic seismic fault interpretation: A systematic review. *Earth Sci Rev.* 2023;243:104509.
doi: 10.1016/j.earscirev.2023.104509
- Waldeland AU, Jensen AC, Gelius LJ, Solberg AHS. Convolutional neural networks for automated seismic

- interpretation. *Lead Edge*. 2018;37(7):529-537.
doi: 10.1190/tle37070529.1
17. Ronneberger O, Fischer P, Brox T. U-Net: Convolutional networks for biomedical image segmentation. In: Navab N, Hornegger J, Wells W, Frangi A, eds. *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015; Lecture Notes in Computer Science (9351)*. Proceedings of the MICCAI 2015, 18th International Conference; October 5-9, 2015; Munich, Germany. Springer; 2015:234-241.
doi: 10.1007/978-3-319-24574-4_28
 18. Huang J, Nowack RL. Machine learning using U-Net convolutional neural networks for the imaging of sparse seismic data. *Pure Appl Geophys*. 2020;177(6):2685-2700.
doi: 10.1007/s00024-019-02412-z
 19. Peters B, Haber E, Granek J. Neural networks for geophysicists and their application to seismic data interpretation. *Lead Edge*. 2019;38(7):534-540.
doi: 10.1190/tle38070534.1
 20. Wu X, Liang L, Shi Y, Fomel S. FaultSeg3D: Using synthetic data sets to train an end-to-end convolutional neural network for 3D seismic fault segmentation. *Geophysics*. 2019;84(3):IM35-IM45.
doi: 10.1190/geo2018-0646.1
 21. Cunha A, Pochet A, Lopes H, Gattass M. Seismic fault detection in real data using transfer learning from a convolutional neural network pre-trained with synthetic seismic data. *Comput Geosci*. 2020;135:104344.
doi: 10.1016/j.cageo.2019.104344
 22. Deng B, Zou G, Peng S, She J, Han C, Liu Y. An approach of 2D convolutional neural network-based seismic data fault interpretation with linear annotation and pixel thinking. *Geophys Prospect*. 2024;72(9):3350-3370.
doi: 10.1111/1365-2478.13606
 23. Li G, Li H, He S, Wang L. Multi-stage progressive network for seismic random noise suppression. *J Seism Explor*. 2025;34(1):43.
doi: 10.36922/jse025240011
 24. Wang W, Chen H, Chang D, Wang X, Wang S, Li D. Dual-branch dense network for seismic background noise elimination. *J Seism Explor*. 2025;34(5):53.
doi: 10.36922/jse025290038
 25. Parsania PS, Virparia PV. A comparative analysis of image interpolation algorithms. *Int J Adv Res Comput Commun Eng*. 2016;5(1):29-34.
doi: 10.17148/IJARCCCE.2016.5107
 26. Jakhetiya V, Kumar A, Tiwari AK. A survey on image interpolation methods. In: Jusoff K, Xie Y, eds. *Volume 7546, Second International Conference On Digital Image Processing*. Proceedings of the Second International Conference on Digital Image Processing; February 26-28, 2010; Singapore. Society of Photo-Optical Instrumentation Engineers (SPIE); 2010:75461T.
doi: 10.1117/12.855799
 27. Han D. Comparison of commonly used image interpolation methods. In: Proceedings of the 2nd International Conference on Computer Science and Electronics Engineering (ICCSEE 2013); Advances in Intelligent Systems Research Volume 34. 2nd International Conference on Computer Science and Electronics Engineering (ICCSEE 2013); March 22-23, 2013; Hangzhou, China. Atlantis Press; 2013:1556-1559.
doi: 10.2991/iccsee.2013.391
 28. Hu K, Zhang D, Xia M, Qian M, Chen B. LCDNet: Lightweight cloud detection network for high-resolution remote sensing images. *IEEE J Sel Top Appl Earth Obs Remote Sens*. 2022;15:4809-4823.
doi: 10.1109/JSTARS.2022.3181303
 29. Lu W, Chen SB, Shu QL, Tang J, Luo B. DecoupleNet: A lightweight backbone network with efficient feature decoupling for remote sensing visual tasks. *IEEE Trans Geosci Remote Sens*. 2024;62:4414613.
doi: 10.1109/TGRS.2024.3465496
 30. Xiao Y, Xu T, Yu X, Fang Y, Li J. A lightweight fusion strategy with enhanced interlayer feature correlation for small object detection. *IEEE Trans Geosci Remote Sens*. 2024;62:1-11.
doi: 10.1109/TGRS.2024.3457155