

Multi-Channel Gramian Angular Field Combined with Attention-Enhanced ResNet for Hydraulic Support Solenoid Valve Fault Diagnosis

Jinxin Wang ^{a,b}, Dong Zhen ^{c,d,*}, Yunfei Wang ^a, Wenxian Yang ^e, Zhonghui Li ^a, Xiaofei Liu ^a, Fengshou Gu ^{e,*}

^a School of Safety Engineering, China University of Mining and Technology, Xuzhou 221116, China.

^b Xuzhou Mining Business Group Co., Ltd., Xuzhou 221111, China.

^c School of Mechanical Engineering, Hebei University of Technology, Tianjin, 300401, China.

^d Tianjin Hegongda Advanced Equipment Research Institute Co., Ltd., Tianjin, 300401, China.

^e School of Computing and Engineering, University of Huddersfield, Huddersfield HD1 3DH, UK.

*Corresponding author: d.zhen@hebut.edu.cn

**Corresponding author: f.gu@hud.ac.uk

Received 27 July, 2026; Revised 14 September 2026; Accepted 28 September 2026; Published online 09 October 2026

Abstract: Solenoid valves are the most critical components in hydraulic supports. However, the solenoid valves suffer severely from multiple faults due to the harsh underground working conditions and intense dynamic impact loading, causing a large number of roof uncontrolled collapse accidents. Existing support diagnosis focuses only on the control system, while the solenoid valves, which deserve more attention, have been largely overlooked. This paper proposes a novel support solenoid valves fault identification approach by integrating multi-channel Gramian angular field with an attention-enhanced ResNet architecture. A scaled support test rig that replicates the operating principle of a ZY3000/09/20 shield support is firstly constructed, and the triaxial vibration responses under typical valve faults are then obtained via reproducing the faults on the test rig. The multi-channel Gramian angular field is introduced to investigate the temporal dependency of the vibration signals by mapping the one-dimensional time-series data into two-dimensional images. To effectively identify the fault type, the ResNet50 backbone is enhanced by the convolutional block attention module, which adaptively re-weights channel- and spatial-wise feature maps to amplify fault-relevant signatures. Ablation experiments

demonstrate that both the multi-channel GAF encoding and the CBAM attention module contribute meaningfully to the final diagnostic performance.

Keywords: Hydraulic Support; Solenoid Valve; Fault Diagnosis; Gramian Angular Field;

Residual Network

1. Introduction

Hydraulic supports serve as the primary roof control equipment in fully mechanized coal mining faces, which provides a stable working space by withstanding the extremely high loads from overlying strata. A hydraulic support is a complex coupled mechanical-hydraulic system, in which solenoid valves are the core actuating components that execute all support actions through precise control of high-pressure emulsion flow. However, the underground mining environment is characterized by intense dynamic impact, high static loads, and heavy dust as well as corrosive conditions. Such a harsh environment makes solenoid valves suffer severely from continuous wear, fatigue, and contamination throughout their service lives. In recent years, uncontrolled roof caving accidents caused by solenoid valve malfunction have been reported frequently, leading to numerous personnel casualties. Typically, the roof accident happened at Xinfeng coal mine, Henan province, China, in the year of 2022 resulted in 4 casualties. Therefore, developing reliable and timely fault diagnosis techniques for hydraulic support solenoid valves is of great importance for ensuring mining safety.

Currently, field investigations by the authors at various coal mines found that modern hydraulic support control system is normally mandatorily equipped with solenoid valves diagnosis sub-module due to the catastrophic consequences of valve malfunction. However, existing support valve diagnosis module mainly focuses on verifying whether a control command has been executed as intended by monitoring the feedback current. The abnormal state detection and fault location of valve are not directly accounted for. Although many works are devoted to the fault diagnosis of solenoid valves in other industrial fields [1-6], e.g., aviation, hydro-power, vehicle engineering, and general industrial production, the findings cannot be directly applied to the solenoid valves in hydraulic supports, due to their special

structural design. Specifically, hydraulic support solenoid valves adopt a two-stage pilot-operated structure, and are designed with large internal clearances to tolerate heavy fluid contamination and ensure fast and reliable logical actions [7,8]. These differences in structure result in distinct fault characteristics and diagnostic requirements, making direct transfer of existing diagnostic methods infeasible.

In the past few decades, mechanical fault diagnosis technology has been attempting to establish a high-fidelity dynamic model to describe the system behavior under normal and faulty conditions, and develop a diagnosis strategy accordingly [9-12]. This model-based diagnosis community facilitates the fundamental understanding of mechanical system dynamics, and has spawned numerous research achievements. With the ever-increasing functional demands of mechanical systems, however, these systems have become increasingly complex, posing an elaborate physical model is almost unavailable in practice. In contrast, data-driven methods learn the fault characteristics directly from sensor measurements without explicit physical modeling, and have gained increasing attention [13-15]. Among these, deep learning architectures have demonstrated remarkable performance in mechanical fault diagnosis by automatically extracting complete fault features from raw signals. However, a plenty of existing deep learning-based approaches operate directly on one-dimensional time-series data [16-20], which limits their ability to fully extract the fault features and fault identification.

To solve this problem, encoding time-series signals into two-dimensional images has been thought to be a promising procedure to fully exploit the advantages of deep learning model on feature extraction. Commonly encoding methodologies include the short-time Fourier transform (STFT) [21], wavelet transform (WT) [22], and Markov transition fields [23]. The Fourier-based signal processing method used in time-frequency representations such as STFT and

WT inherently miss the instantaneous phase relationships which is important to characterize a transient fault behavior, while Markov transition fields can only capture statistical state transition probabilities, losing the amplitude values and the absolute temporal structure of the original signal. As revealed in our previous work, the vibration signals of hydraulic support solenoid valves are characterized by transient impulsive components, and fault-induced signal is encoded in instantaneous phase relationships and localized amplitude patterns. This property makes transfer methods above limited to analyze the solenoid valves signals.

Gramian angular field (GAF) can preserve both temporal dependencies and amplitude impulse when transforms one-dimensional signal into images, which has attracted particular attention from researcher. Recent studies have successfully applied GAF-based encoding to the fault diagnosis of bearing [24], gearbox [25], and motor [26], demonstrating its advantages on feature extraction. Nevertheless, the application of GAF encoding to solenoid valve fault diagnosis remains largely unexplored. Meanwhile, existing works normally construct GAF images from a single vibration channel, which fails to uncover the complementary information contained in multi-axis sensor signals.

As for fault identification, residual networks (ResNet) have become a promoting deep classification model by introducing skip connections that mitigate the vanishing gradient problem and enable effective training of very deep networks [27]. Some researchers also attempt to introduce various attention mechanism into ResNet architectures to improve the feature selection ability of the model, the convolutional block attention module (CBAM) being one of them [28-30]. These findings suggest that combining GAF-based signal imaging with attention-enhanced ResNet offers a promising way for achieving high-precision solenoid valve fault diagnosis.

Motivated by the above considerations, this paper proposes a novel fault diagnosis method for hydraulic support solenoid valves by integrating multi-channel GAF image encoding with a CBAM-enhanced ResNet architecture. The main contributions of this paper are summarized as follows: (i) for the first time, this paper investigates the fault diagnosis of hydraulic support solenoid valves, and proposes a tailored diagnostic framework that considers the transient

impulsive characteristics of support solenoid valve vibrations; (ii) a multi-channel encoding strategy is developed to fuse complementary directional fault features by transforming multi-axis vibration signals into a unified two-dimensional representation for deep learning-based classification; and (iii) an attention-enhanced architecture is constructed to amplify fault-relevant signatures by integrating channel and spatial attention mechanisms into the residual network backbone.

The remainder of this paper is organized as follows. Section 2 describes the multi-channel GAF encoding methodology for solenoid valve fault diagnosis. Section 3 presents the CBAM-enhanced ResNet50 architecture and the model training strategy. Section 4 develops a scaled support test rig, and validates the proposed methodology using fault reproduction experiments. Finally, Section 5 summarizes the paper.

2. Multi-Channel GAF Imaging of Vibration Signals

2.1 Fundamentals of Gramian Angular Field

The gramian angular field is a transformation operation which have the ability to map one-dimensional time-series data into two-dimensional pictures while keep the temporal dependency of the raw signals unchanged. Let $S = \{s_1, s_2, \dots, s_n\}$ be the detected data, the normalization is firstly applied to scale the data into a standard range.

$$\bar{s}_i = \frac{(s_i - \max(S)) + (s_i - \min(S))}{\max(S) - \min(S)} \quad (1)$$

The data vector is then projected onto the polar coordinates for the image generation afterwards, where the angular coordinate is used to represent the signal amplitude and the radial coordinate denotes the temporal order. The polar angle ϕ_i and radius r_i can be computed as Eqs. (2) and (3).

$$\phi_i = \arccos(\bar{s}_i), \quad 0 \leq \phi_i \leq \pi \quad (2)$$

$$r_i = \frac{t_i}{N}, \quad i = 1, 2, \dots, N \quad (3)$$

A specialized inner product is defined to capture the temporal correlations between each pair of measurements. The gramian angular summation field (GASF) is defined using the cosine of the sum of two angles.

$$\text{GASF} = \begin{bmatrix} \cos(\phi_1 + \phi_1) & \cos(\phi_1 + \phi_2) & \cdots & \cos(\phi_1 + \phi_n) \\ \cos(\phi_2 + \phi_1) & \cos(\phi_2 + \phi_2) & \cdots & \cos(\phi_2 + \phi_n) \\ \vdots & \vdots & \ddots & \vdots \\ \cos(\phi_n + \phi_1) & \cos(\phi_n + \phi_2) & \cdots & \cos(\phi_n + \phi_n) \end{bmatrix} \quad (4)$$

This can be expressed more computable form as

$$\text{GASF} = \tilde{X}^T \cdot \tilde{X} - \sqrt{I - \tilde{X}^2}^T \cdot \sqrt{I - \tilde{X}^2} \quad (6)$$

where $\tilde{X}$ is the normalized time series and I is the unit vector.

The GASF primarily captures the global correlation structure and phase relationships among data points, making it particularly effective at representing the overall trends and periodicities inherent in the signal. Both horizontal and vertical axes of a GASF image represent time, and the element at position (i, j) of the GASF matrix represents the correlation between the i -th and j -th time points. Therefore, its structure reflects the reciprocal nature of temporal correlations, which is beneficial for deep learning model-based feature extraction as it provides consistent representations regardless of the direction of analysis.

The symmetric structure of GASF matrix reflects the reciprocal nature of temporal correlations, which is beneficial for CNN-based feature extraction as it provides consistent representations regardless of the direction of analysis. Specifically, the diagonal elements of GASF represent the autocorrelation of each measurement, while the other elements temporal correlations of the corresponding measurements. This nature makes the GASF matrix having the ability to encoding the dynamic evolution pattern of signals, which is suitable for the convolution operation in a deep learning model to characterize a fault type.

2.2 Multi-Channel GAF Encoding Strategy

In solenoid valve fault diagnosis, the vibration signals are typically acquired in three orthogonal directions i.e. X-, Y-, and Z- axes using a triaxial accelerometer mounted on the valve body surface. Each axis captures distinct dynamic characteristics of the valve during operation.

To fully exploit the complementary information embedded in the triaxial vibration signals, a three-channel GAF encoding strategy is proposed that transforms each axial signal into a GASF image and assigns it to a distinct RGB channel. Specifically, the GASF-transformed X-axis signal is assigned to the red channel, the GASF-transformed Y-axis signal to the green channel, and the GASF-transformed Z-axis signal to the blue channel.

Let the triaxial vibration signals be denoted as $S_X = \{s_{X,1}, s_{X,2}, \dots, s_{X,N}\}$, $S_Y = \{s_{Y,1}, s_{Y,2}, \dots, s_{Y,N}\}$, and $S_Z = \{s_{Z,1}, s_{Z,2}, \dots, s_{Z,N}\}$, where N is the number of sampling points in each channel. The GASF matrices for the X-, Y-, and Z- axes are then computed according to Eq. (4), denoted as GASF_X , GASF_Y , GASF_Z , respectively, each of dimension $N \times N$. The three-channel RGB image is then constructed by stacking the three matrices.

$\text{GAF}_{\text{RGB}} = [\text{GASF}_X, \text{GASF}_Y, \text{GASF}_Z]$ (7) where the three channels correspond to the red, green, and blue color planes, respectively. All generated images are resized to 224×224 pixels to ensure compatibility with the ResNet50 architecture described in Section 3. The complete encoding procedure is illustrated in Figure 1.

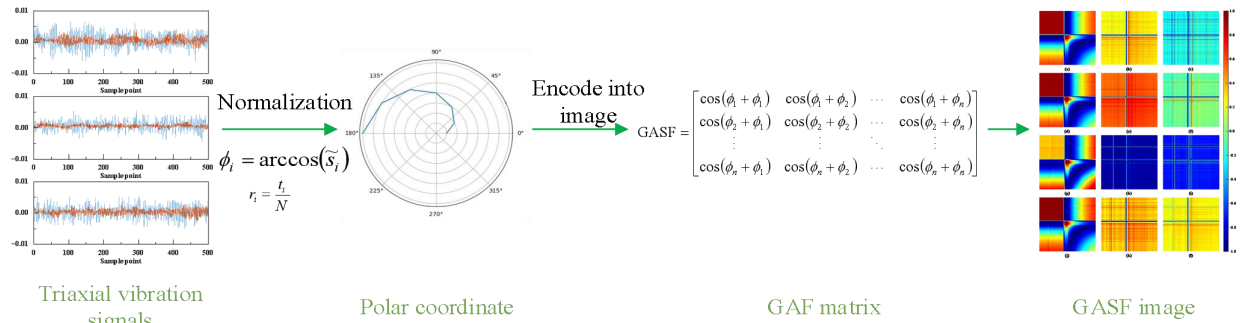


Fig. 1. Schematic diagram of the multi-channel GAF encoding strategy.

3. Fault Diagnosis Based on Attention-Enhanced ResNet

3.1 ResNet50 with Transfer Learning

ResNet50 is adopted as the backbone network for GAF image classification. As shown in Figure 2, the core innovation of ResNet50 lies in the bottleneck residual block, which mitigates the degradation problem in deep networks through skip connections. For an input feature map $\mathbf{x}$, a traditional network directly learns a desired underlying mapping $H(\mathbf{x})$, whereas the residual block instead learns a residual mapping $F(\mathbf{x}) = H(\mathbf{x}) - \mathbf{x}$. The output of the block is therefore

$$\mathbf{y} = F(\mathbf{x}) + \mathbf{x} \quad (8)$$

where $F(\mathbf{x})$ represents the stacked convolutional operations. This skip connection allows the gradient to flow directly through the identity mapping during backpropagation

$$\frac{\partial L}{\partial \mathbf{x}} = \frac{\partial L}{\partial \mathbf{y}} \cdot \frac{\partial L}{\partial \mathbf{x}} + \frac{\partial L}{\partial \mathbf{y}} \quad (9)$$

This formulation ensures that gradients can propagate through the network without attenuation even when the residual mapping F approaches zero.

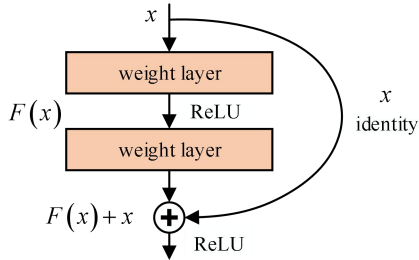


Fig. 2. Skip connection in ResNet50.

The structure of ResNet50 is shown in Figure 3. The input structure is a 7×7 convolutional layer with 64 filters and a stride of 2, followed by a 3×3 max pooling layer with a stride of 2, which rapidly reduces the spatial dimensions of the input image. The subsequent feature extraction is organized into four main stages, each composed of a stack of 3 bottleneck residual blocks. Stage 1 contains 3

bottleneck blocks, operating on 56×56 feature maps with 256 output channels. Stage 2 contains 4 bottleneck blocks, producing 28×28 feature maps with 512 channels. Stage 3 consists of 6 bottleneck blocks, generating 14×14 feature maps with 1024 channels. Stage 4 is built with 3 bottleneck blocks, generating 7×7 feature maps with 2048 channels. Within each stage, the first bottleneck block performs spatial downsampling using a stride-2 convolution in its 3×3 layer, while the remaining blocks maintain the feature map size. After the final convolutional stage, a global average pooling layer collapses the spatial dimensions into a feature vector, and a fully connected layer with softmax activation produces the final classification output.

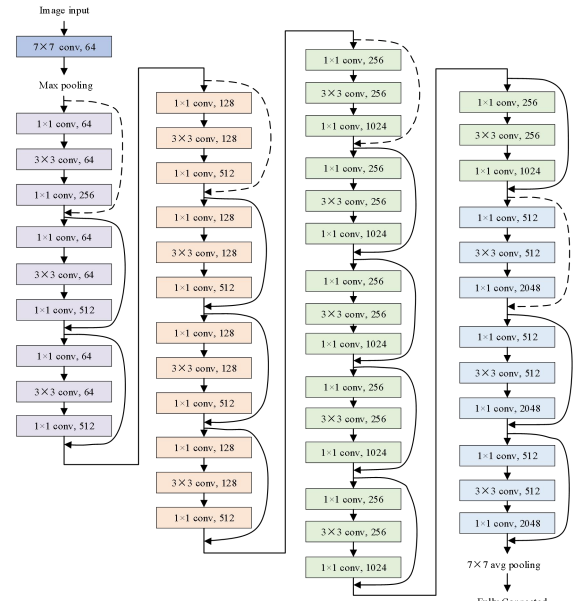


Fig. 3. Architecture of the ResNet50 network.

3.2 CBAM Attention Mechanism

To enhance feature discriminability, the convolutional block attention module is embedded into each residual block. CBAM sequentially computes channel and spatial attention maps, adaptively recalibrating features along both dimensions.

The principle of CBAM is shown as Figure 4. Given an intermediate feature map $\mathbf{F} \in \mathbb{R}^{C \times H \times W}$, the channel attention map $\mathbf{M}_c \in \mathbb{R}^{C \times 1 \times 1}$ is computed by aggregating spatial information through both average and max pooling.

$$\mathbf{M}_c(\mathbf{F}) = \sigma(\text{MLP}(\text{AvgPool}(\mathbf{F})) + \text{MLP}(\text{MaxPool}(\mathbf{F}))) \quad (10)$$

where MLP is a shared two-layer perceptron, and σ denotes the sigmoid function. The channel-refined feature is

$$\mathbf{F}' = \mathbf{M}_c(\mathbf{F}) \otimes \mathbf{F} \quad (11)$$

The spatial attention map $\mathbf{M}_s \in \mathbb{R}^{1 \times H \times W}$ is then applied to $\mathbf{F}'$

$$\mathbf{M}_s(\mathbf{F}') = \sigma(f^{7 \times 7}([\text{AvgPool}(\mathbf{F}'), \text{MaxPool}(\mathbf{F}')])) \quad (12)$$

where $f^{7 \times 7}$ denotes a convolution with kernel size 7×7 , and $[\cdot]$ indicates channel-wise concatenation. The final output is

$$\mathbf{F}'' = \mathbf{M}_s(\mathbf{F}') \otimes \mathbf{F}' \quad (13)$$

CBAM is inserted after the convolutional layers in each residual block, prior to the skip connection, enabling progressive feature refinement throughout the network with minimal computational overhead.

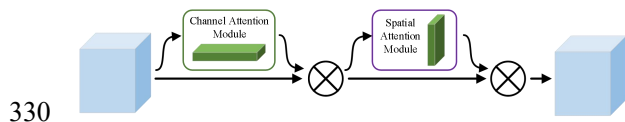


Fig. 4. Architecture of the convolutional block attention module.

3.3 Model Training Strategy

The triaxial vibration signals are firstly transformed into GAF pictures according technical scheme presented in Section 2. The GAF pictures are then assigned to RGB channels for fault identification, and each image is resized to 224×224 pixels to input into the CBAM-enhanced ResNet50 network. The training datasets, validation datasets, and test datasets are randomly selected from the original datasets, with a ratio of 70:15:15.

The training error is evaluated using multi-class cross-entropy loss function, i.e.,

$$L = -\frac{1}{N} \sum_{i=1}^N \sum_{j=1}^K y_{ij} \log(\hat{y}_{ij}) \quad (14)$$

where N represents the batch size, K denotes the number of categories, including normal and fault types, y_{ij} is the labels of categories, and $\hat{y}_{ij}$ is the output probability for the j -th class of the i -th sample.

In order to improve the robustness of the model, data augmentation is applied to enrich training

images. The image augmentation operation includes random flips, random rotations, brightness adjustments and random noise addition. Each picture has 20% probability to be independently applied using the augmentation operation above. The model is trained and deployed using the PyTorch framework. The batch size is set as 32. Both training and validation accuracy and loss curves is monitored to evaluate the training process. The final model evaluation is conducted on the test datasets to report the unbiased classification performance.

4. Experiment Results

4.1 Hydraulic Support Test rig

4.1.1 Overall architecture

Figure 5 illustrates the self-developed scaled-down support test rig that replicates the operating principle of a ZY3000/09/20 shield support, and the hydraulic schematic is shown as Figure 6. The test rig consists of a drive unit, a loading unit, and a real-time measurement and control system. The drive unit integrates a variable-frequency pump, a servo pump, an unloading valve, long hydraulic pipelines, a solenoid valve assembly, and a set of cylinders that replicate the support working actions. The solenoid valve assembly under investigation is ZMJ FHDC200A(8)A52X model. The cylinders in the test rig includes one leg cylinder and three advancing-ram cylinders, where cylinders of different sizes are employed to reproduce the various motions of the hydraulic support. The loading unit consists of a proportional relief valve, a loading cylinder, and a hydraulic bridge circuit. The applied load, which emulates underground strata pressure, is regulated by the proportional relief valve, and the hydraulic bridge enables bidirectional loading. A real-time control and data acquisition platform is established using Simulink Real-Time, which performs synchronous acquisition of system pressures and provides closed-loop control of the servo pump output throughout the test.

4.1.2 Design of the key operating parameters

The dimensions of the simulation cylinders are determined based on similarity principles according to the parameters of the actual mining support, i.e. ZY3000/09/20. Each cylinder is butt-connected to a dedicated loading cylinder in an

opposed configuration to simulate the underground
load. The leg cylinder has a bore diameter of 180
mm, a rod diameter of 160 mm, a stroke of 400
mm, and a setting load of 420 kN. Each advancing-
ram cylinder has a bore diameter of 140 mm, a rod
diameter of 80 mm, and a stroke of 800 mm,
providing a conveyor-pushing force of 35.6 kN
and a support-pulling force of 61.2 kN. The
loading cylinders have a bore diameter of 140 mm,
a rod diameter of 80 mm, and a stroke of 1000 mm.

The pump station pressure is designed to satisfy
the working pressure required by the leg setting
load and the rams, i.e.

$$p_b \geq k_1 p_m \quad (15)$$

where p_m is the setting working pressure derived
from the leg or ram, and k_1 is the pressure loss
coefficient along the pipeline from the pump
station to the support, typically taken as 1.1-1.2.
The required pump station pressure for the test rig
is estimated to be 18 MPa.

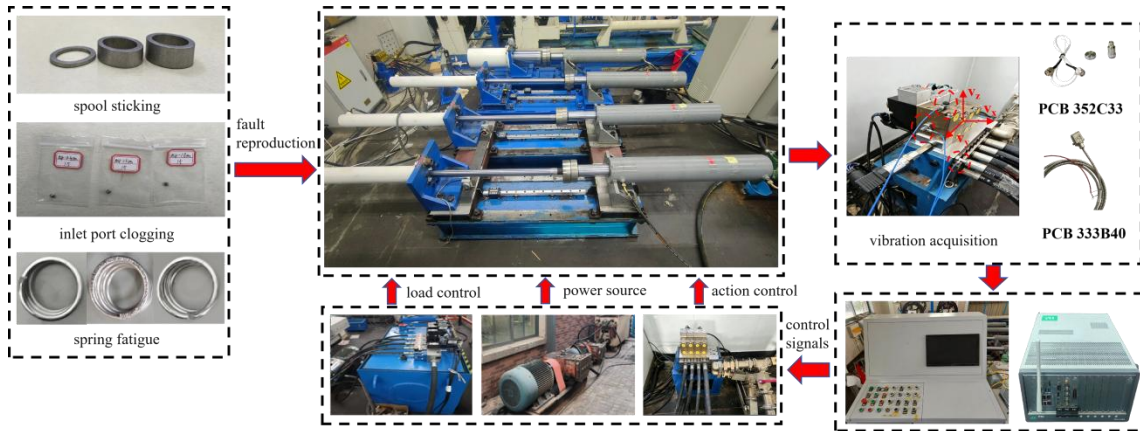


Fig. 5. A principle-equivalent and scaled-down test rig of the ZY3000/09/20 shielded hydraulic support.

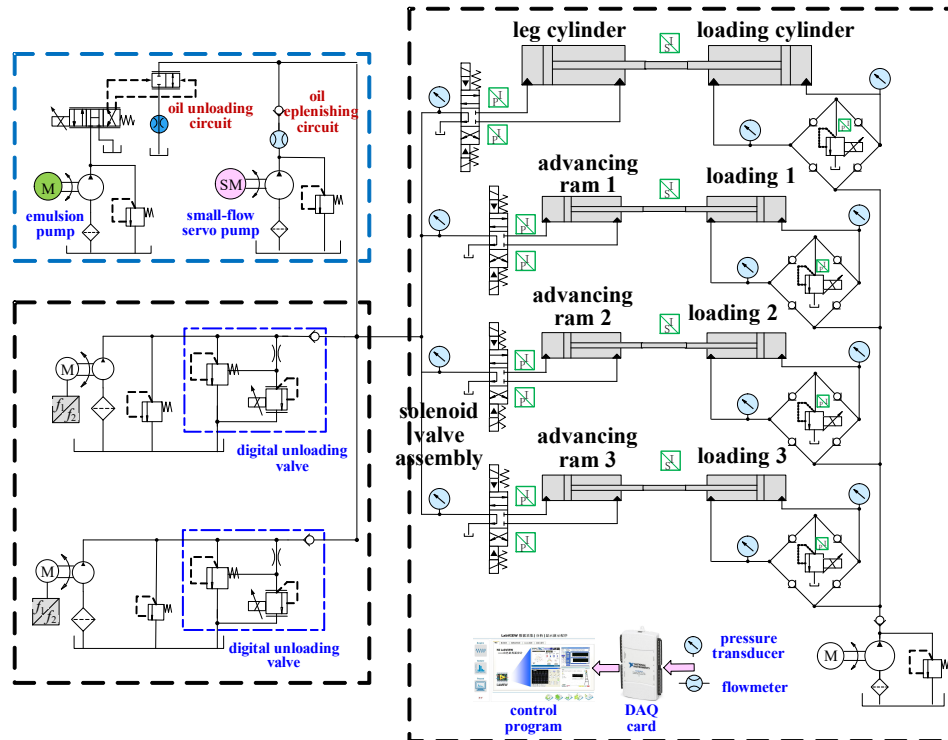


Fig. 6. Hydraulic schematic of the test rig.

The pump station flow rate is generally estimated by assuming that all legs and rams of one support operate at the same time. The flow rate Q_b of the emulsion pump station should satisfy

$$Q_b \geq k_2 \left(\sum Q_i \right) \frac{v_q}{A} \times 10^{-3} \quad (16)$$

where $\sum Q_i$ is the total volume of emulsion required for all legs and rams of one support to complete all actions, v_q is the haulage speed of the shearer, A is the support center distance, and k_2 is the leakage loss coefficient from the pump station to the support pipeline, typically taken as 1.1-1.3. Based on Eq. (16), the estimated flow rate for the test rig is 205 L/min.

For the raising action, the load force mainly originates from the dead weight M of the leg and the friction force f_m between the piston rod and the cylinder bore of the leg emulsion cylinder. Denoting the effective area of the leg loading cylinder as A_L , the load pressure P_L is

$$P_L = \frac{2Mg + f_m}{A_L} \quad (17)$$

where g is the gravitational acceleration. The mass of each leg of the ZY3000/09/20 hydraulic support is 2450 kg, so the load pressure is 3.43 MPa.

For the lowering action, the load force originates primarily from the friction force f_m between the piston rod and the cylinder bore of the leg emulsion cylinder, and the load pressure P_j is

$$P_j = \frac{f_m}{A_L} \quad (18)$$

therefore the pressure is calculated as 0.5 MPa.

According to the monitoring data on-site, the pressure fluctuation amplitudes for different support actions are 2.36 MPa for lowering, 2.13 MPa for pulling, 5.29 MPa for raising, and 9.63 MPa for sequential pushing of the armored face conveyor. The corresponding pressure fluctuation amplitudes measured on the test rig are 2.26 MPa for lowering, 2.15 MPa for pulling, and 5.1 MPa for raising, which are in good agreement with the underground data and demonstrate that the test rig can faithfully reproduce the pressure

characteristics of the actual support supply system during different operations.

4.1.3 Fault reproduction and data preprocessing

According to the on-site maintenance reports, three solenoid valve faults commonly occurred, i.e., spool sticking, inlet port clogging, spring fatigue, are reproduced in the test rig with 3 graded severity levels, respectively. Specifically, inlet port clogging is reproduced by mounting replaceable throttling orifices with three different diameters (1.8 mm, 0.8 mm, and 0.4 mm) at the port. Spool sticking is simulated by installing aluminum spacer rings with three different widths (0.5 mm, 2.0 mm, and 2.5 mm) at the spring seat. As for spring fatigue, the gradual loss of restoring force is reproduced by progressively reducing the effective turns of the spring coil by 0.25, 1.00, and 1.25 turns, respectively. The faulty components are presented in Figure 5, and the testings are carried out with working loading 15 MPa, which is a scaled-down loads of ZY3000/09/20. The vibration signals are acquired using a NI PXIe-1092 chassis equipped with NI PXIe-4464 module. The sampling frequency is set as 5 kHz, and each experiment is independently carried out for 50 times to obtain an accurate description of the working process.

The vibration signals of the solenoid valve mainly consist of switching-induced impact transients. A denoising method based on compressed sensing is applied to remove the background noise while preserving the switching-induced transients. An energy contribution criterion is used to truncate the coefficients adaptively. Figure 7 compares the signals before and after denoising. It can be seen that the high-frequency random fluctuations are effectively suppressed, and the periodic impulsive characteristics are well retained. This result helps the subsequent fault analysis. However, for incipient faults, the useful signal features are still heavily contaminated by residual noise after denoising, which makes early fault diagnosis very challenging. To solve this problem, a dedicated diagnostic method is proposed in the next section.

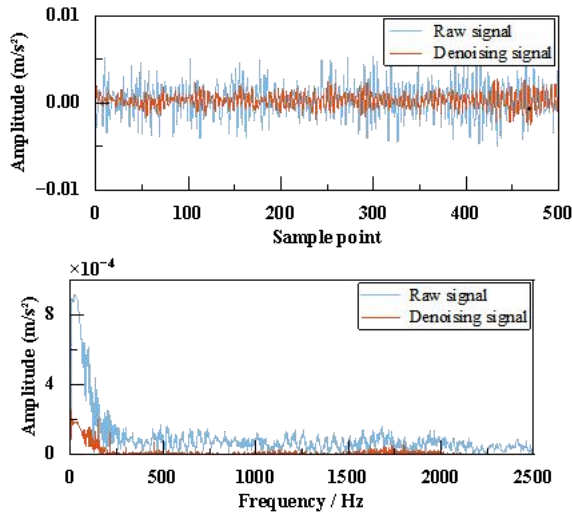


Fig. 7. Time-domain and frequency-domain comparison of raw and denoised signals for spring fatigue.

4.2 Classification Performance and Visual Interpretation

Figure 8 presents the GASF images of triaxial vibration signals under different states, where the first row denotes the health condition and last 3 row represent the three fault conditions, i.e., spool sticking, inlet port clogging, spring fatigue, respectively. It should be noted that the normalization is performed independently for each axis, since the three-axis vibration signals have different amplitude values. Distinct horizontal and vertical intersecting stripes can be observed in GASF image, which indicates the periodic impulsive component in solenoid valve vibration signals and represent the temporal correlation characteristics of switching-induced transients in the two-dimensional image domain. Notably, the density, regularity, and intensity distribution of these stripes differ markedly between the healthy condition and the faulty states, as well as among the three fault types. These differences demonstrate that the GASF images can effectively capture the fault-induced periodic and impulsive components of the vibration signals. Furthermore, observable local bright areas near the diagonal line are presented in the X-axis GASF images, corresponding to the instantaneous impact peaks of the X-axis vibration signal. According to the operating sequence of the solenoid valve, these impact signals are caused by the switching events of each support action.

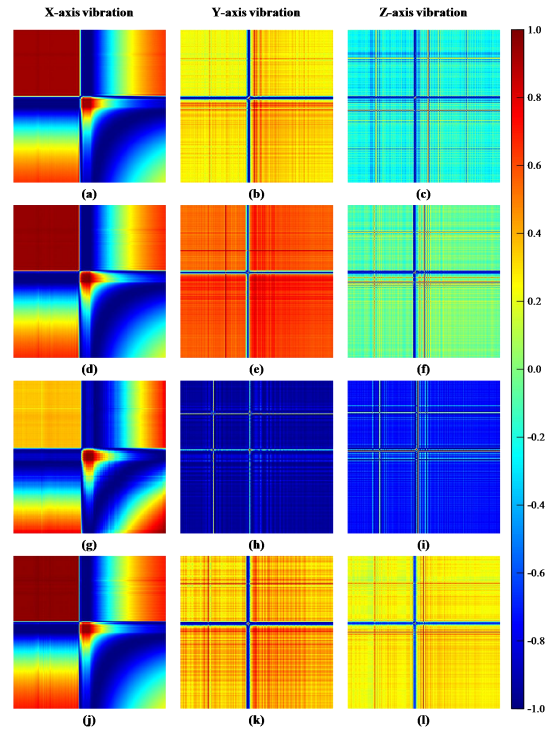


Fig. 8. Multi-channel GASF sample images for each condition.

Figure 9 presents the training process of the CBAM-enhanced ResNet50 model fed by multi-channel GASF image datasets. The diagnostic performance is evaluated via the training and validation loss and accuracy curves. It can be seen that the training process is smooth and the model converges within the first 20 epochs. The smooth training process also indicates the GASF images effectively capture the discriminative characteristics of different fault types.

Figure 10 shows the confusion matrix and t-SNE visualization of the model, where label 0 represents the health condition, label 1 denotes the spool sticking, label 2 depicts inlet port clogging, and label 3 is spring fatigue. The proposed method achieves an overall accuracy of 98.82% across all conditions, and the precision, recall, and macro F1 scores all exceeding 98%. The false alarm rate and miss alarm rate are only 1.89% and 0.94%, respectively, indicating satisfactory performance for multiple fault identification. The t-SNE visualization directly displays the fault classification results. The clear inter-class separation also demonstrates the effectiveness of

the proposed method for solenoid valve fault identification.

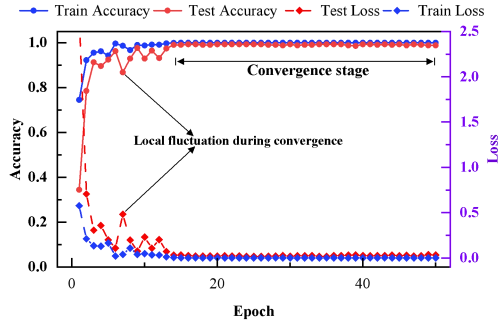


Fig. 9. Training and validation loss and accuracy curves

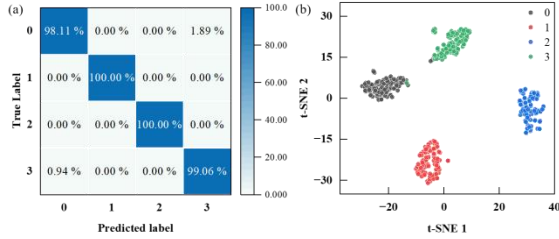


Fig. 10. Confusion matrix and t-SNE on the test set.

As for the model interpretability, i.e. which regions of the GASF images contribute most for fault identification, Grad-CAM is used to visualize the attention heatmaps, shown as Figure 11. The left column displays the original GASF images of health condition as well as three fault types, i.e. spool sticking, inlet port clogging, spring fatigue, respectively, while the right column shows corresponding Grad-CAM images. The image regions highlighted by Grad-CAM represent the image features contribute most for fault characterization. It may be noted that main attention has been given to the intersections of horizontal and vertical stripes, which represent the intrinsic vibration response characteristics of different faults according to the analysis above. These observations confirm that the CBAM module effectively guides the ResNet model to focus on fault-sensitive features in the GASF images, and thereby improves the diagnostic accuracy.

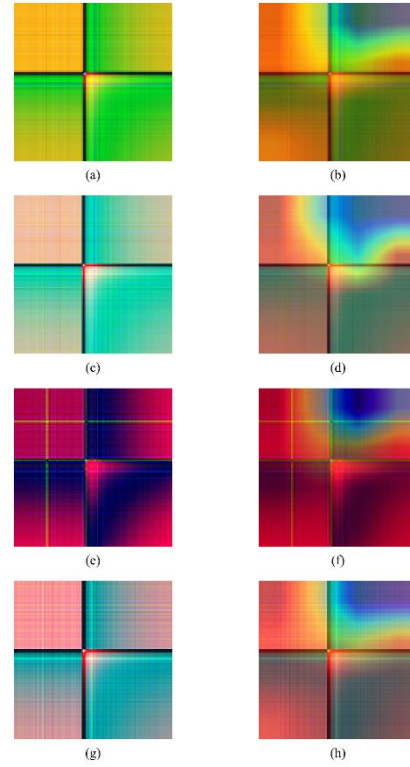


Fig. 11. Attention heatmap visualization on representative GAF images.

4.3 Ablation Study on GAF Encoding and Attention Module

A set of ablation experiments is carried out to investigate the individual contributions of each model components. All ablation experiments are implemented on all fault types to investigate the overall performance. The first variant is removing the CBAM module from the model, and the second one only uses two vibration signals, and the last variant uses the X-axis GASF image as a single-channel input. The classification results of original and ablation models are summarized in Figure 12. The three variant all shows a noticeable performance decline compared with original model, where the single-channel model has the lowest performance with accuracy dropping from 98.82% to 83.02%, and macro F1-score decreases from 98.83% to 82.26%. Meanwhile, the dual-channel model also shows a performance decline. These comparisons indicate that multi-channel input, i.e. triaxial vibration, has a complete description of faults behavior and thereby obtain a satisfactory diagnosis results.

628 Removing the CBAM module from the original
629 model, i.e. the first variant, also lead to a slight but
630 noticeable performance decrease. This result
631 demonstrates that even when the three-channel
632 encoding already provides rich feature input,
633 CBAM module can further improve the diagnosis
634 accuracy via guiding ResNet model to focus on
635 fault discriminative features using the channel and
636 spatial attention mechanisms.

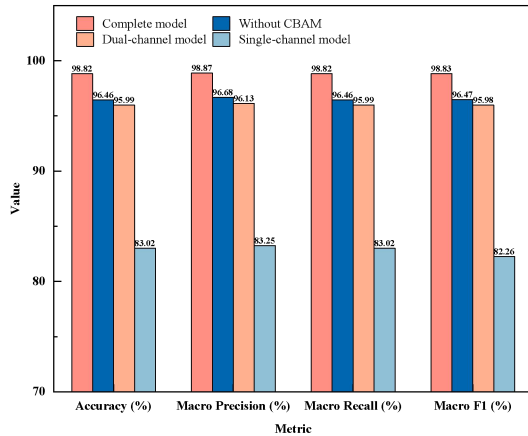


Fig. 12. Ablation results under different GAF channel configurations and attention settings.

5. Conclusion

(1) The three typical solenoid valve faults, i.e. spool sticking, inlet port clogging, spring fatigue, mainly alter the temporal correlation of the impulsive vibration signals, which is reflected in the GASF images as distinct variations in the cross-stripe patterns. Grad-CAM heatmaps demonstrate that these textures are the primary image feature for fault identification.

(2) Experiments and ablation test demonstrate that the multi-channel GASF images, i.e. triaxial vibration, has a complete description of faults behavior and thereby leads to a satisfactory diagnosis results, i.e. achieving an accuracy of 98.82% in our test, significantly outperforming dual-channel and single-channel models.

(3) CBAM module can further improve the diagnosis accuracy via guiding ResNet model to focus on fault discriminative features using the channel and spatial attention mechanisms. Grad-CAM heatmaps shows that the highlighted regions match well with the discriminative fault image features, which demonstrates the effectiveness of CBAM module.

Acknowledgment

This work was supported by the National Natural Science Foundation of China (Grant No.: 52304271, 52404268), Natural Science Foundation of Jiangsu Province, China (Grant No.: BK20221117) and the Basic Research Project of Xuzhou City, China (Grant No.: KC22001).

References

- [1] D. Ma et al., “Fault diagnosis of multi-step electromagnetic hydraulic valve group based on localized current signal CS-SVM,” *Meas.*, vol. 245, pp. 116632, 2025.
- [2] W. Li et al., “Electromagnetic valve fault diagnosis based on multi-source signal feature fusion,” *Meas. Sci. Technol.*, vol. 36, no. 7, pp. 075010, 2025.
- [3] V. Vantilborgh et al., “Data-driven virtual sensing for probabilistic condition monitoring of solenoid valves,” *IEEE Trans. Autom. Sci. Eng.*, vol. 21, no. 2, pp. 1297–1311, 2024.
- [4] J. Liniger et al., “Early detection of coil failure in solenoid valves,” *IEEE/ASME Trans. Mechatron.*, vol. 25, no. 2, pp. 683–693, 2020.
- [5] M. N. Utah et al., “Fault state detection and remaining useful life prediction in AC powered solenoid operated valves based on traditional machine learning and deep neural networks,” *Nucl. Eng. Technol.*, vol. 52, no. 9, pp. 1998–2008, 2020.
- [6] G. Mazaev et al., “Bayesian convolutional neural networks for remaining useful life prognostics of solenoid valves with uncertainty estimations,” *IEEE Trans. Ind. Inform.*, vol. 17, no. 12, pp. 8418–8428, 2021.
- [7] Y. Liao et al., “Optimization of an intrinsically safe solenoid valve and the static and dynamic characteristics,” *Int. J. Appl. Electromagn. Mech.*, vol. 61, pp. 111–112, 2019.
- [8] H. Wang et al., “Design and dynamic characteristics measurement of a water hydraulic proportional flow valve with a discrete pilot stage,” *IEEE Trans. Instrum. Meas.*, vol. 74, pp. 7507011, 2025.
- [9] A. Srivastava et al., “Development of a digital model of a gear rotor system for fault diagnosis using the finite element method and machine learning,” *J. Dyn. Monit. Diagn.*, vol. 4, pp. 121–136, 2025.

- [10] Y. Han et al., “The use of model-based voltage and current analysis for torque oscillation detection and improved condition monitoring of centrifugal pumps,” *Mech. Syst. Signal Process.*, vol. 222, pp. 111781, 2025.
- [11] Z. Teng et al., “Dynamic relative advantage-driven multi-fault synergistic diagnosis method for motors under imbalanced missing data rates,” *J. Dyn. Monit. Diagn.*, vol. 4, pp. 111–120, 2025.
- [12] Q. Chao et al., “Physics informed neural networks for detecting the wear of friction pairs in axial piston pumps,” *Reliab. Eng. Syst. Saf.*, vol. 261, pp. 111144, 2025.
- [13] P. Chopra et al., “Open-set fault classification for rotatory machine by DNN’s neuron activation similarity score,” *IEEE Sens. J.*, vol. 26, no. 10, pp. 14831–14840, 2026.
- [14] H. Bilal et al., “Online fault diagnosis of industrial robot using IoRT and hybrid deep learning techniques: An experimental approach,” *IEEE Internet Things J.*, vol. 11, no. 19, pp. 31422–31437, 2024.
- [15] X. Chen et al., “Large models for machine monitoring and fault diagnostics: Opportunities, challenges, and future direction,” *J. Dyn. Monit. Diagn.*, vol. 4, pp. 76–90, 2025.
- [16] H. Ruan et al., “An enhanced non-local weakly supervised fault diagnosis method for rotating machinery,” *Meas.*, vol. 189, pp. 110433, 2022.
- [17] Z. Wu et al., “A deep reinforcement transfer convolutional neural network for rolling bearing fault diagnosis,” *ISA Trans.*, vol. 129, pp. 505–524, 2022.
- [18] J. Tian et al., “A test-time adaptation method using evidential deep learning for online machinery fault diagnosis,” *Knowl.-Based Syst.*, vol. 331, pp. 114831, 2025.
- [19] C. Yang et al., “A novel deep parallel time-series relation network for fault diagnosis,” *IEEE Trans. Instrum. Meas.*, vol. 72, pp. 3507413, 2023.
- [20] T. Lin et al., “A novel multi-sensor information fusion method for fault diagnosis of rotating machinery with missing signals,” *Adv. Eng. Inform.*, vol. 68, pp. 103595, 2025.
- [21] A. K. Sharma et al., “A novel vision transformer with selective residual in multihead self-attention for pattern recognition,” *Pattern Recognit.*, vol. 172, pp. 112497, 2026.
- [22] Y. Yu et al., “A novel digital twin-enabled three-stage feature imputation framework for non-contact intelligent fault diagnosis,” *Adv. Eng. Inform.*, vol. 66, pp. 103434, 2025.
- [23] A. Teta et al., “Robust and fast fault diagnosis framework for grid-connected photovoltaic systems via adaptive channel-wise representations and transfer learning,” *Meas.*, vol. 258, pp. 119290, 2026.
- [24] J. Wang et al., “A novel graph representation framework for bearing fault diagnosis via wavelet packet decomposition and Gramian angular field weighting,” *Nondestruct. Test. Eval.*, to be published, 2025.
- [25] Q. Yang et al., “Self-attention parallel fusion network for wind turbine gearboxes fault diagnosis,” *IEEE Sens. J.*, vol. 23, no. 19, pp. 23210, 2023.
- [26] A. T. Hindi et al., “Modified Grammian angular field for spatially informed induction motor fault diagnosis using feature fusion for fault tolerant control applications,” *Nondestruct. Test. Eval.*, vol. 41, no. 8, pp. 5066–5103, 2026.
- [27] V. S. B. Rama et al., “A fixed kernel and cloud-integrated deep learning model for wind turbine fault diagnosis with imbalanced SCADA data,” *Int. J. Control Autom. Syst.*, to be published, 2026.
- [28] X. Zhang et al., “Research on rolling bearing fault diagnosis based on parallel depthwise separable ResNet neural network with attention mechanism,” *Expert Syst. Appl.*, vol. 286, pp. 128105, 2025.
- [29] X. Zhao et al., “Compound fault diagnosis of diesel engines by combining CDCGAN and multistage transfer learning,” *J. Ind. Inf. Integr.*, vol. 50, pp. 101073, 2026.
- [30] X. Yan et al., “Machinery multi-sensor fault diagnosis based on adaptive multivariate feature mode decomposition and multi-attention fusion residual convolutional neural network,” *Mech. Syst. Signal Process.*, vol. 202, pp. 110664, 2023.