GEOPHYSICS, VOL. 91, NO. 1 (JANUARY–FEBRUARY 2026); P. WA123–WA135, 16 FIGS., 4 TABLES.
10.1190/GEO-2024-0534

Rapid retrieval and classification of passive-source body wave events using a convolutional self-attention encoder: Application to gas storage monitoring

Yinghe Wu[1,2], Shulin Pan[1,3], Dawei Liu[2,4,] Kai Chen[1,3], Yaojie Chen[1,3], Ziyu Qin[5,6], Shengbo Yi[7], and Zeyang Liu[7]

ABSTRACT

Passive-source seismic interferometry (SI) demonstrates significant potential in geophysical monitoring owing to its low cost and nondestructive attributes. However, uneven noise source distributions, pervasive surface waves, and other sources of noise often obscure body wave reflections, leading to very low signal-to-noise ratios in reconstructed signals. Additionally, long-term data acquisition generates large, high-noise data sets, increasing processing costs. Thus, rapidly retrieving body wave-dominated segments is essential for preprocessing. In this study, a convolutional self-attention encoder (CSE) was introduced to address these challenges. This method adopted multimodal inputs consisting of noise segments and their corresponding frequency attributes, enhancing overall performance. An attention mechanism was integrated to construct a binary classifier based on the CSE, enhancing contextual understanding. These steps helped the CSE to better learn features from the data sets, thereby improving generalization ability and prediction accuracy. Accordingly, the network was updated through incremental learning on small data sets, enabling continuous prediction and gradual adaptation to new data. This enabled us to rapidly generate large, high-quality training data sets and to train models with high performance. To further improve prediction data quality, the intermediate features extracted from the deep convolutional layers corresponding to the retrieved body wave events were extracted, and these features were clustered to perform quality classification of the retrieved data. Synthetic simulation data and field data from Northwest China validated that the proposed method retrieves and classifies body wave reflections with high accuracy and efficiency, providing robust data for subsequent monitoring and enhancing imaging results. The proposed method provides efficient and reliable data preprocessing for the application of passive-source SI in energy transition scenarios, such as subsurface fluid migration monitoring and carbon dioxide geological sequestration, and facilitates the development of low-carbon exploration and green transformation in the energy industry.

INTRODUCTION

In the context of energy transition and climate change, technologies such as the low-carbon development of unconventional oil and gas and geological carbon dioxide (CO_2) sequestration have become critical to achieving carbon neutrality goals. Passive-source seismic interferometry (SI) (Wapenaar et al., 2008), characterized by its environmentally friendly nature and lack of reliance on artificial seismic sources, demonstrates unique advantages in these scenarios by enabling low-cost, sustainable subsurface dynamic monitoring. By applying interferometry and processing to long-term background noise recordings, surface and body waves similar to those from active sources can be extracted from the noise for subsequent imaging. The passive-source method was first proposed

Manuscript received by the Editor 29 July 2024; revised manuscript received 5 August 2025; published ahead of production 26 October 2025; published online 2 March 2026.
[1] Southwest Petroleum University, School of Geoscience and Technology, Chengdu, China. Email: yinghe.wu@foxmail.com; shulinpan@swpu.edu.cn (Corresponding author); dgchenkai@gmail.com; yaojie.chen@foxmail.com.
[2]University of Alberta, Department of Physics, Faculty of Science, Edmonton, Canada. Email: 409791715@qq.com.
[3]The State Key Laboratory of Oil and Gas Reservoir Geology and Exploitation, Chengdu, China.
[4]Xi'an Jiaotong University, Xi'an, China.
[5]Chengdu Technological University, School of Computer Engineering, Chengdu, China. Email: ziyu_qin@126.com.
[6] Chengdu Normal University, School of Computer Science, Chengdu, China.
[7]China University of Petroleum-Beijing, Changping, Beijing, China. Email: 13159962121@163.com; lzyllq123@126.com.

by Aki (1957), who proposed an innovative technique for inverting subsurface velocity structures using seismic noise recordings.

Claerbout (1968) demonstrated that the autocorrelation of seismic recordings transmitted from the subsurface when received at the free surface of a horizontally layered medium is equivalent to the zero-offset record. They referred to this correlation calculation method as "daylight imaging." Rickett et al. (1999) verified Claerbout's theory by obtaining a zero-offset profile that matched the active source profile through autocorrelation. Schuster et al. (2001) processed the virtual shot collection obtained through cross-correlation during the active source imaging process, obtained images of subsurface reflection structures, and renamed the original "daylight imaging" to "seismic interferometry." Bensen et al. (2007) proposed a method for processing background noise data. These studies became the foundation of SI theory.

In the field of natural earthquakes, the passive-source seismic SI method has advanced significantly and has been successfully applied to crustal imaging (Nakata et al., 2011; Olugboji et al., 2016; Lu et al., 2018) and subsurface monitoring (Bohnhoff et al., 2010; Mao et al., 2022). In the field of exploration seismology, due to the ubiquitous surface wave sources, background noise data recorded by geophones often contain a large amount of surface wave information. Therefore, the passive surface wave method based on SI has rapidly advanced and has been widely used in near-surface shear wave velocity modeling (Cheng et al., 2016; Zhang et al., 2019; Mordret et al., 2020; Ning et al., 2021; Wu et al., 2023), reaching a relatively mature stage of development.

Years of research and experiments have proven that passive-source SI body wave imaging has significant application potential in tasks such as subsurface structure imaging (Wegler et al., 2007; Draganov et al., 2009; Xu et al., 2012; Zhang et al., 2022), mine monitoring (Cheraghi et al., 2015; Li et al., 2015; Gu et al., 2021), CO_2 geological and gas storage monitoring (Boullenger et al., 2015; Cheraghi et al., 2017), and oil exploration and development (da Silva et al., 2021; Likhacheva et al., 2021; Wu et al., 2025). Passive body wave imaging still suffers from low signal-to-noise ratios (SNR), limited resolution, and low imaging efficiency, which significantly constrain its large-scale application. Consequently, extracting body waves useful for imaging from seismic background noise has become a prominent research focus, attracting considerable research attention from scholars.

One type of method focuses on filtering the original data to highlight the body wave signal. These include coherence filtering (Nakata et al., 2015), polarization filtering (Dangwal et al., 2021), and filtering based on singular value decomposition (Moreau et al., 2017), among others. Such methods are more suitable for regions with favorable data collection conditions. Another type of method focuses on selective superposition, that is, retaining time panels related to body waves while rejecting time segments dominated by other waves. Effective solutions include beamforming (Rost and Thomas, 2017), illumination diagnosis (Almagro Vidal et al., 2014; Chamarczuk et al., 2019), statistical time series classification (Groos et al., 2012), the local similarity approach (Li et al., 2018), SNR calculation (Liu et al., 2021), and other methods. This type of method is more suitable for areas with poor data collection conditions; however, most of these data selection approaches are either semiautomatic or performed on cross-correlated data, which increases processing costs. Each of these traditional methods has distinct characteristics; however, they are ineffective when applied to large volumes of background noise data spanning tens of days or even months.

Machine learning and deep learning have offer rapid inference capabilities and have been widely applied in recent years to automate various tasks in geophysics, including seismic data denoising, full waveform inversion, seismic imaging, first-break picking, data reconstruction, and seismic inversion, with satisfactory results (Pan et al., 2020; Zhang and Gao, 2021; Zhang et al., 2021; Wu et al., 2022; Liu et al., 2023a, 2023b; Torres et al., 2023). These applications demonstrate the applicability of deep learning in seismic exploration. Deep learning methods have also been used in passive-source seismic exploration; however, they are mostly limited to extracting surface wave signals from natural earthquakes or near-surface noise (Zhang et al., 2020; Chen et al., 2023). Few practical studies have applied deep learning to passive-source body wave exploration. Some scholars have used deep learning networks to recover virtual shot records from short-term noise panels, achieving good synthetic results (Sun et al., 2023; Ma et al., 2024; Zhao et al., 2024).

In regions with complex data acquisition environments, retrieving body wave events is a crucial step; however, traditional methods are insufficient when handling massive background noise data spanning weeks or even months. Deep learning has fast inference capabilities; however, there are challenges in building reliable training data sets. Because high-quality, short-time noise segments containing effective body waves are scarce in collected data and data quality is limited, the cost of generating high-quality training data sets is high. Therefore, there is limited research on using deep learning to retrieve body wave signals from ambient noise. Chamarczuk et al. (2019) applied, for the first time, a support vector machine combined with body wave velocity constraints in the tau-p domain to evaluate a set of virtual shots after correlation. They successfully retrieved body wave events in the collected data. Next, Mezyk et al. (2021) used a convolutional neural network (CNN) combined with a specific visual geometry module to construct a binary classification model. This model analyzes multiple image features of noisy segments fed into the network, enabling body-wave events in the noisy segments to be detected prior to correlation operations and classifying them. However, the above methods require considerable time to produce data sets.

In this study, we propose a novel, fast retrieval and classification method for passive-source seismic body wave events using a convolutional self-attention encoder (CSE). This approach can quickly and robustly extract high-quality body wave events from massive ambient noise and is applied to the real-time monitoring of gas storage. It shows great potential in monitoring scenarios, such as unconventional low-carbon oil and gas development and CO_2 geological storage, providing a low-cost and sustainable novel subsurface dynamic monitoring solution for the energy industry. First, the data set construction process based on illumination diagnosis and incremental learning is introduced, enabling rapid generation of high-quality training data by combining traditional methods with machine learning. Next, the network training mechanism and retrieval-classification process are described, which enhance prediction accuracy through multimodal feature input and optimization of the attention mechanism. Finally, the effectiveness of the proposed method is demonstrated on synthetic data and field data.

METHODS

Review of the SI method

In a closed-domain context characterized by a nonuniform and lossless medium, we assumed acquisition geometry where the source completely surrounds the receiver. The seismic sources are located in the subsurface, with two receivers, *A* and *B*, on the ground surface. By performing cross-correlation on the passive recordings at these two receivers, it is possible to derive Green's function representing the reflection response between them as follows:

$$\{G(x_B,x_A,t)+G(x_B,x_A,-t)\}*\gamma(t)\approx\frac{1}{\rho c}S(x_B,t)*S(x_A,-t), \quad (1)$$

where $G(x_B,x_A,t)$ is the seismic reflection record (Green's function) excited at receiver location x_A and received at receiver location x_B, which is also referred to as the causal part; $G(x_B,x_A,-t)$ is the noncausal part; $\gamma(t)$ denotes the result of the convolution of the noise component contained in the noisy signal; ρ and c are the constant mass density and velocity of the medium, respectively; $S(x_A,t)$ and $S(x_B,t)$ denote the seismic response received by the two receivers at x_A and x_B, respectively; and $*$ denotes the convolution operation.

Currently, equation 1 is typically used to reconstruct Green's function through cross-correlation and stack processing of long-term background noise data. Notably, the reconstruction result is affected by the source wavelet; however, the normalization and spectral whitening applied during the data preprocessing in this study can mitigate this effect. However, few segments in the field data contain effective body wave events, and the abundance of noise affects the reconstructed virtual source function. Therefore, it is crucial to extract body waves that are beneficial for imaging from seismic ambient noise.

Construction of training data sets based on incremental learning

Inspired by incremental learning (Rebuffi et al., 2017), we used a two-step method that combined the traditional illumination diagnosis method with deep learning to construct the training data sets. The illumination diagnosis method (Almagro Vidal et al., 2014) primarily relies on the velocity difference between body and surface waves within the study area. By setting a threshold for the slowness parameter, the method determines whether body waves dominate the cross-correlated noise panels. Building upon this, the 3D illumination diagnosis method (Chamarczuk et al., 2019) enhances accuracy by incorporating arrival time information of subsurface transient sources propagating to different receiver lines. In general, subsurface seismic sources are transient during passive-source data acquisition. The illumination characteristics of these transient sources can be analyzed in the corresponding virtual source panel, described as follows:

$$C^S\left(x_B,x_A,t\right)=\frac{1}{\rho c}\left(u^{\mathrm{obs}}\left(x_A,x_S,-t\right)*u^{\mathrm{obs}}\left(x_B,x_S,t\right)\right). \quad (2)$$

Suppose that a source is located at receiver x_A, emitting energy to multiple receivers x_B within a finite angular window. Let $*$ denote time convolution and $u^{\mathrm{obs}}(x_A,x_S,-t)$ represent the reverse-time wavefield observed at x_A from a transient source at x_S. The ray parameter of the event at time t = 0 s, observed at the position of receiver x_A (virtual source position), provides a measure of the illumination characteristics of the primary transient source captured in a given noise panel. This ray parameter serves as a potential tool for distinguishing between surface waves and body waves. Then, we used the slant-stack transform of field v, $\widetilde{v}(p,\tau)=\int v(x,\ \tau\ +\ px)dx$, where P is the ray parameter, x is the offset, and τ is the intercept time at $P=0$. We subsequently evaluated the slant stack at $\tau=0$ s for each virtual common-source panel C^s:

$$\tilde{C}^S(X_A,P)=\int C^s[x_B,x_A,P\cdot(x_B-x_A)]dx_B, \quad (3)$$

where C^S represents the virtual-source function of the transient source S in the $\tau-p$ domain. $\widetilde{C}^S$ describes the dominant ray-parameter contribution from the transient source to the virtual source located at x_A and recorded at x_B. Based on the expected velocities of body waves and surface waves in the study area, ray parameter limits can be established, and the discrimination criteria defined as follows:

$$C_L^S\left(X_BX_A,t\right)=\begin{cases}0, & if \quad \max\left(\left\|\widetilde{C}_L^S\left(x_A,p\right)\right\|\right)>p_{\mathrm{limit}},\\ \frac{1}{\rho c}\left(u^{\mathrm{obs}}\left(x_A,x_S,-t\right)*u^{\mathrm{obs}}\left(x_B,x_S,t\right)\right), & if \quad \max\left(\left\|\widetilde{C}_L^S\left(x_A,p\right)\right\|\right)<p_{\mathrm{limit}},\end{cases} \quad (4)$$

where $\max(\|\widetilde{C}_L^S(x_A,P)\|)$ is the dominant ray-parameter value and p_{limit} is the expected minimum slowness of P-waves in the recording area. The above steps describe the traditional illumination diagnosis method. Next, we performed this analysis simultaneously on a noise panel recorded across several parallel receiver lines, denoted by the subscript L, thereby extending the illumination analysis to 3D. The time difference between body wave events arriving from the subsurface to different receiving lines on the surface is significantly smaller than that of surface waves arriving at different receiving lines on the surface. Therefore, we scanned the noise panels using a sliding time window, evaluated the ray parameters of the dominant events within each window, and identified the event with the highest slant-stack value across the entire panel. Given the difference in arrival times Δt_i for this event from the ith parallel receiver line and the distance Δx_i between them, the apparent slowness p_i in the crossline direction is estimated as follows:

$$p_i=\frac{\Delta t_i}{\Delta x_i}. \quad (5)$$

According to the predefined body wave speed limit, the final discrimination criterion is defined as follows:

$$S_L^{\mathrm{obs}}\left(t\right)=\begin{cases}0, & if \quad p_1\wedge p_2>p_{\mathrm{limit}},\\ s_{\mathrm{body}}^{\mathrm{obs}}, & if \quad p_1\wedge p_2\le p_{\mathrm{limit}},\end{cases} \quad (6)$$

where $S_L^{\mathrm{obs}}(t)$ represents the noise panel after passing the first step and $s_{\mathrm{body}}^{\mathrm{obs}}$ represents the noise panel classified as containing body wave events. Readers are referred to Almagro Vidal et al. (2014) and Chamarczuk et al. (2019) for a detailed description on the illumination diagnosis procedure.

Because noisy segments containing body wave events are scarce and the illumination diagnosis method requires correlation operations on raw data, the cost of generating a large number of high-quality data sets using this method is significant. Deep learning has the advantage of fast inference. As shown in Figure 1, incremental learning is an efficient deep learning approach that updates the model by gradually learning new data without retraining the entire model. This method is particularly suitable for processing large-scale or continuously evolving data sets, as it can not only significantly reduce running time but also enhance the model's adaptability.

Inspired by incremental learning (Rebuffi et al., 2017) and aiming to generate large, high-quality data sets more rapidly while achieving better training results, we propose the workflow shown in Figure 2. First, a small initial training data set is generated through the illumination diagnosis method, with each sample verified for accuracy. Then, the model is preliminarily trained using the initial data sets to provide it with basic retrieval capabilities. The trained model is subsequently applied to new, unlabeled data to obtain preliminary prediction results.

These prediction results are corrected manually or semi-automatically, and the corrected samples are added to the training data set to further optimize and enhance the model's performance. As model performance improves, new training data are continuously added for testing and correction, gradually expanding the size of the training data set. Each optimization cycle builds on the results of the previous round of corrections and training. After multiple adjustments, the size and quality of the training data set are greatly improved, and the model eventually achieves a stable, high-precision level. This gradual expansion enables the rapid acquisition of large, high-quality training data sets while avoiding retraining on the entire data set, thereby greatly reducing computational costs.

Network architecture

We developed an end-to-end CSE to learn different features from various noise data segments for data classification, enabling efficient data-driven retrieval of body wave events. This approach enables direct classification of data before cross-correlation operations. Figure 3a depicts the network architecture, which primarily consists of three parts: the self-attention mechanism (SAM), the CNN, and fully connected layers. The principle of this model is to feed the preprocessed seismic data and their corresponding actual and imaginary components, obtained via Fourier transformation, into the network. First, the CNN layer extracts nonlinear spatial features. By jointly inputting time-domain and frequency features, the network better captures the characteristics of body wave events. Then, SAM (Vaswani et al., 2017) is applied to the CNN layer output to discard irrelevant information and emphasize features critical for retrieving body wave events. Then, the fully connected layer adjusts the output dimension, and finally, the output corresponds to the binary logic value of different data types. The network architecture is relatively simple, and the running cost is low. Because our objective was to distinguish between segments containing body wave events and other noise segments for binary classification, the loss function of the network is

$$BCELoss = -(y * \log(P) + (1-y) * \log(1-P)), \tag{7}$$

where P is the predicted value and y is the label value. Each CNN layer consists of two convolutional layers, one batch normalization layer and a maximum pooling layer. The hyperparameters for each layer of the network architecture and the training hyperparameters are shown in Tables 1 and 2, respectively.

We used six 2D convolutional layers. The first through sixth layers used 16, 32, 64, 32, 16, and 1 convolution kernels for feature extraction, respectively. Each convolutional layer performs deeper feature extraction on the output of the previous layer. The size of each convolution kernel was 3 × 3, with a stride of 1 and padding of 1. By incorporating SAM, the CNN output layer was used as the query, key, and value vectors to redistribute feature weights. This enables the model to more accurately select and use significant features, thereby improving its performance. The output y of the attention layer is defined as follows:

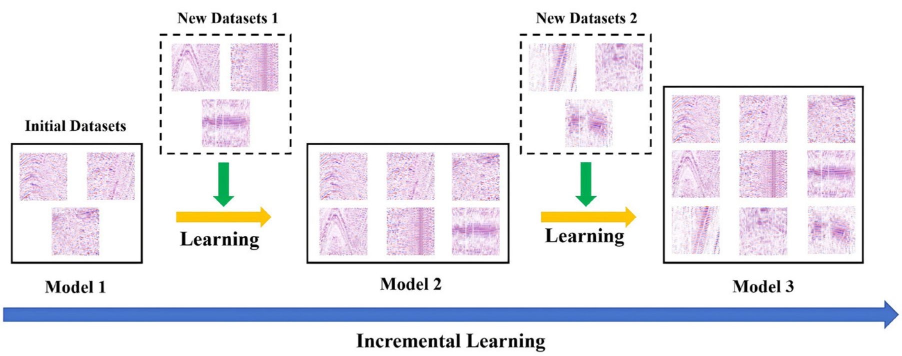

Figure 1. Incremental learning diagram.

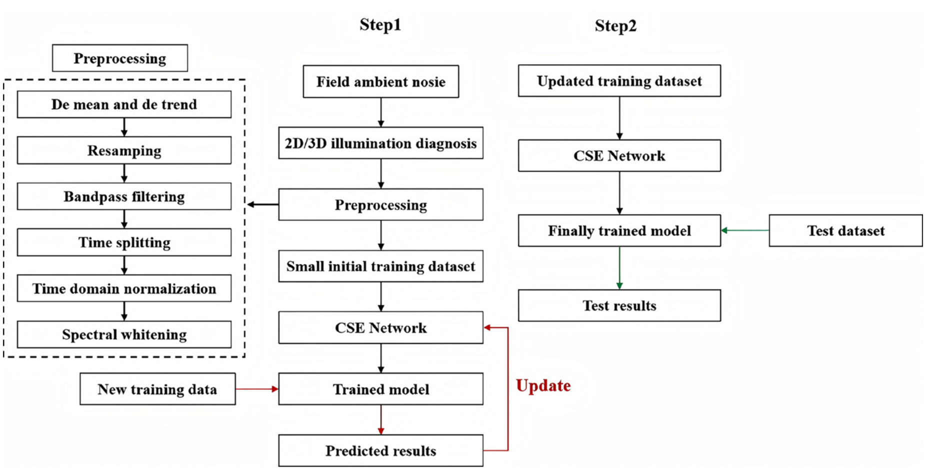

Figure 2. Workflow of dataset construction and training.

$$\begin{cases} I = X \cdot W_I, \quad I = Q, K, V, \\ scores = \dfrac{Q \cdot K^T}{\sqrt{d'}}, \\ \alpha = Soft\max(scores), \\ y = \alpha \cdot V, \end{cases} \tag{8}$$

where X is the CNN output; W_Q, W_K, and W_V are the weight matrices learned during training; Q, K, and V are the query, key, and value matrices, respectively; d' is the dimension of the query and key vectors; and α is the weight of SAM.

Evaluation and classification of retrieval results

In binary classification problems, prediction accuracy is typically evaluated based on the difference between the actual and predicted values. To measure the overall performance of the classifier under different experimental conditions and to highlight the effectiveness and superiority of the proposed method, we introduced five evaluation indicators: runtime, confusion matrix.

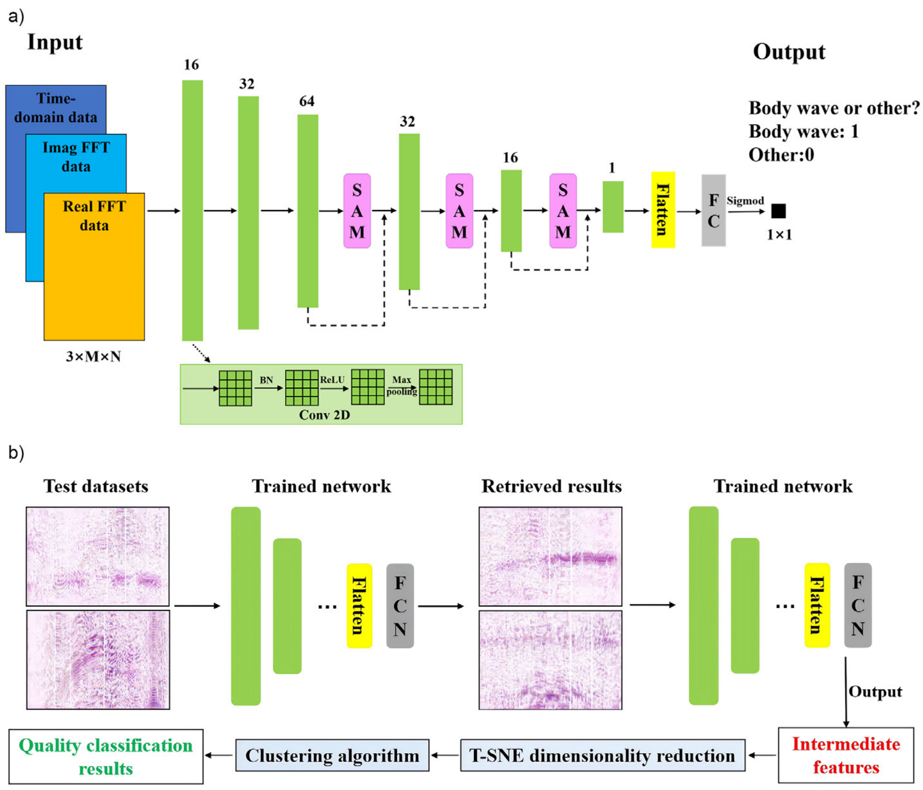

Figure 3. (a) CSE network architecture for body wave event retrieval. (b) The quality classification process of retrieved results.

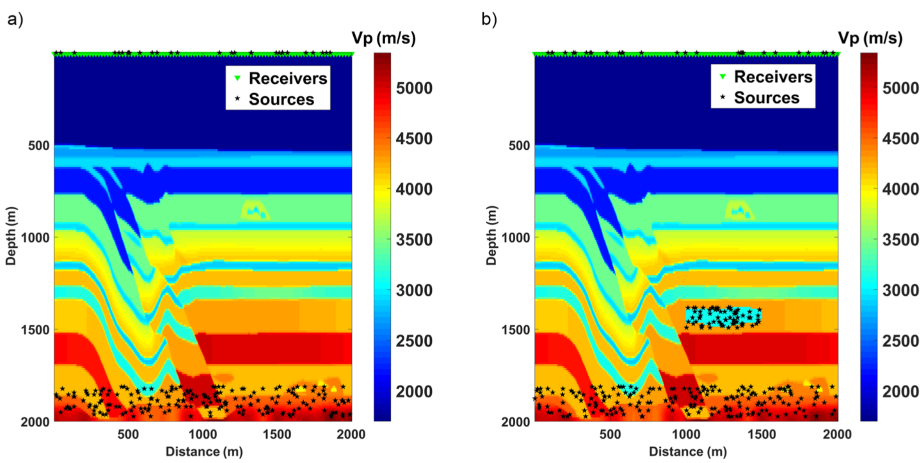

Figure 4. Two-dimensional synthetic P-wave velocity model. (a) Before gas injection. (b) After gas injection.

In addition, as shown in Figure 3b, we reinput the retrieved body wave events into the trained network model and extracted the intermediate features from the deep convolutional layer. After dimensionality reduction, unsupervised clustering analysis was performed to classify the retrieved data effectively. Notably, we output the intermediate features after the flatten operation, with each noise segment corresponding to a row vector of features. This was performed to avoid the loss of key information caused by overly small feature dimensions, as well as to prevent increased computational cost and the introduction of irrelevant features resulting from excessively large dimensions. This approach allows us to discard noise segments with incorrect predictions and body wave event segments with low SNR, retaining only the retrieved segments with high SNR to further improve imaging quality.

To facilitate the visualization and interpretation of the clustering results, we first applied the t-distributed stochastic neighbor embedding (T-SNE) method (Vander Maaten et al., 2008) to capture local similarities between points in high-dimensional space and to separate projected samples represented by detected body wave segments from other irrelevant samples (see Appendix A for T-SNE details). Then, the retrieved body wave events can be distinguished by further clustering the projection points corresponding to the body wave segments.

EXPERIMENTS AND RESULTS

Synthetic data test

To verify the feasibility of the method, we established a synthetic model simulating the process before and after gas injection in the storage facility and generated extensive synthetic background noise data for testing. During

Table 1. Main structural hyperparameters of the network.

Stage	Module	Hyperparameters
Network architecture	Conv2D_1	Input channels: 3 Output channels: 16
	BN	Number of features: 16
	Max pooling	Pooling size: 3 Stride:2
	Conv2D_2	Input channels: 16 Output channels: 32
	BN	Number of features: 32
	Max pooling	Pooling size: 3 Stride: 2
	Conv2D_3	Input channels: 32 Output channels: 64
	BN	Number of features: 64
	Max pooling	Pooling size: 3 Stride: 2
	Self-attention layer	Hidden dimension: 64 Output dimension: 64
	Conv2D_4	Input channels :64 Output channels: 32
	BN	Number of features: 32
	Max pooling	Pooling size: 3 Stride: 2
	Self-attention layer	Hidden dimension: 32 Output dimension: 32
	Conv2D_5	Input channels: 32 Output channels: 16
	BN	Number of features: 16
	Max pooling	Pooling size: 3 Stride: 2
	Self-attention layer	Hidden dimension: 16 Output dimension: 16
	Conv2D_6	Input channels: 16 Output channels:1
	BN	Number of features: 1
	Max pooling	Pooling size: 3 Stride: 2
	Full connected layer	Output: 1

Table 2. List of training hyperparameters.

Stage	Hyperparameters
Training	Optimizer: Adam
	Batch size: 8
	Learning rate: 0.001
	Loss function: BCE
	Epochs: 50
	Activation function: Sigmod
	GPU: 4×1080Ti

gas injection into the ground, the seismic velocity state of the reservoir changes; therefore, a new model is created by modifying the formation velocity after gas injection. Because gas injection is a complex process, we focused on quickly retrieving noise segments containing body wave events and used only simple velocity changes to replace the changes in the ground before and after gas injection.

To more intuitively observe the changes before and after gas injection, a 2D synthetic model was established for experiments. Figure 4a shows the original model, whereas Figure 4b depicts the model after gas injection, where the P-wave velocity in the target gas injection area decreased by 800 m/s. We used the finite-difference 2D elastic wave seismic wave field simulation method. Notably, the field data were collected as single-component data; therefore, only the Z-component data were used. To better simulate actual subsurface conditions, 300 seismic sources were randomly distributed on the surface (generating noise such as surface waves), at injection locations, and at the bottom of the model.

The 101 receivers were located on the surface, spaced 20 m apart, with a sampling interval of 4 ms. The source wavelet was a Ricker wavelet with a dominant frequency of 35 Hz. To obtain long-term background noise recordings for testing, passive-source seismic records of 4 h each were simulated on the model before and after gas injection through the passive-source observation system. Surface wave noise generated by different seismic sources was randomly superimposed and added to the passive-source recordings from different time periods to create segments dominated alternately by surface waves and body waves.

For comparison, we first simulated active-source shot records and performed reverse time migration (RTM) imaging (Figure 5). As shown in Figure 5b, a new reflection event emerged at approximately 1.2–1.3 s, resulting from strong reflection coefficients caused by reduced formation velocity. This velocity reduction induces phase variations in subsurface reflection events, causing misalignment between events before and after the velocity change. Notably, the position of this new reflection event in Figure 5d perfectly corresponds to the designed gas injection model.

Next, the long-term ambient noise was divided into 5-s-long segments, and all segments were interfered and stacked to obtain the passive-source virtual shot records shown in Figure 6, followed by RTM imaging. Owing to the influence of surface noise sources, the passive-source imaging results, compared with those from active sources, were largely dominated by noise such as surface waves. The effective event axis was mostly obscured, and no effective signal could be observed.

We then applied the workflow shown in Figure 1 to create data sets and train the CSE network shown in Figure 2. The data sets were divided into a training set and a validation set at a 5:1 ratio, and 2 h of data before and after gas injection were retained as the test set. Figure 7 illustrates the segments containing body wave events and those without obvious body wave events, as identified by the illumination diagnosis method for initial training. The absolute value of the slowness corresponding to the energy peak was smaller for segments dominated by body waves, that is, the peak appeared in the central region. Spectral analysis showed that segments containing body wave events differed from other segments, supporting the inclusion of frequency-domain information for training. The corresponding outputs when generating labels were 1 and 0, respectively.

To demonstrate the effectiveness and superiority of our method, we conducted ablation experiments to validate the contribution of each component and compared the results with those of traditional methods (see Appendix A for specific experimental information). The results indicated that incorporating attention mechanisms and using incremental learning to expand the data sets can significantly improve network performance. Furthermore, integrating time-domain and Fourier transform data for training can further optimize prediction results without significantly increasing computational time.

Compared to traditional methods, the proposed CSE network offers greater efficiency in body wave retrieval, with its time-saving advantage becoming increasingly significant as the volume of predicted data increases. Next, we reconstructed virtual shot records using the 617 segments containing body wave events retrieved by the CSE network from a total of 2880 segments. As depicted in Figure 8, the reflection signal of the imaging results after retrieval was recovered; however, the resolution and SNR were not as good as those of active-source recordings. A new strong reflection event appeared at 1.2 s, providing a basis for subsequent monitoring.

To further improve the quality of the virtual shot records, we classified the segments with body wave events retrieved by the CSE network. Figure 9 illustrates the T-SNE dimensionality reduction and clustering of intermediate feature data for the retrieved body wave events, successfully dividing them into three categories. The corresponding noise segments in the time domain are shown in Figure 10. Finally, we discarded segments with identification errors and low SNR, retaining only the higher-quality segments containing body wave events for virtual shot record reconstruction and imaging.

As shown in Figure 11, the SNR data were improved, and the abnormal areas that changed in the imaging results are clearly presented, providing reliable information for monitoring. Notably, constrained by the fundamental synthetic assumptions of SI (e.g., the need for sufficiently long valid body wave recordings and a uniform spatial distribution of subsurface noise sources), the reconstructed virtual shot records and corresponding imaging in Figure 11 do not fully reach the quality standard of the active-source results shown in Figure 5. However, compared to the original passive-source results presented in Figure 6, substantial improvements are evident in the virtual shot records and imaging quality, validating the efficacy of our methodology.

Field data test

Encouraged by the synthetic data retrieval results, we further verified the reliability and practicability of the proposed method

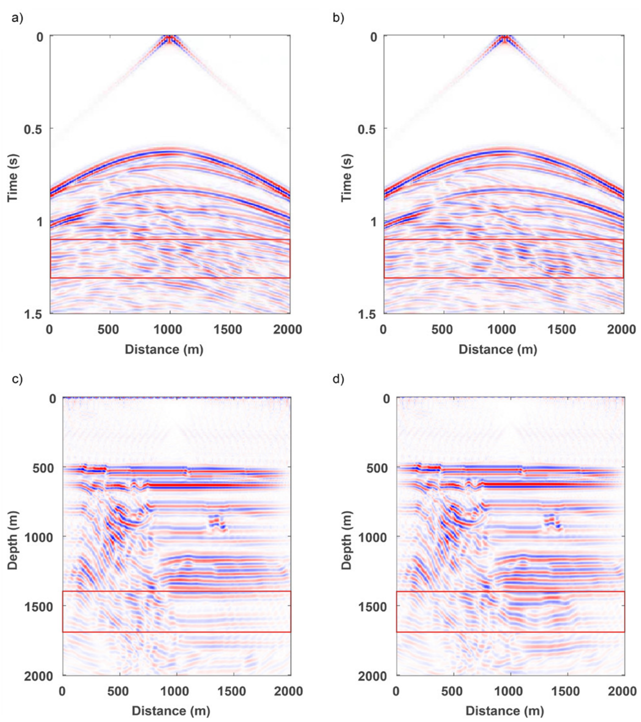

Figure 5. (a) Active source shot records acquired before gas injection. (b) Active source shot records acquired after gas injection. (c) Depth-domain RTM image generated from (a). (d) Depth-domain RTM image generated from (b).

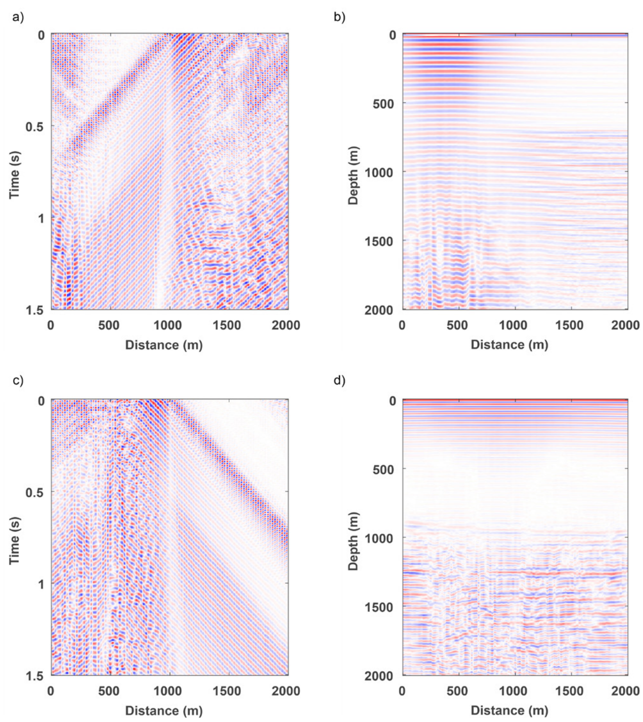

Figure 6. (a) Reconstructed virtual shot record before gas injection (before data screening). (b) Depth-domain RTM image generated from (a). (c) Reconstructed virtual shot record after gas injection (before data screening). (d) Depth-domain RTM image generated from (c).

by applying it to the body wave retrieval task using actual ambient noise. These data were obtained through a 3D linear array deployed at a gas storage project in Northwest China, with the observation system shown in Figure 12. For data collection, we selected the SmartSolo high-resolution single-component intelligent seismic sensor. There were 16 receiving lines, with a sampling interval of 1 ms, and data collection spanned several months.

For the experiments, we selected two 48-h periods of data from three lines (L3–L5) recorded at different times. The raw data and their corresponding frequency domain characteristics are shown in Figure 12b and 12c. Spectral analysis showed that the raw data exhibited a relatively low SNR. The main frequency band was concentrated between 0 and 30 Hz, with dramatic spectral changes. Frequencies between 0 and 10 Hz mostly consisted of low-frequency noise and were considered uninformative. Consequently, a bandpass filter of 10–30 Hz was applied.

To improve processing efficiency, each line used data from only 140 geophones (see the red box in Figure 12a). The long-term seismic data were divided into 5-s long segments, and the data were resampled from 1 to 8 ms after preprocessing. Similarly, we applied the workflow shown in Figure 1 to create the data sets and used the CSE network shown in Figure 2 for training and prediction. The data sets were divided into a training set and a validation set in a 5:1 ratio, and 2 days of 20-h-long data were retained as the test set. Figure 13 shows the positive and negative samples generated for initial training using the 3D illumination diagnosis method.

Due to the poor quality of the collected data, we produced a total of 195 segments containing valid body wave events from the 5760 segments in the 8-h-long data of the L4 line in the training set. Then, we

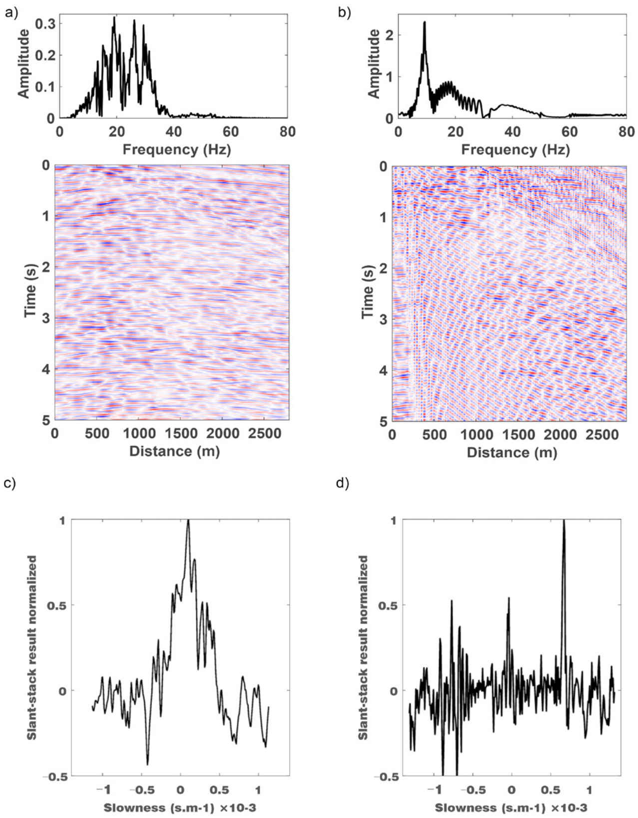

Figure 7. (a) Samples with body wave events and corresponding spectrum. (b) Samples without significant body wave events and corresponding spectrum. (c) Illumination diagnosis for samples in (a). (d) Illumination diagnosis for samples in (b).

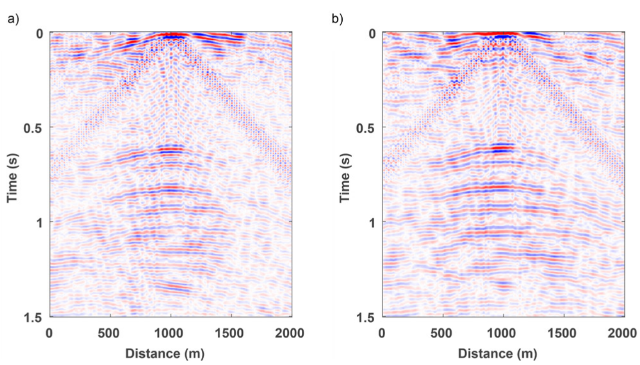

Figure 8. Reconstructed virtual shot records by CSE retrieved results. (a) Before gas injection. (b) After gas injection.

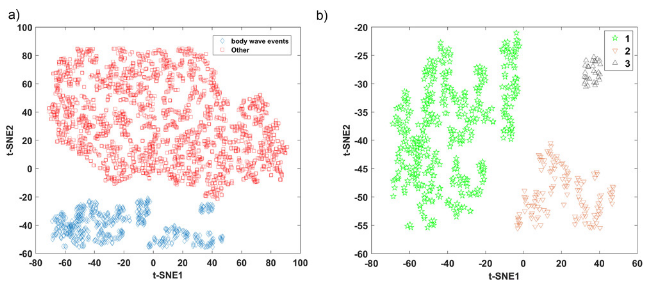

Figure 9. (a) Results of T-SNE dimensionality reduction of the intermediate features corresponding to the prediction results. (b) Results of clustering the intermediate features corresponding to the retrieved body wave events using the K-means clustering method.

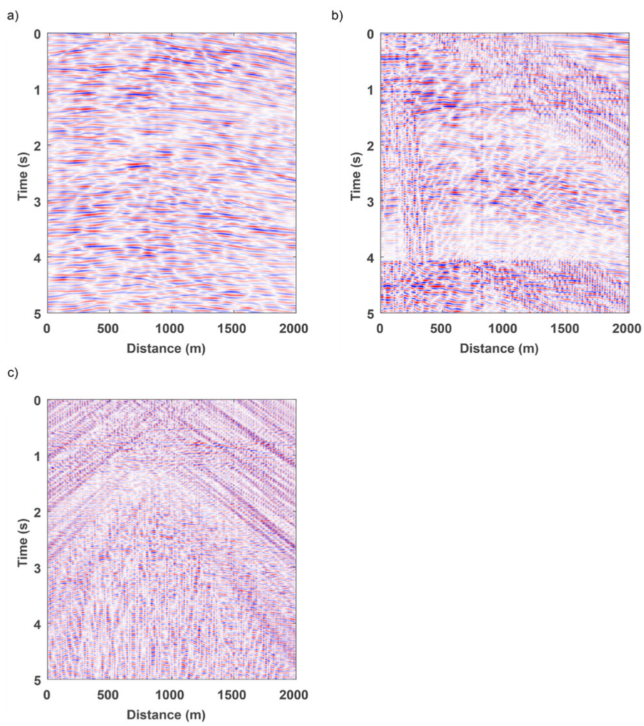

Figure 10. Quality classification of retrieved results. (a) High-quality segment. (b) Segment with low SNR. (c) Misidentified segment.

used the hyperparameters presented in Table 1 for training. The experimental results shown in Table 3 (note that, due to the large amount of field test data, we calculated only precision, recall, and the F1 score for comparison) indicate that the proposed method successfully retrieved 987 body wave events from the 40-h-long L4 line test data (28,800 segments) over 2 days, demonstrating high recall. Compared to the conventional method, our method greatly improved work efficiency.

Next, we classified the segments containing body wave events retrieved by the CSE network. Figure 14 shows the T-SNE dimensionality reduction and clustering process of the intermediate feature data corresponding to these events, which were successfully classified into four categories. The corresponding noise segments in the time domain are shown in Figure 15. We subsequently discarded segments with recognition errors and low SNR for imaging. Due to limitations in data quality, the process of recovering virtual shot records and imaging not only amplifies invalid noise but also results in the loss of useful information. Therefore, we directly performed auto-correlation stack imaging (equivalent to zero-offset imaging) on the high-quality segments containing body wave events. To verify the reliability and accuracy of the imaging results, we first presented the active-source profile corresponding to the L4 line (Figure 16a), where the red box marks the reservoir location at a depth of approximately 1 km within a time window of 0.8–1 s.

The passive-source imaging results (Figure 16b-16e) showed that, due to the lower main frequency band of the passive monitoring data and the fact that actual data acquisition conditions deviated further from the SI hypothesis compared to synthetic experiments, there were significant differences between the imaging results and the active-source profile. Therefore, we adopted the event axis comparison method rather than direct image subtraction to analyze subsurface changes. Specifically, we first marked the strong amplitude events caused by prominent reflection interfaces (indicated by black lines) on the active-source profile and then projected them onto the passive-source profile for comparison.

The analysis indicated that the full-day data imaging results without effective noise segment retrieval (Figure 16b and 16c) exhibited the following issues: the event axis showed poor consistency with the active source; there were anomalous differences in the event axis between the 2 days' imaging results in the shallow layers and marked positions, which did not align with the actual geological conditions. In contrast, the imaging results based on the retrieved noise segments (Figure 16d and 16e) demonstrated significant improvements: the event axis exhibited good consistency with the active source; the event axis in the shallow layers and marked positions remained stable across the 2 days' imaging results, with noticeable changes only beginning at the green line position. In particular, the variations in the event axis within the red box clearly reflected the impact of gas injection operations on the reservoir, and the imaging results were reasonable. Application to field data showed that the method proposed in this study correctly retrieves noise segments containing body wave events, improve the quality of passive-source imaging, provide a reliable basis for time-lapse analysis, and validate the practicality and effectiveness of the method.

DISCUSSION

Note that the method proposed in this study can not only effectively retrieve body wave events from massive passive seismic data but also establish an effective passive-source time-lapse monitoring technique to reveal spatiotemporal subsurface changes.

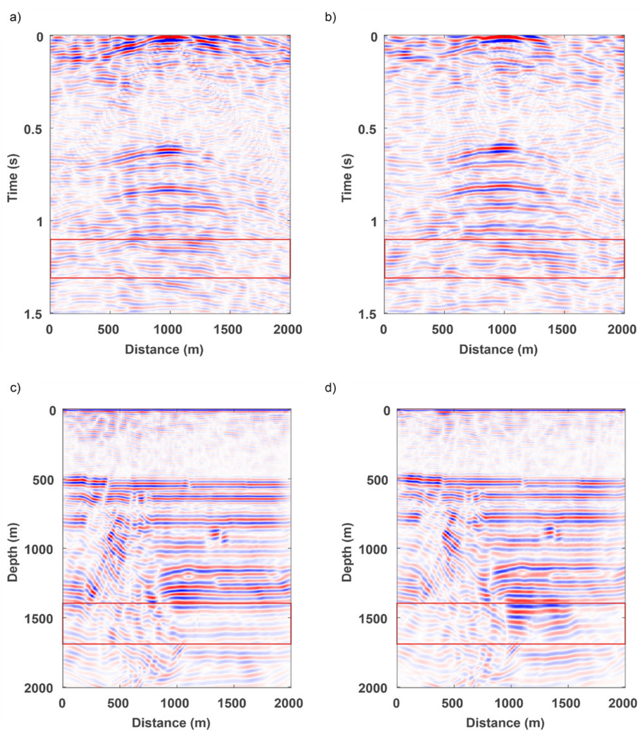

Figure 11. (a) Final reconstructed virtual shot records before gas injection. (b) Final reconstructed virtual shot records after gas injection. (c) Final depth domain RTM imaging result before gas injection. (d) Final depth domain RTM imaging result after gas injection.

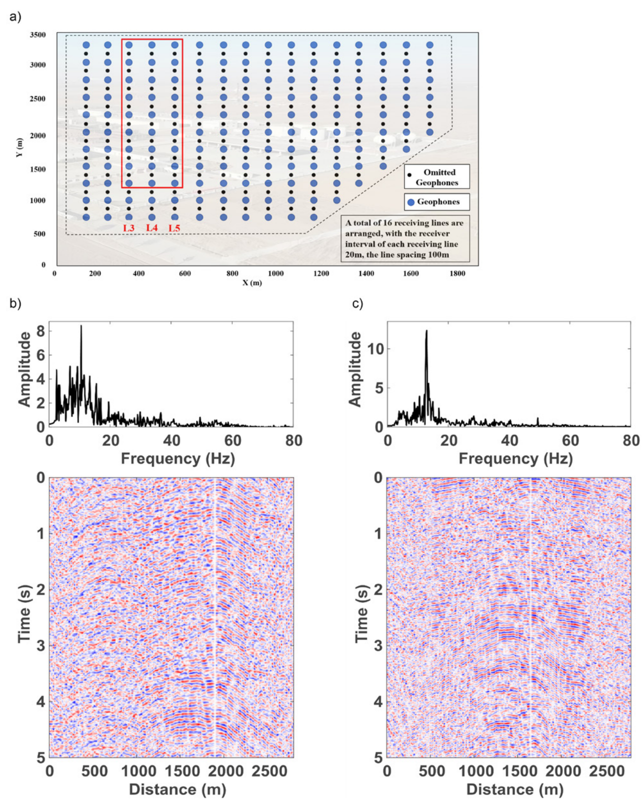

Figure 12. (a) Field observation system. (b and c) Raw data and corresponding spectrum at different time periods.

In the context of global efforts to promote energy transition and address climate change, this method may prove particularly valuable across various applications. For example, in the field of CO_2 geological storage, it can help evaluate the integrity of CO_2 storage bodies, mitigate potential leakage risks, and thus assist in achieving carbon neutrality goals. In oilfield development scenarios, this technique can estimate changes in subsurface oil reservoirs to provide data support for optimizing production plans, thereby improving oil and gas recovery rates. Additionally, it can reduce resource waste and carbon emissions during energy exploration and production processes. In these applications, efficient and rapid retrieval of body waves is crucial for estimating real-time subsurface changes. The CSE-based body wave event retrieval scheme has demonstrated its capability to intelligently identify noise segments dominated by body wave energy within massive noisy data sets, thereby improving the quality of passive-source imaging results.

Traditional SI relies on the assumptions of long-term effective recording and uniform spatial distribution of seismic sources. Consequently, virtual cannon recordings reconstructed in passive-source time-lapse monitoring generally exhibit inferior quality compared to that of active-source data. This is also the main reason for the scarcity of practical applications of this technology. Due to limitations in existing active-source data and logging data, as well as the inherent resolution constraints of passive-source imaging, we have not yet been able to quantitatively interpret the monitoring results. However, by innovatively solving the problem of effective data recognition and combining autocorrelation imaging with phase axis analysis techniques, we have achieved a qualitative and reasonable interpretation of

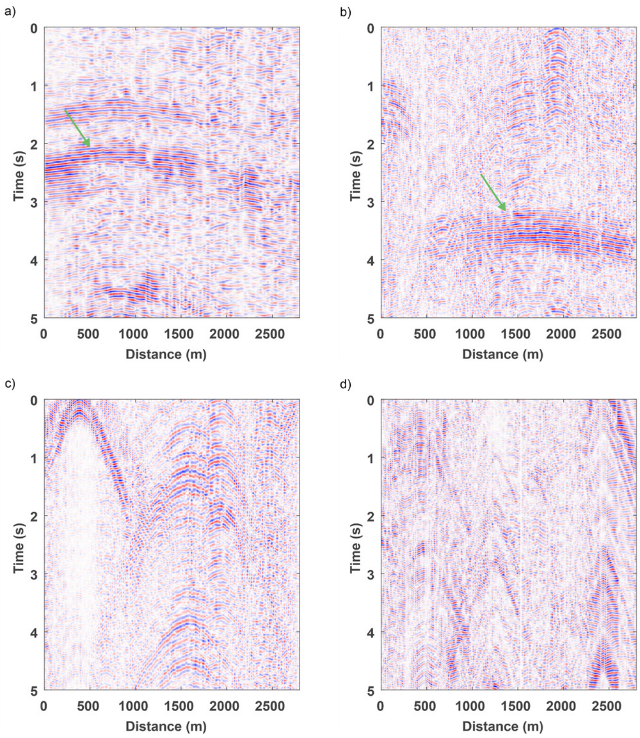

Figure 13. (a and b) Samples with body wave events (green arrows are body wave events). (c and d) Samples without significant body wave events.

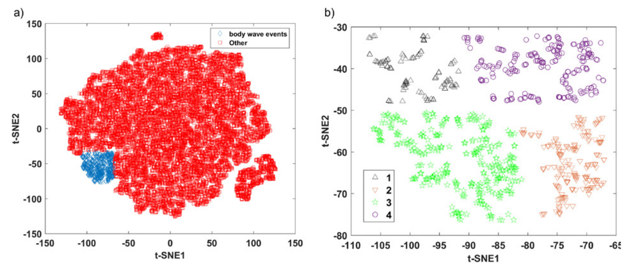

Figure 14. (a) Results of T-SNE dimensionality reduction of the intermediate features corresponding to the prediction results. (b) Results of clustering the intermediate features corresponding to the retrieved body wave events using the K-means clustering method.

Table 3. Comparison of field data experimental results.

Method	Train data sets (Size of positive samples)	Data type	Time (train + test) (h)	Accuracy (%)	Precision (%)	Recall (%)	F1 value
Traditional	/	Time	9.5	100	100	100	1
CSE	Updated (3 × 215 × 625 × 101)	Time + FFT	1.5	/	93.2	96.5	0.948

Table 4. The ablation experiment results of the test datasets.

Method	Train datasets (size of positive samples)	Data type	Time (train + test) (min)	Accuracy (%)	Precision (%)	Recall (%)	F1 value
Traditional	/	Time	57	100	100	100	1
DL (no SAM)	Initial (184 × 1250 × 101)	Time	9	76.3	73.8	75.6	0.747
CSE	Updated (184 × 1250 × 101)	Time	11	80.6	78.5	80.3	0.794
CSE (incremental learning)	Updated (566 × 1250 × 101)	Time	26	93.1	91.6	92.8	0.921
CSE （frequency domain）	Updated (2 × 566 × 1250 × 101)	FFT	25	91.9	89.4	90.9	0.907
CSE (this study)	Updated (3 × 566 × 1250 × 101)	Time + FFT	36	**98.1**	**95.5**	**98.4**	**0.969**

passive-source time-lapse imaging, providing key methodological guidance for the development of the field. Notably, this study primarily focused on the preprocessing of passive-source data and does not further explore methods to improve the accuracy and resolution of time-lapse monitoring.

The academic community is actively exploring high-precision reconstruction techniques for passive-source data to achieve results comparable to those of active-source exploration, such as using deep learning to overcome the synthetic limitations of traditional SI (Sun et al., 2023; Ma et al., 2024; Zhao et al., 2024). Although such research has achieved results comparable to those of active-source shot records in synthetic data experiments, there have been no public reports of applications to actual data. Nevertheless, this is of great significance for advancing the practical use of passive-source wave exploration and warrants further research and breakthroughs.

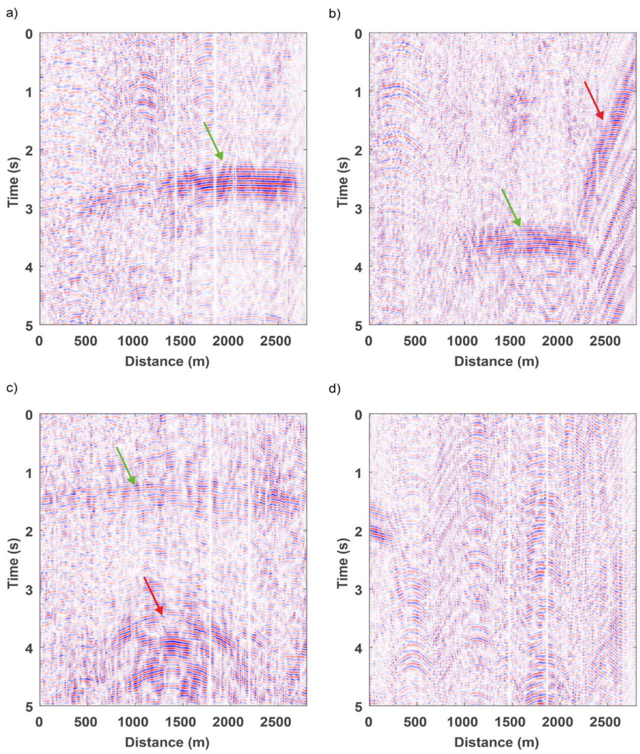

Figure 15. Quality classification of retrieved results (green arrows are body wave events and red arrows are dominant noise). (a) High-quality segment. (b) Segment with significant air wave. (c) Segment with significant surface wave. (d) Misidentified segment.

CONCLUSION

This study proposes a fast retrieval and classification method for passive-source seismic body wave events using CSE, enabling direct classification of noise segments prior to cross-correlation operations. To better learn features, we input multimodal data consisting of three components: time-domain data and the real and imaginary components obtained through Fourier transformation. During data set generation, we used illumination diagnosis and incremental learning methods to rapidly update and construct high-quality training data sets. In addition, we visualized the intermediate features extracted by the deep convolution layer corresponding to the retrieved body wave events and performed unsupervised clustering to complete the quality classification, which was then used to further improve imaging quality. The application results on synthetic and field data demonstrate that the method offers high accuracy and efficiency, significantly improves the SNR of the imaging results, and provides a solid foundation for subsequent monitoring and interpretation. It shows great potential for application in low-carbon exploration and monitoring using passive-source SI in the context of energy transition scenarios.

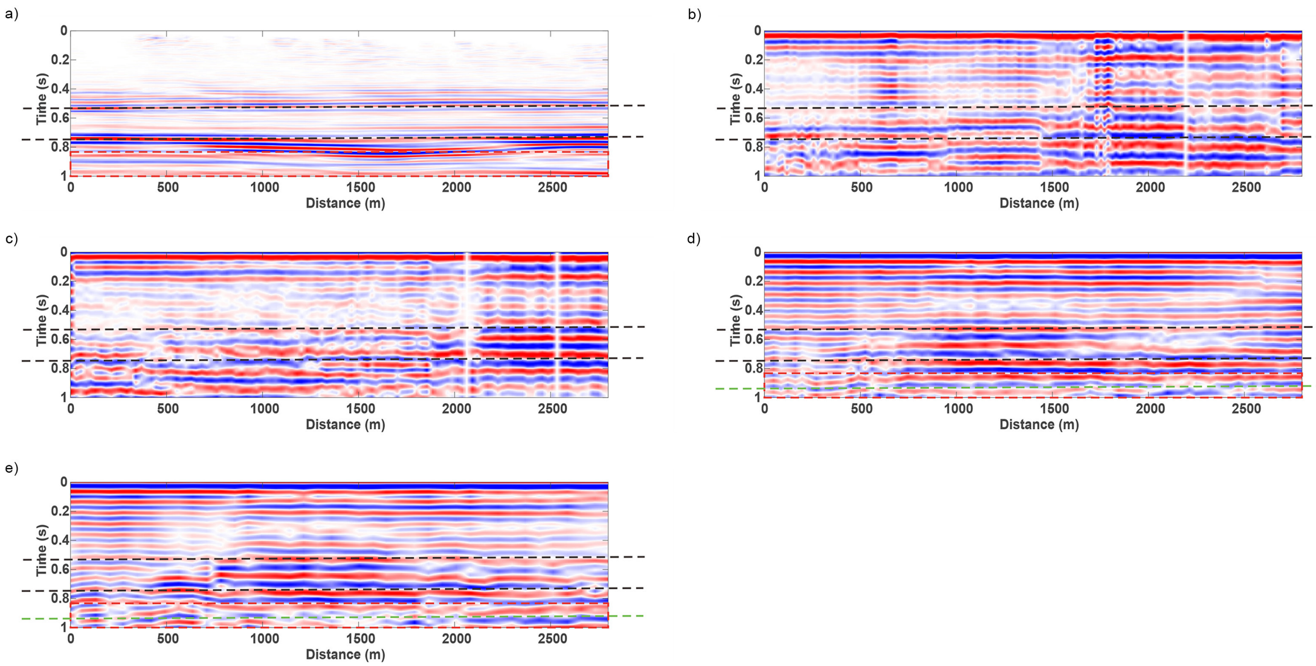

Figure 16. (a) Field active source poststack profile. (b) August 28th autocorrelation imaging result (before processing). (c) October 5th autocorrelation imaging result (before processing). (d) August 28th autocorrelation imaging result (after processing). (e) October 5th autocorrelation imaging result (after processing).

FUNDING INFORMATION

This work was supported in part by the Technology Cooperation Project of the CNPC-SWPU Innovation Alliance (2020CX020000) and in part by the Natural Science Foundation of Sichuan Province (24NSFSC0808). Yinghe Wu was supported by the China Scholarship Council (202306440144).

DATA AND MATERIALS AVAILABILITY

Data associated with this research are confidential and cannot be released.

APPENDIX A

Basic principle of T-SNE

The core idea of T-SNE is to define a probability distribution over pairs of data points in the high-dimensional space to represent their similarities and then construct a corresponding distribution in a low-dimensional space. By minimizing the difference between these distributions using Kullback–Leibler (KL) divergence, the algorithm maps high-dimensional data into a low-dimensional space for visualization.

Step 1. Calculate similarities in high-dimensional space:

For each data point x_i, we calculate the conditional probability $p_{j|i}$ of each other point x_j given x_i. This probability represents the likelihood that point x_j is a neighbor of point x_i. It is calculated using the following formula:

$$p_{j|i} = \frac{\exp(-\|x_i - x_j\|^2 / 2\delta_i^2)}{\sum_{k \neq i} \exp(-\|x_i - x_k\|^2 / 2\delta_i^2)}, \tag{9}$$

where δ_i is the Gaussian variance with data point x_i as the center, which determines the neighborhood range. The value of δ_i may differ for each point x_i and is determined based on "perplexity." Finally, to obtain a symmetric similarity matrix, we average $p_{j|i}$ and $p_{i|j}$ to obtain p_{ij}:

$$p_{ij} = \frac{p_{j|i} + p_{i|j}}{2n}. \tag{10}$$

Using this approach, a symmetric similarity matrix is obtained in which each element p_{ij} reflects the similarity between data points x_i and x_j in high-dimensional space.

Step 2. Calculate the similarity between points in low-dimensional space:

In low-dimensional space, we calculate the similarity q_{ij} between points y_i and y_j as follows:

$$q_{ij} = \frac{(1 + \|y_i - y_j\|^2)^{-1}}{\sum_{k \neq l} (1 + \|y_k - y_l\|^2)^{-1}}. \tag{11}$$

Next, the T-SNE algorithm minimizes the difference between the similarity matrix p_{ij} and the similarity matrix q_{ij} to obtain an appropriate low-dimensional space representation.

Step 3. Optimize the positions of points in low-dimensional space:

The positions of points in low-dimensional space are optimized by minimizing the KL divergence. KL divergence is used to measure the difference between the similarity distributions in high- and low-dimensional spaces:

$$C = \sum_{i \neq j} p_{ij} \log \frac{p_{ij}}{q_{ij}}. \tag{12}$$

Next, the gradient descent method is used to minimize the KL divergence and update the positions of the points in low-dimensional space:

$$\frac{\delta C}{\delta y_i} = 4 \sum_j (p_{ij} - q_{ij})(y_i - y_j)(1 + \|y_i - y_j\|^2)^{-1}. \tag{13}$$

Ablation experiment details

To clarify the contributions of each component of the CSE network, we conducted ablation experiments while keeping other training parameters constant. The experimental results were compared with those of a traditional method, and the relevant data are summarized in Table 4. In the experiments, after incorporating the attention mechanism into the basic network architecture, key indicators such as the accuracy and recall rate in the body wave retrieval task improved significantly, demonstrating its effectiveness in focusing on key information.

Through incremental learning via gradually increasing the size of the training data set, the prediction accuracy of the network for body wave events improved substantially, indicating that this approach facilitates learning more sample features and enhances generalization ability. When training the network separately using only time-domain data and Fourier transform data, we found that the difference in prediction accuracy between the two was minimal. However, combining time-domain and Fourier transform data for training further improved the network's prediction results without significantly increasing training or prediction time. Additionally, predictions made using a small training data set contained errors that required manual screening. However, as the volume of prediction data increased, the method's time-cost and overall efficiency advantages in body wave retrieval tasks became increasingly significant, fully demonstrating its superiority compared to traditional methods.

REFERENCES

Aki, K, 1957, Space and time spectra of stationary stochastic waves, with special reference to microtremors: Bulletin of the Earthquake Research Institute, **35**, 415–456.

Almagro Vidal, C., D. Draganov, J. vander Neut, et al, 2014, Retrieval of reflections from ambient noise using illumination diagnosis: Geophysical Journal International, **198**, 1572–1584, doi: 10.1093/gji/ggu164.

Bensen, G. D., M. H. Ritzwoller, M. P. Barmin, et al, 2007, Processing seismic ambient noise data to obtain reliable broad-band surface wave dispersion measurements: Geophysical Journal International, **169**, 1239–1260, doi: 10.1111/j.1365-246X.2007.03374.x.

Bohnhoff, M., G. Dresen, W. L. Ellsworth, et al, 2010, Passive seismic monitoring of natural and induced earthquakes: case studies, future directions and socio-economic relevance: New Frontiers in Integrated Solid Earth Sciences, 261–285, doi: 10.1007/978-90-481-2737-5_7.

Boullenger B., A. Verdel, B. Paap, et al, 2015, Studying CO_2 storage with ambient-noise seismic interferometry: A combined numerical feasibility study and field-data example for Ketzin, Germany: Geophysics, **80**, no. 1, Q1–Q13, doi: 10.1190/geo2014-0181.1.

Chamarczuk, M., M. Malinowski, Y. Nishitsuji, et al, 2019, Automatic 3D illumination-diagnosis method for large- N arrays: Robust data scanner and machine-learning feature provider: Geophysics, **84**, no. 3, Q13–Q25, doi: 10.1190/geo2018-0504.1.

Chen, X. H., H. X. Xing, J. Y. Pang, et al, 2023, Detection of low-velocity layer using passive surface-wave data and deep learning: An application

in the Urban Area of Hangzhou, China: In Journal of Physics: Conference Series, **2651**, 012028, doi: 10.1088/1742-6596/2651/1/012028.
Cheng, F., J. H. Xia, Y. H. Luo, et al, 2016, Multichannel analysis of passive surface waves based on crosscorrelations: Geophysics, **81**, no. 5, EN57–EN66, doi: 10.1190/geo2015-0505.1.
Cheraghi, S., A. J. Craven, and G. Bellefleur, 2015, Feasibility of virtual source reflection seismology using interferometry for mineral exploration: A test study in the Lalor Lake VMS mining area, Manitoba, Canada: Geophysical Prospecting, **63**, 833–848, doi: 10.1111/1365-2478.12244.
Cheraghi, S., J. White, D. Draganov, et al, 2017, Passive seismic reflection interferometry: A case study from the Aquistore CO_2 storage site, Saskatchewan, Canada: Geophysics, **82**, no. 3, B79–B93, doi: 10.1190/geo2016-0370.1.
Claerbout, J. F, 1968, Synthesis of a layered medium from its acoustic transmission response: Geophysics, **33**, no. 2, 264–269, doi: 10.1190/1.1439927.
da Silva, D. W. T., A. F. do Nascimento, and F. L. de Santana, 2021, Characterization of seismic noise in an oil field using passive seismic data from a hydraulic fracturing operation: Pure and Applied Geophysics, **178**, no. 10, 3849–3868, doi: 10.1007/s00024-021-02854-4.
Dangwal, D., and M. Behm, 2021, Interferometric body-wave retrieval from ambient noise after polarization filtering: Application to shallow reflectivity imaging: Geophysics, **86,** no. 6, Q47–Q58, doi: 10.1190/geo2020-0768.1.
Draganov, D., X. Campman, J. Thorbecke, et al, 2009, Reflection images from ambient seismic noise: Geophysics, **74**, no. 5, A63–A67, doi: 10.1190/1.3193529.
Groos, J. C., S. Bussat, and J. R. R. Ritter, 2012, Performance of different processing schemes in seismic noise cross-correlations: Geophysical Journal International, **188**, no. 2, 498–512, doi: 10.1111/j.1365-246X.2011.05288.x.
Gu, N., M. Chamarczuk, J. Gao, et al, 2021, Passive seismic structure imaging of a coal mine by ambient noise seismic interferometry on a dense array: Acta Geologica Sinica, **95**, no. S1, 37–39, doi: 10.1111/1755-6724.14826.
Li, X. L., D. D. Ling, and D. C. Li, 2015, Application of interferometric imaging in Shuangan goaf detection. Progress in Geophysics (in chinese): **30**, 2188–2192.
Li, Z. F., Z. G. Peng, D. Hollis, et al, 2018, High-resolution seismic event detection using local similarity for Large-N arrays: Scientific Reports, **8**, no. 1, 1646, doi: 10.1038/s41598-018-19728-w.
Likhacheva, O., G. Kashin, and G, V. Mironychev, 2021, Use of passive seismic to identify oil-bearing reservoirs in the Udmurt Republic, Russia: Journal of Geotechnology and Energy, **38**, no. 1, doi: 10.7494/jge.2021.38.1.4290.
Liu, D. W., M. D. Sacchi, X. K. Wang, et al, 2023, Unsupervised deep learning for ground roll and scattered noise attenuation: IEEE Transactions on Geoscience and Remote Sensing, **61**, 1–13, doi: 10.1109/TGRS.2023.3325324.
Liu, D. W., W. L. Niu, X. K. Wang, et al, 2023, Improving vertical resolution of vintage seismic data by a weakly supervised method based on cycle generative adversarial network: Geophysics, **88**, no. 6, V445–V458, doi: 10.1190/geo2023-0006.1.
Liu, G. F., Y. Liu, X. H. Meng, et al, 2021, Surface wave and body wave imaging of passive seismic exploration in shallow coverage area application of Inner Mongolia: Chinese Journal of Geophysics, **64**, 937–948.
Lu, Y., L. Stehly, A. Paul, et al, 2018, High-resolution surface wave tomography of the European crust and uppermost mantle from ambient seismic noise: Geophysical Journal International, **214**, no. 2, 1136–1150, doi: 10.1093/gji/ggy188.
Ma, L. Y., L. G. Han, and Q. Feng, 2024, Real-time passive seismic interferometry with deep transfer learning: IEEE Transactions on Geoscience and Remote Sensing, **62**, 1–9, doi: 10.1109/TGRS.2024.3365688.
Mao, S., A. Lecointre, RD van der Hilst, et al, 2022, Space-time monitoring of groundwater fluctuations with passive seismic interferometry: Nature Communications, **13**, no. 1, 4643, doi: 10.1038/s41467-022-32194-3.
Mężyk, M., M. Chamarczuk, and M. Malinowski, Automatic image-based event detection for large-N seismic arrays using a convolutional neural network: Remote Sensing, **13**, no. 3, 389, doi: 10.3390/rs13030389.
Mordret, A., R. Courbis, F. Brenguier, et al, 2020, Noise-based ballistic wave passive seismic monitoring-Part 2: Surface waves: Geophysical Journal International, **221**, no. 1, 692–705, doi: 10.1093/gji/ggaa016.
Moreau L., L. Stehly, P. Boué, et al, 2017, Improving ambient noise correlation functions with an SVD-based Wiener filter: Geophysical Journal International: **211**, no. 1, 418–426, doi: 10.1093/gji/ggx306.
Nakata N., R. Snieder, T. Tsuji, et al, 2011, Shear wave imaging from traffic noise using seismic interferometry by crosscoherence: Geophysics, **76**, no. 6, SA97–SA106, doi: 10.1190/geo2010-0188.1.
Nakata, N., J. P. Chang, J. F. Lawrence, et al, 2015, Body wave extraction and tomography at Long Beach, California, with ambient-noise interferometry: Journal of Geophysical Research: Solid Earth, **120**, no. 2, 1159–1173, doi: 10.1002/2015JB011870.
Ning, L., T. Dai, Y. Liu, et al, 2021, Application of multichannel analysis of passive surface waves method for fault investigation: Journal of Applied Geophysics, **192**, 104382, doi: 10.1016/j.jappgeo.2021.104382.
Olugboji, T., and Xue, S. 2022. A short-period surface-wave dispersion datasets for model assessment of Africa's crust: ADAMA: Seismological Society of America, **93**, 1943–1959, doi: 10.1785/0220210355.
Pan, S. L., K. Chen, J. Y. Chen, et al. 2020, A partial convolution-based deep-learning network for seismic data regularization: Computers & Geosciences, **145**, 104609, doi: 10.1016/j.cageo.2020.104609.
Rebuffi, S., A. Kolesnikov, G. Sperl, et al, 2017, iCaRL: Incremental Classifier and Representation Learning: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2001–2010.
Rickett, J., J. F. Claerbout, 1999, Acoustic daylight imaging via spectral factorization: Helioseismology and reservoir monitoring: The Leading Edge, **18**, no. 8, 957–960, doi: 10.1190/1.1438420.
Rost S., and C. Thomas, 2002, Array seismology: Methods and applications: Review of Geophysics, **40**, no. 3, 1008, doi: 10.1029/2000RG000100.
Schuster, G. T., J. Rickett, 2001, Daylight imaging in V(x,y,z) media: Stanford Exploration Project and SEP Report, **105**, 209–227.
Sun, H. Y., and L. Demane, 2023, Beyond correlations: Deep learning for seismic interferometry: IEEE Transactions on Neural Networks and Learning Systems, **34**, 3385–3396, doi: 10.1109/TNNLS.2022.3172385.
Torres, K., and M. D. Sacchi, 2023, Deep decomposition learning for reflectivity inversion: Geophysical Prospecting, **71**, 963–982, doi: 10.1111/1365-2478.13366.
Vander Maaten, L., and G. Hinton, 2008, Visualizing data using t-SNE: Journal of Machine Learning Research, **9**, 2008.
Vaswani, A., N. Shazeer, and N. Parmar, 2017, Attention is all you need: Advances in Neural Information Processing Systems, **30**, 1–7.
Wapenaar, K., D. Draganov, and J. O. Robertsson, 2008, Seismic interferometry: History and present status: Society of Exploration Geophysicists, **26**, doi: 10.1190/1.9781560801924.
Wegler, U., C. Sens-Schönfelder, 2007, Fault zone monitoring with passive image interferometry. Geophysical Journal International: **168**, no. 3, 1029–1033, doi: 10.1111/j.1365-246X.2006.03284.x.
Wu Y. H., S. L. Pan, J. Y. Chen et al, 2022, A surface-wave inversion method based on FHLV loss function in LSTM: IEEE Geoscience and Remote Sensing Letters, **19**, 1–5, doi: 10.1109/LGRS.2022.3187020.
Wu. Y. H., S. L. Pan, S. B. Yi, et al, 2023, An Automatic Screening Method for the Passive Surface-Wave Imaging Based on the F-K Domain Energy Characteristics: IEEE Transactions on Geoscience and Remote Sensing, **61**, 1–12, doi: 10.1109/TGRS.2023.3321786.
Wu, Y. H., S. L. Lan, H. Q. Lan, et al. 2025, Dynamic reservoir monitoring using similarity analysis of passive source time-lapse seismic images: Application to waterflooding front monitoring in Shengli Oilfield, China. Petroleum Science, **22**, 1062–1079, doi: 10.1016/j.petsci.2024.12.008.
Xu, Z., C. Juhlin, O. Gudmundsson, et al, 2012, Reconstruction of subsurface structure from ambient seismic noise: an example from Ketzin, Germany: Geophysical Journal International, **189**, no. 2, 1085–1102, doi: 10.1111/j.1365-246X.2012.05411.x.
Zhang, F. J., Z. Xu, C. Juhlin, et al, 2022, Passive seismic interferometry imaging. An example from the Ketzin pilot CO2 geological storage site: Journal of Applied Geophysics, **205**, 104765, doi: 10.1016/j.jappgeo.2022.104765.
Zhang, P., L. G. Han, Y. C. Yin, et al, 2019, Passive seismic full waveform inversion using reconstructed body-waves for subsurface velocity construction: Exploration Geophysics, **50**, 124–135, doi: 10.1080/08123985.2019.1575644.
Zhang, W., and J. H. Gao, 2021, Deep-learning full-waveform inversion using seismic migration images: IEEE Transactions on Geoscience and Remote Sensing, **60**, 1–18, doi: 10.1109/TGRS.2021.3062688.
Zhang, W., J. H. Gao, T. Yang, et al, 2021, Least-squares reverse time migration using convolutional neural networks: Geophysics, **86**, no. 6, R959–R971, doi: 10.1190/geo2021-0006.1.
Zhang, X. T., Z. Jia, Z. E. Ross, et al, 2020, Extracting dispersion curves from ambient noise correlations using deep learning: IEEE Transactions on Geoscience and Remote Sensing, **58**, 8932–8939, doi: 10.1109/TGRS.2020.2992043.
Zhao, B. H., L. G. Han, P. Zhang, et al, 2024, Noise Reduction and Encrypted Reconstruction of Passive Source Virtual Shot Records Based on GMF-RS Network: IEEE Transactions on Geoscience and Remote Sensing, **62**, 1–10, doi: 10.1109/TGRS.2024.3357724.

Biographies and photographs of the authors are not available.