Robust prestack seismic facies analysis using shearlet transform-based deep learning

Pu Liu¹,²,³,*, Rui Li¹, Yuehua Yue⁴, Songjie Liao⁴ and Feng Qian⁴

¹ Chengdu University of Technology, Chengdu, 610059, Sichuan, China
² Mountain Geophysical Technology Test Center, China National Petroleum Corporation, Chengdu, 611731, Sichuan, China
³ Bureau of Geophysical Prospecting Inc, China National Petroleum Corporation, Zhuozhou, 072751, Hebei, China
⁴ School of Information and Communication Engineering, Center for Information Geoscience, University of Electronic Science and Technology of China, Chengdu, 611731, Sichuan, China

*Corresponding author: Pu Liu. E-mail: lpislow123@163.com

Received 13 November 2021, revised 20 February 2022
Accepted for publication 8 March 2022

Abstract

One of the primary purposes of seismic stratigraphy is to evaluate the components of seismic layer relationships within a depositional chronology. Prestack seismic images contain a wealth of information, such as variations in the offset and azimuth of a seismic event, and naturally produce higher-resolution seismic facies analysis results than poststack data. However, prestack data usually suffer from potential unreliability issues due to low signal-to-noise ratios. As this is often overlooked, the present facies analysis methods sometimes fail to extract accurate features from prestack images, which inevitably influences the facies analysis results. To address this issue, this article provides a robust data-driven technique for extracting offset-temporal features via shearlet transform-based deep convolution autoencoders (STCAEs). Unlike the present facies analysis in the time domain, STCAE can optimally represent prestack images at multiple scales and directions through the two-dimensional shearlet transform, which preserves fine edges while suppressing noise in prestack images. Subsequently, robust features are extracted from prestack images in a data-driven manner through a contractive convolutional autoencoder network. We compare our method with other advanced methods and demonstrate the advantages of the proposed approach in classifying seismic layers in prestack data.

Keywords: Seismic facies classification, robust, deep convolution autoencoders, shearlet transform

1 Introduction

Seismic phase classification is the interpretation of seismic facies in terms of sedimentary environments and facies distributions (Dumay 1988; Coléou et al. 2003). With the advent of 3D seismic technology and the consequent expansion of survey areas, the manual approach for seismic facies interpretation has become extremely time-intensive. However, seismic surface analysis is a machine-learning method that automatically constructs 2D seismic facies. Like most machine-learning problems, determining the appropriate representations for seismic data is the fundamental approach to automatic seismic facies analysis for seismic facies classification (SFC) (Zhao et al. 2015).

The current seismic literature builds upon two significant model-driven and data-driven techniques to achieve SFC. Model-driven seismic facies analysis usually extracts main features through a handcrafted design, which generally applies two methods, namely, poststack (Saraswat & Sen...
Based deep convolution autoencoder, or STCAE, method). For poststack classification methods, model-driven seismic facies analysis identifies diverse waveform signal patterns with time-series transforms, and the features can be divided into three main categories: time-domain (Dumay 1988; Saraswat & Sen 2012), frequency-domain (Dumay 1988; Xie et al. 2017), and time-frequency domain features (Matos et al. 2007; Du et al. 2015; Amendola et al. 2017). In contrast, deep learning (DL) differs from the above model-based feature extraction methods in that it is a fusion of data-driven feature extraction and classification into one, with typical methods being supervised (Zhao 2018; Li et al. 2020; Puzyrev & Elders 2020; Shi & Fomel 2020; Feng et al. 2021; Zhang et al. 2021), semi-supervised (Liu et al. 2020), and unsupervised DL (Duan et al. 2019). It is considered that prestack seismic images contain more information than poststack seismic waveforms; therefore, prestack seismic data classification results are more convincing than poststack seismic data using image processing-based feature extraction methods, such as prestack-texture (Song et al. 2016) and Gabor transform (Song et al. 2015) methods. However, model-driven methods fail to adapt to the polytrope of individual images, which inevitably affects the extracted features. As an alternative, a data-driven strategy is introduced that utilizes deep learning (e.g., autoencoders; Qian et al. 2018; Shi & Fomel 2020; Puzyrev & Elders 2020; Zhu et al. 2021) to learn feature representations from prestack images. However, the existing data-driven approaches originated for common image processing purposes. It is well known that prestack seismic images have more noise and are more unreliable than regular images (Gao et al. 2013).

To overcome the deficiencies mentioned above, we introduce a powerful data-driven offset-time feature extraction method, which relies on the combination of a shearlet transform (Wang 2010; Liu et al. 2014) and deep convolution autoencoders (Masci et al. 2011) (called the shearlet transform-based deep convolution autoencoder, or STCAE, method). The advantages of the STCAE-based approach are as follows. (1) Robust representation: when used on prestack seismic images, the shearlet transform preserves as much detail as possible while transforming images without information loss. Additionally, the shearlet transform’s superior frequency characteristics provide an ideal balance of denoising performance and detail retention. (2) Multiscale information: the shearlet transform is a multiscale analytic technique with improved directional sensitivity. It combines multiscale decomposition with prestack seismic images to capture anisotropic characteristics efficiently. (3) Data-driven approach: the STCAE algorithm attempts to extract characteristics from prestack images in the shearlet domain. The polytrope of individual shearlet coefficient pictures can be better adapted by a data-driven approach, thus compensating for model-specified feature inadequacies.

Relying on the advantages listed above, here we offer a robust strategy for data-driven feature extraction through the combination of a shearlet transform and deep convolution autoencoders. Specifically, this article makes the following contributions.

1. We propose a novel SFC workflow for inaccurate prestack seismic images. Specifically, unlike traditional deep convolutional autoencoders (DCAEs), we first introduce the shearlet transform to represent a prestack image sparsely and then automatically learn its features using STCAE to obtain noise-robust features.

2. To attenuate the image noise in the shearlet domain, we first extract the texture feature from shearlet coefficients to highlight valuable feature information from noisy shearlet images. Then, we introduce penalties sensitive to feature changes that form a STCAE network within the DCAE framework to learn features from ST-based texture images automatically.

3. We validate the STCAE algorithm for automatic feature extraction and demonstrate its efficacy in terms of extraction quality using data from a synthetic dataset and actual surveys.

The structure of the rest of the paper is as follows. Section 2 discusses the proposed model. We discuss the validity of the model using theoretical data and data from a real-world work zone in Section 3. Section 4 provides a summary of this study.

2 Proposed method

2.1 Overview

As mentioned before, typical facies analysis methods based on dimensionality reduction suffer from a fatal flaw: the extracted features are susceptible to noise, especially severe noise, which potentially affects the accuracy of seismic facies analysis. Hence, this section presents the proposed STCAE model to circumvent this problem, and the model workflow consists of the following four stages.

1. Data preparation: as a prestack image, select prestack seismic data from a suitable time window on a target horizon. Then, input the prepared prestack image into the shearlet transform.

2. ST-based texture feature extraction: take the shearlet transform of all prestack images and extract the texture feature from shearlet coefficients.

3. ST-based texture feature compression: train the STCAE network on the texture feature images; after training,
feed the network the same texture feature images to extract the deep features.

(4) Self-organizing feature map (SOM)-based unsupervised pattern recognition: to create the final facies map, cluster the deep features together using an unsupervised pattern recognition algorithm.

### 2.2 Weighted ST-based texture feature extraction

Hence, we postulate that prestack images may be represented as $S_n$, where $n = 1, \ldots, N$ in the space $\mathbb{R}^{T \times R}$ and can be used as input variables. Here, $T$ denotes the number of time samples, $N$ denotes the total number of common-depth points (CDPs), and $R$ represents the number of offset sections. This stage automates the extraction of the texture feature $F \in \mathbb{R}^{T \times R}$ from ST coefficients $Y \in \mathbb{R}^{T \times R}$ that best represent $S_n$, where $L$ is the number of scales.

Thus, it is pronounced that this step consists of the following two processes.

#### 2.2.1 Discrete ST (DST)

Formally, for a given prestack image $S$ subject to the following continuous ST is given by (Wang 2010)

$$ Y(a, s, l) = \langle S, \psi_{a,s,l} \rangle, \ l \in [L]. \ (1) $$

Here,

$$ \psi_{a,s,l}(x) = a^{-3/4} \psi \left( A^{-1} B^{-1} (x - t) \right), \ (2) $$

where $\psi_{a,s,l}(x)$ is the shearlet function, $t \in \mathbb{R}^2$ is the translation parameter, and $A$ and $B$ are an anisotropic expansive matrix and a shear matrix, respectively, and are expressed as follows:

$$ A = \begin{pmatrix} a & 0 \\ 0 & \sqrt{a} \end{pmatrix}, \quad B = \begin{pmatrix} 1 & s \\ 0 & 1 \end{pmatrix}, \ (3) $$

where $a \in \mathbb{R}^+$ denotes the scaling factor and $s \in \mathbb{R}$ denotes the shear factor.

#### 2.2.2 Weighted ST-based texture feature extraction

For the 3D ST coefficients that have been solved, we propose using a weight fusion approach to extract the potential robust texture features implicit in the ST coefficients. The average $l_1$ norms of the ST coefficient subbands are used as the ST texture features of the prestack images, as described below:

$$ e = \frac{1}{TR} \sum_{i=1}^{T} \sum_{j=1}^{R} |Y(i, j, l)|, \ l \in [L]. \ (4) $$

The energy distribution of the transform domain can be seen in the coefficients of the subbands shown in figure 1. The direct clustering of the subbands shield the small-scale local features of the ST features from the large-scale global features, which is not conducive to discovering the details of the original prestack image. Therefore, it is necessary to weight the energy of the ST features at each scale.

Then, to balance the contribution of the features at each scale for cluster analysis, the eigenvectors are weighted based on the energy ratio of the transform coefficients for different scales. First, the coefficient energy for each scale is determined as follows:

$$ e_j = \sum_{i=1}^{n_i} e_{jp} \ (5) $$

where $n_i$ is the number of directions at scale $j$ and $e_{ij}$ is the subband energy at scale $j$ and in direction $i$. Thus, the following equation can be used to compute the weight coefficient:

$$ w_{ij} = \begin{cases} 1, j = 1 \\ e_{ij}, j \neq 1. \end{cases} \ (6) $$

Then, combined with the weight coefficient and the texture measure calculated using equation (4), a normalized eigenvector can be obtained as follows:

$$ \tilde{e}_{ij} = \psi_j e_{ij}. \ (7) $$

### 2.3 ST feature compression based on the STCAE network

To remove residual noise and reduce dimensionality, this section proposes using a convolutional autoencoder (CAE) network to further compress ST texture features. In addition, a penalty term sensitive to feature changes is introduced to form the STCAE network and extract prestack ST features robustly. The structure of the STCAE network is depicted in figure 2. During the encoding stage, a convolution operation is performed on the input array of multiscale coefficients $X_i \in \mathbb{R}^{C \times T \times R}$ and the convolution kernel of the $i$th layer, and $Y_E^i \in \mathbb{R}^{C \times C \times T \times R}$ ($C$ denotes the number of channels in the input data, $n$ denotes the size of the 2D convolution kernel, and the number of convolution kernels is denoted by $k$; the $i$th layer outputs $k$ feature maps). Then, the $k$th feature of $X_i$ in the input data is calculated as follows:

$$ \Phi_k(X_i) = \sigma \left[ \Psi \left( Y_E^i \otimes X_i + b_r^i \right) \right], \ (8) $$

where the function $\Psi(\cdots)$ represents a maximum pooling operation and $b_r^i$ is a $k$-dimensional vector, each element of which depends on the corresponding convolution kernel.

Through calculations in $l$ hidden layers, an encoding result $h$ is obtained, and it can be expressed with the following equation:

$$ Y_l = \Phi \left( Y_{l-1} \otimes \tilde{Y}_{l-1} + b_{r}^l \right). \ (9) $$

The output of the $i$th hidden layer for decoding with regard to the received input $Y_{l-1}$ is as follows during the decoding stage:

$$ Y_l = \tilde{Y} \left( Y_{l-1} \otimes \tilde{Y}_{l-1} + b_{r}^l \right), \ (10) $$

Downloaded from https://academic.oup.com/jge/article/19/3/521/6617874 by guest on 05 July 2022
where symbol $\otimes$ denotes a deconvolution operation, $Y_i$ is the $i$th feature map, $W_d^i$ is the deconvolution kernel, and $\tilde{f}(\cdots)$ represents an unpooling operation. Generally, pooling operations are irreversible. In this study, during the reconstruction process, we randomly place the calculated value at a certain location in the unpooling layer while filling the other locations in the layer with zero to produce a relatively reasonable approximation.

The convolution operations in the $I$th hidden layers produce the sparsely encoded reconstructed data $\hat{\chi}_i$. The whole decoding process can be expressed using the following equation:

$$\hat{\chi}_i = g(y_i) = D(y_i; W_d^1, W_d^2, \ldots, W_d^D).$$

(11)

Assume that the following is the definition of the loss function:

$$J(\theta) = \sum_{X \in D} \mathcal{L}(X, g[f(X)])$$

(12)

where $\mathcal{L}(\cdots)$ is the reconstruction error and $\theta = \{W_\phi, W_d^i, b_\phi, b_d^i\}, i \in [I]$ is the parameter set of the CAE network. Generally, the minimum mean square error is selected as follows:

$$J(\theta) = \frac{1}{N} \sum_{i \in D} |\chi_i - \hat{\chi}_i|^2_F.$$  

(13)

The loss function is then modified to include a penalty term to construct the STCAE network:

$$\left|J_f(x)\right|^2_F = \sum_{i \in [d_x], j \in [d_y]} \left(\frac{\partial f_j(x)}{\partial x_i}\right)^2.$$  

(14)

By including the penalty term in equation (14) in the loss function, the final STCAE network's objective function becomes the following:

$$J_{STCAE}(\theta) = \sum_{X \in D} \left(\mathcal{L}(X, g[f(X)]) + \lambda J_f(x)\right)$$

(15)

Training is posed as an optimization problem that is performed layer by layer. Prior to training, random values are used to initialize the network parameters. The texture features of the prestack ST coefficients are input to the automatic derivative function in TensorFlow one-by-one. The parameters of each layer are updated in the direction that minimizes the loss function.
Once the whole sample set is trained, the final network parameters $\theta$ are generated, which can then be used to extract the hidden-layer features from the sample data:

$$F_i = E \left( X_i; \mathcal{W}^1, b^1, \ldots, \mathcal{W}^I, b^I \right),$$

where $F_i$ denotes the features of sample $i$.

### 2.4 Unsupervised pattern recognition

Feature vectors are generated from the prestack seismic images and by clustering the feature vector by the STCAE network, enabling the next stage of the facies map. There are numerous unsupervised clustering approaches, such as SOM (Su & Chang 2000) and fuzzy c-means clustering (Wang et al. 2006), that can be used. Here, we use an SOM technique to recognize seismic facies without sacrificing generality. An SOM approach randomly initializes $M$ model vectors $p_i \in \{ p_1, p_2, \ldots, p_M \}$ and calculates the distance between each input sample $x_i \in \{ x_1, x_2, \ldots, x_N \}$ and each model vector, which is defined as follows:

$$i_a = \min_i |x_a - p_i|^2.$$

To determine the $i_a$th neuron that is the closest to the sample, the SOM algorithm updates the $i_a$th neuron and its nearby
neuron model vector, as described below:

\[ p_j(t+1) = p_j(t) + \alpha(t) h_{ia} [x_k - p_{ia}(t)] , \]

\[ h_{ia} = \exp \left\{ -|r_u - r_i| \right\}, \tag{18} \]

where \( \alpha(t) \) denotes the learning rate, which falls monotonically as the number of training iterations grows, \( h_{ia} \) denotes the neighbourhood of the activated neuron, and \( r_u \) and \( r_i \) are the locations of the updated neuron and the activated neuron in the lower-dimensional space, respectively.

\[ p_j(t+1) = p_j(t) + \alpha(t) h_{ia} [x_k - p_{ia}(t)]. \]

\[ h_{ia} = \exp \left\{ -|r_u - r_i| \right\}, \tag{18} \]

where \( \alpha(t) \) denotes the learning rate, which falls monotonically as the number of training iterations grows, \( h_{ia} \) denotes the neighbourhood of the activated neuron, and \( r_u \) and \( r_i \) are the locations of the updated neuron and the activated neuron in the lower-dimensional space, respectively.

2.5 Performance evaluation

2.5.1 Number of reflection pattern types. The Davies–Bouldin index (DBI) (Kärkkäinen & Fränti 2000) is chosen for the quantitative assessment and as the clustering algorithm evaluation measure to evaluate the goodness of split by a SOM algorithm for a given number of clusters. The DBI is essentially a data-driven metric used to determine the number of clusters, and it is calculated as follows:

\[ DBI = \frac{1}{N} \sum_{i=1}^{N} \max_{i \neq j} \left( \frac{\bar{c}_i - \bar{c}_j}{d_{ij}} \right), \tag{19} \]

where \( \bar{c}_i \) and \( \bar{c}_j \) are the average distances between the data points assigned to clusters \( I \) or \( J \) and the corresponding centroids, respectively, and \( d_{ij} \) represents the distance between two cluster centres. That is, the DBI is calculated by comparing each cluster’s average similarity to the cluster that is most similar to it. The lower the average similarity is, the more distinct the clusters are and the more accurate the clustering result is.

2.5.2 Uncertainty evaluation of the clustering results. The distinction index \( D \) is generally used to evaluate the uncertainty of the SFC results quantitatively as follows:

\[ \text{DBI} = \frac{1}{N} \sum_{i=1}^{N} \max_{i \neq j} \left( \frac{\bar{c}_i - \bar{c}_j}{d_{ij}} \right), \tag{19} \]

where \( \bar{c}_i \) and \( \bar{c}_j \) are the average distances between the data points assigned to clusters \( I \) or \( J \) and the corresponding centroids, respectively, and \( d_{ij} \) represents the distance between two cluster centres. That is, the DBI is calculated by comparing each cluster’s average similarity to the cluster that is most similar to it. The lower the average similarity is, the more distinct the clusters are and the more accurate the clustering result is.
Figure 6. The $D$ value for each location. (a) Based on the poststack data. (b) Based on the prestack ST features. (c) Based on the weighted prestack ST features.

(Jiang et al. 2011): 

$$D = 1 - \frac{|x - c_1|_2}{|x - c_2|_2},$$  \hspace{1cm} (20)

where $x$ is the cluster sample, $c_1$ and $c_2$ are the centroids of the clusters that are the closest and second closest to the sample, respectively, and $|\cdot|_2$ is the $l_2$ norm of the vector. Note that the distinction index can reflect the reasonableness of the clustering results. If the distinction index is small, it means that the sample is closer to the centre of mass of both clusters; then, it is more difficult to determine which cluster it belongs to. In contrast, the higher the distinction index, the more reliable the clustering result. For real-world data, the statistical percentage of $D$ usually also needs to be calculated to evaluate the recognition results quantitatively.

$$P(d) = \frac{n_d}{n} \cdot 100\%.$$  \hspace{1cm} (21)

Here, the number of the full sample is denoted by $n$, and $n_d$ denotes the total number of samples classified as having $D = d$.

3 Model verification

In this part, we demonstrate the performance of our STCAE network using synthetic and field data sets. Section 3.1 begins by describing the data sets used throughout the review process. Then, using both theoretical data (see Section 3.2) and data from a real-world work zone, we compare...
Figure 7. Classification results of the prestack data with SNR = 10 dB. (a) Prestack for point A in the model with added noise. (b) CAE-SOM-based classification result. (c) STCAE-SOM-based classification result.

our proposed strategy to the compared methodologies (see Section 3.3).

3.1 Data

To test the effectiveness of the proposed scheme, we used synthetic and field data sets to evaluate our STCAE model. The following provides an overview and the specifics of the artificial data and real seismic data sets.

- Synthetic data: the synthetic example is composed of four medium layers, as shown in figure 4. In this example, the velocities in layers 1, 3, and 4 were constant, and the velocity in layer 2 varied in the transverse direction.
- Field data: the prestack seismic data set covers the Liziba (LZB) region of southern China’s Sichuan Basin and consists of 950 inlines by 550 crosslines, totaling over 50000 CDPs recorded at 2 ms intervals.

3.2 Network architecture

The convolutional layers, which give the network its name, are at the heart of the STCAE-based model. In both the encoder and decoder, this model contains 18 convolutional layers (kernel size $3 \times 3$), each using the ReLU activation function. In the encoder and decoder, respectively, a new max-pooling and concatenate layer with dimensions of $2 \times 2$ is introduced after the final convolutional layer. The Adam optimization approach is used to train the DCAE-based model for 20 epochs.

3.3 Validation on synthetic data

For the synthetic data, we analyse the seismic facies of the bottom boundary in layer 2 to evaluate the effectiveness of the proposed ST features in characterizing the features of the prestack seismic data. The reflection coefficient of each self-reflection horizon was obtained by the Aki–Richards approximation equation. In addition, we produce prestack seismic datasets and poststack seismic datasets by using the convolutional reflection coefficient of a 30 Hz Ricker wavelet. To illustrate the robustness of the features, we add random noise with a signal-to-noise ratio (SNR) of 12 dB to the synthetic data. Finally, we use the SOM algorithm to classify the recognized features, as shown in figure 5.

The clustering results of the poststack data, the prestack ST features, and the weighted prestack ST features are revealed in figures 5a–c, respectively. To compare the seismic facies maps of the three feature extraction methods, figure 5a shows that the method based on the poststack data is weak in distinguishing different categories and obtains the wrong classification for the second and third categories. The clustering results of the latter two methods show clear boundaries of various seismic facies (see figures 5b–c). Through this comparison, we demonstrate the advantages of prestack data pattern analysis.

For an even better demonstration, we use the $D$ values to evaluate the results of seismic facies classification. Figures 6a–c show the $D$ value for each location of the poststack data, the prestack ST features, and the weighted prestack ST features, respectively. Although the methods based on prestack data features have high clustering discrimination, the method based on weighted ST features has a higher discrimination when the rock properties of the two seismic facies are similar. This leads us to the conclusion that the method based on weighted ST features retains more data difference features and is more conducive to clustering field data with uneven distributions of feature differences.

To illustrate the robustness of our STCAE model in extracting weighted prestack ST features, we add noise to the synthetic data generated by the geological model to obtain
noisy prestack data with SNRs of 10 dB, 8 dB, 4 dB, and 2 dB. Then, we use the CAE model and the STCAE model to reduce the dimensionality of the weighted ST feature by the CAE network and by the STCAE network, respectively, and use the SOM algorithm for clustering to obtain the seismic facies map (see figures 7–10). This comparison shows that our method has excellent performance in terms of noise resistance.

Similarly, we calculate the $D$ values for the clusters of the features extracted by the CAE and STCAE networks. Figures 11a–d shows the discriminations of the clustering results of the prestack data with different levels of noise by the

Figure 8. Classification results of the prestack data with SNR = 8 dB. (a) Prestack for point A in the model with added noise. (b) CAE-SOM-based classification result. (c) STCAE-SOM-based classification result.

Figure 9. Classification results of the prestack data with SNR = 4 dB. (a) Prestack for point A in the model with added noise. (b) CAE-SOM-based classification result. (c) STCAE-SOM-based classification result.

Figure 10. Classification results of the prestack data with SNR = 2 dB. (a) Prestack for point A in the model with added noise. (b) CAE-SOM-based classification result. (c) STCAE-SOM-based classification result.
CAE network. Figure 11e–h shows the $D$ values corresponding to the STCAE network. For the second and third seismic facies with small lateral differences in waveform amplitude, compared to the CAE network, our proposal extracted more distinguishing features. This result shows that, with the reduction in the SNR of the prestack data, the discrimination degree based on the STCAE network is significantly better than that of the CAE network.

### 3.4 Validation of field data

We further demonstrate the superiority of the proposed STCAE method for seismic facies analysis using 4-D prestack field data. This article chooses a horizon in the LZB prestack data that includes faults, and the amplitude attribute of this horizon is displayed in figure 12. We extract data from 24 ms time windows above and below the target horizon and then use the CAE network and STCAE network to extract the compressed representation of the ST weighted features of the prestack data. For the CAE and STCAE networks, the feature learning rate was set to 0.002. We use $13 \times 3 \times 3 \times 10$ convolutional layers to obtain 10 layers of feature maps and subsequently vectorize them to generate their eigenvectors. The number of clustering centres of the SOM algorithm was set to 10, and the clustering analysis of the poststack and prestack data characteristics is carried out, forming seismic facies classification maps (see figures 13–15). It is worth noting that figure 3 shows the deep features extracted from the prestack images in Shearlet domain by the proposed STCAE, which inputs into the SOM to implement facies clustering.

The clusters of the poststack data produced by the SOM algorithm are illustrated in figure 13. The red circles show that the fault identification is not clear. In the red circle of figures 14 and 15, the classification results of both the CAE-SOM and STCAE-SOM methods allow for the clear presentation of large fault structures in seismic data. However, the STCAE-SOM network has some advantages over the CAE-SOM network in identifying more minor faults in geological interpretation. As circled in figures 13 and 14, there is a distribution of faults, which are vaguely identified in the
Figure 12. The amplitude attribute of the LZB prestack seismic data.

Figure 13. SOM-based classification map of the poststack seismic data.

Figure 14. CAE-SOM-based classification map of the prestack seismic data.

results based on the CAE-SOM method, which does not respond well to faults. In contrast, the STCAE-SOM method can identify the presence of faults. This comparison shows that our proposed model performs better than the other two models in seismic interpretation.

To explain the numerical advantages of the features extracted by the STCAE network for improving the clustering effect, we calculate the discrimination of the clustering results obtained by the three methods (see figure 16). Figure 16a illustrates that the $D$ value of the SOM algorithm
Figure 15. STCAE-SOM-based classification map of the prestack seismic data.

Figure 16. The $D$ values for the clusters. (a) The $D$ values for the clusters of the poststack data produced by the SOM algorithm. (b) The $D$ values for the clusters of the prestack seismic data produced by the CAE-SOM method. (c) The $D$ values for the clusters of the prestack seismic data produced by the STCAE-SOM method.

Figure 17. Statistical distribution of the $D$ values for the classification results for the LZB prestack data.

is relatively low, indicating that the reliability of the classification results is low. The $D$ values corresponding to the CAE-SOM and STCAE-SOM methods are concentrated in a high-value range and a low-value range, as shown in figures 16b and c. As mentioned earlier, the higher the distinction index, the more reliable the result. Figure 17 depicts the statistical distribution of the $D$ values for the classification results on the LZB prestack data. The results of STCAE-SOM are concentrated on the high-value segment, while those of CAE-SOM are concentrated on the low-value segment. This result indicates that the STCAE-SOM method is more reliable than the CAE-SOM method. This test demonstrates that the STCAE-SOM approach numerically outperforms the other two methods.

4 Conclusion
This article provides a data-driven method that uses prestack seismic data to recognize seismic facies based on the STCAE model. The proposed STCAE network structure, which
combines weighted ST-based texture features and STCAE network techniques, can directly learn offset-temporal features from prestack seismic images. STCAE outperforms standard methods in classifying data more precisely, defines multiscale fractures more clearly, and shows seismic surface characteristics more accurately, as demonstrated by the experiments on synthetic and genuine seismic data.

Acknowledgements

The authors would like to thank the associate editor, Ximing Wu, and the three anonymous reviewers, whose constructive suggestions helped to greatly improve and clarify this article. This work was supported by the National Natural Science Foundation of China under Grant No. 41874155 and Grant No. 42130812.

References

Amendola, A., Gabrielelli, G., Dell’Aversana, P. & Marini, A.J., 2017. Seismic facies analysis through musical attributes, *Geophysical Prospecting*, **65**, 49–58.

Coléou, T., Poupon, M. & Azbel, K., 2003. Unsupervised seismic facies classification: A review and comparison of techniques and implementation, *The Leading Edge*, **22**, 942–953.

Du, H.K., Cao, J.X., Xue, Y.J. & Wang, X.J., 2015. Seismic facies analysis based on self-organizing map and empirical mode decomposition, *Journal of Applied Geophysics*, **112**, 52–61.

Duan, Y., Zheng, X., Hu, L. & Sun, L., 2019. Seismic facies analysis based on deep convolutional embedded clustering, *Geophysics*, **84**, IM87–IM97.

Dumay, J., 1988. Multivariate statistical analyses applied to seismic facies recognition, *Geophysics*, **53**, 1151–1159.

Feng, R., Balling, N., Grana, D., Dramsch, J.S. & Hansen, T.M., 2021. Bayesian convolutional neural networks for seismic facies classification, *IEEE Transactions on Geoscience and Remote Sensing*, **59**, 8933–8940.

Gao, J., Sacchi, M.D. & Chen, X., 2013. A fast reduced-rank interpolation method for prestack seismic volumes that depend on four spatial dimensions, *Geophysics*, **78**, V21–V30.

Jiang, B., Pei, J., Tao, Y.F. & Lin, X.M., 2011. Clustering uncertain data based on probability distribution similarity, *IEEE Transactions on Knowledge and Data Engineering*, **25**, 751–763.

Kärkkäinen, I. & Franti, P., 2000. Minimization of the value of Davies-Bouldin index, in *Proceedings of the LASTED International Conference on Signal Processing and Communications*, Malaga, 2000, ACTA Press, Calgary, Canada, P426–P435.

Li, F., Zhou, H., Wang, Z. & Wu, X., 2020. ADDCNN: An attention-based deep dilated convolutional neural network for seismic facies analysis with interpretable spatial-spectral maps, *IEEE Transactions on Geoscience and Remote Sensing*, **59**, 1733–1744.

Liu, C., Wang, D., Wang, T., Feng, F. & Meng, G., 2014. Random seismic noise attenuation based on the Shearlet transform, *Acta Petrolei Sinica*, **35**, 692–699.

Liu, M.L., Jervis, M., Li, W.C. & Ninlet, P., 2020. Seismic facies classification using supervised convolutional neural networks and semisupervised generative adversarial networks, *Geophysics*, **85**, O47–O58.

Masci, J., Meier, U., Ciresan, D. & Schmidhuber, J., 2011. Stacked convolutional auto-encoders for hierarchical feature extraction, in *Proceedings 21st International Conference on Artificial Neural Networks*, Espoo, 2011, Springer, Berlin, Germany, p.52.

Matos, M.D., Osorio, P.L. & Johann, P.R., 2007. Unsupervised seismic facies analysis using wavelet transform and self-organizing maps, *Geophysics*, **72**, P9–P21.

Molino-Minero-Re, E., Rubio-Acosta, E., Benítez-Pérez, H., Brandi-Purata, J.M., Perez-Quezadas, N.J. & García-Nocetti, D.F., 2018. A method for classifying pre-stack seismic data based on amplitude-frequency attributes and self-organizing maps, *Geophysical Prospecting*, **66**, 673–687.

Puzyr, V. & Elders, C., 2020. Deep convolutional autoencoder for unsupervised seismic facies classification, in EAGE/AAPG Digital Subsurface for Asia Pacific Conference, Kuala Lumpur, 2020, European Association of Geoscientists & Engineers, Houten, the Netherlands, p.1.

Qian, F., Miao, Y., Liu, X.Y., Wang, Y.J., Cai, L. & Hu, G.M., 2018. Unsupervised seismic facies analysis via deep convolutional autoencoders, *Geophysics*, **83**, 1–19.

Saraswat, P. & Sen, M.K., 2012. Artificial immune-based self-organizing maps for seismic-facies analysis, *Geophysics*, **77**, O45–O53.

Song, C., Li, L., Li, L. & Li, K., 2021. Robust K-means algorithm with weighted window for seismic facies analysis, *Journal of Geophysics and Engineering*, **18**, 618–626.

Song, C., Li, Z., Liu, Z. & Hu, G., 2015. Prestack reflection pattern based seismic facies analysis, *SEG Technical Program Expanded Abstracts*, **2015**, 1633–1637.

Song, C., Liu, Z., Cai, H., Wang, Y., Li, X. & Hu, G., 2017a. Unsupervised seismic facies analysis with spatial constraints using regularized fuzzy c-means, *Journal of Geophysics and Engineering*, **14**, 1535–1543.

Song, C.Y., Liu, Z.L., Wang, Y.J., Li, X.M. & Hu, G.M., 2017b. Multi-waveform classification for seismic facies analysis, *Computers & Geosciences*, **101**, 1–9.

Song, C.Y., Liu, Z.N., Cai, H.P., Qian, F. & Hu, G.M., 2016. Pre-stack-texture-based reservoir characteristics and seismic facies analysis, *Applied Geophysics*, **13**, 69–79.

Shi, Y., Wu, X. & Fomel, S., 2020. Waveform embedding: automatic horizon picking with unsupervised deep learning, *Geophysics*, **85**, WA67–WA76.

Su, M.C. & Chang, H.T., 2000. Fast self-organizing feature map algorithm, *IEEE Transactions on Neural Networks*, **11**, 721–733.

Wang, Q.L., 2010. The discrete shearlet transform: A new directional transform and compactly supported shearlet frames, *IEEE Transactions on Image Processing*, **19**, 1166–1180.

Wang, W., Zhang, Y., Yi, L. & Zhang, X., 2006. The global fuzzy c-means clustering algorithm, in 6th World Congress on Intelligent Control and Automation, Dalian, 2006, IEEE, New York, NY, USA, p. 3604.

Wei, Y. & Zhang, Q., 2012. *Common Waveform Analysis: A new and practical generalization of Fourier analysis*, Springer, p. 9.

Xie, T., Zheng, X. & Zhang, Y., 2017. Seismic facies analysis based on speech recognition feature parameters, *Geophysics*, **82**, O23–O35.

Zhang, H.R., Chen, T.S., Liu, Y., Zhang, Y.X. & Liu, J., 2021. Automatic seismic facies interpretation using supervised deep learning, *Geophysics*, **86**, IM15–IM33.

Zhao, T., 2018. Seismic facies classification using different deep convolutional neural networks, *SEG Technical Program Expanded Abstracts*, **2018**, 2046–2050.

Zhao, T., Jayaram, V., Roy, A. & Marfurt, K.J., 2015. A comparison of classification techniques for seismic facies recognition, *Interpretation*, **3**, SAE29–SAE58.

Zhu, D.L., Cui, J.B., Li, Y., Wan, Z.H. & Li, L., 2021. Adaptive Gaussian mixture model and convolution autoencoder clustering for unsupervised seismic waveform analysis, *Interpretation*, **10**, T181–T193.