Unsupervised Seismic Facies Deep Clustering Via Lognormal Mixture-Based Variational Autoencoder

Haowei Hua, Feng Qian, *Member, IEEE*, Gulan Zhang, *Member, IEEE*, and Yuehua Yue

***Abstract*—Seismic facies analysis (SFA) is a crucial step in the interpretation of subsurface structures, with the core challenge being the development of automatic approaches for the analysis of 4D prestack seismic data. The dominant isolated learning-based SFA schemes have gained considerable attention and primarily focus on learning the best representation of prestack data and generating facies maps by clustering the extracted features. However, in isolated learning, the independent nature of feature extraction and clustering leads to the ineffectiveness of clustering loss guidance on feature extraction, thereby resulting in derived features that unnecessarily facilitate the clustering task. As an alternative, we proposed a new unsupervised, end-to-end learning-based SFA method, which is referred to as the lognormal mixture-based variational autoencoder (LMVAE) and enhanced the existing Gaussian mixture variational autoencoder-based deep clustering framework (GMVAE framework). In this approach, both the extraction and clustering of seismic features are simultaneously performed by determining from which mode of the latent mixture distribution the seismic data were generated. Furthermore, the LMVAE extends the Gaussian mixture modeling of seismic features in the GMVAE framework to lognormal mixture modeling, improving the adaptability of SFA to field data. The effective performance of the LMVAE is demonstrated in synthetic and field prestack seismic data.**

***Index Terms*—Deep clustering, end-to-end learning, lognormal mixture-based variational autoencoder (LMVAE), Seismic facies analysis (SFA).**

I. Introduction

SEISMIC facies analysis (SFA) endeavors to perpetuate the interpretation of the depositional environment and facies distribution, being solely observed from both the seismic reflector and horizon information [1], [2], [3]. Traditionally, SFA has relied on the expertise and collaboration of multidisciplinary teams, including geophysicists, geologists, and petrophysicists, to manually interpret seismic data [4]. However, the exponential increase in the seismic data volume has rendered this process both difficult and time-consuming [5], [6]. Consequently, there is an urgent need to develop automatic approaches for SFA to improve the efficiency and accuracy.

The prevailing automatic SFA mainly focuses on two principal types of machine learning techniques, i.e., isolated learning-based SFA and end-to-end learning-based SFA. In isolated learning-based SFA, feature extraction and clustering are performed separately. For feature extraction, both model-driven methods and data-driven methods are employed to obtain a more insightful representation of seismic data. In model-driven methods, time-series transforms are generally implemented to provide nonredundant insights into the underlying prior properties of seismic data [7], [8], including time-domain [1], [9], [10], [11], [12], time-frequency-domain [13], [14], [15], [16], [17], and frequency-domain [1], [18], [19], [20]. In contrast, data-driven methods are aimed at automatically learning local patterns in seismic data without prior knowledge or assumptions, which builds upon the learning capability of the autoencoder [21], [22], [23], [24], [25], [26], or recurrent network [27], [28], [29]. Thereafter, isolated learning-based SFA achieves SFA via multifarious feature clustering algorithms, such as centroid-based clustering [30], [31], [32], [33], [34], [35], [36], probabilistic model clustering [37], [38], [39], and spectral clustering [11], [40], [41]. However, due to the isolation between feature extraction and clustering, the guidance of clustering loss is ineffective for feature extraction. Therefore, the SFA results of isolated learning may be suboptimal since the derived features are not necessarily conducive to the clustering.

To avoid the aforementioned suboptimal results, as an alternative, end-to-end learning-based SFA is performed to establish a straight pathway from seismic data to the facies map without feature appraisal. This approach can be categorized into supervised and unsupervised methods. In the case of supervised methods, the focus lies in directly learning the relationship between the input seismic samples and their corresponding seismic facies category directly from labeled data [42], [43]. Alternatively, employing leveraging image segmentation techniques to partition seismic data into distinct regions for the purpose of classifying seismic facies, where seismic samples belonging to the same region share the same seismic facies. Typical segmentation methods include the following: attention mechanism-based networks [44], [45], [46], U-net-like networks [47], [48], [49], [50], and segmentation with less data [51], [52]. These supervised methods offer a significant boost in the accuracy of facies

Manuscript received 1 August 2023; revised 26 September 2023 and 12 October 2023; accepted 17 October 2023. Date of publication 19 October 2023; date of current version 30 October 2023. This work was supported by the National Natural Science Foundation of China under Grant 41874155, Grant 42130812, and Grant 41874168. *(Corresponding authors: Feng Qian; Gulan Zhang.)*

Haowei Hua and Feng Qian are with the School of Information and Communication Engineering, Center for Information Geoscience, University of Electronic Science and Technology of China, Chengdu 611731, China (e-mail: haoweihua1999@qq.com; fengqian@uestc.edu.cn).

Gulan Zhang is with the School of Geoscience and Technology, Southwest Petroleum University, Chengdu 610500, China (e-mail: gulanzhang@163.com).

Yuehua Yue is with the Satellite Communication Technology Institute, Nanjing Panda Handa Technology Company, Ltd., Nanjing 210000, China (e-mail: 3341452205@qq.com).

Digital Object Identifier 10.1109/JSTARS.2023.3325969

classification by leveraging labeled training data. Nevertheless, annotated data are often constructed by interpreters from limited well-log information [53], making labeled data scarce and valuable. Therefore, the entire survey's depositional environment cannot be fully represented. These constraints ultimately hinder the application of supervised methods. Conversely, unsupervised end-to-end learning-based methods do not need labeled training data and try to discover hidden patterns in data, which routinely relies on embedding clustering layers or classifiers in neural networks to cluster features [5], [54], [55]. However, despite their effectiveness, the aforementioned embedding methods are not purely end-to-end, as they require a pretraining stage for cluster centre initialization [5]. Notably, the loss function used in the formal training stage differs from that in the pretraining stage, resulting in isolated optimization for the two stages. Consequently, these unsupervised SFA methods exacerbate the complexity of the analysis procedure.

To address the abovementioned issue, the variational autoencoder (VAE) provides a promising way. The VAE tends to encode input data into a distribution over a latent space, from which new data points can be generated through latent space sampling. This capacity of the VAE has found substantial application in seismic data processing, often leveraged for the generation of labeled data sources [56], [57], [58]. Beyond data generation, the potential to define the distribution of the latent space as a multimodal distribution [59] has extended the utility of VAE to data clustering tasks, such as the Gaussian mixture variational autoencoder-based deep clustering framework (GMVAE framework) [60], [61], [62]. Inspired by the GMVAE framework, we propose a lognormal mixture-based variational autoencoder (LMVAE) for the unsupervised SFA of prestack data. The proposed method effectively inherits the merits of the GMVAE framework. These advantages include the following: *1) No need for labeled data:* As a VAE approach, the LMVAE accomplishes unsupervised seismic facies clustering for prestack data without the reliance on labeled data. *2) End-to-end learning:* The network is capable of directly generating final results without the necessity for pretraining or any additional supplementary operations. In addition, the LMVAE seeks to model the features as a lognormal mixture distribution. Experimental results suggest that this modeling choice has enhanced the performance of LMVAE in handling field seismic data. The substantial contributions of our work are presented as follows.

1) In this study, the GMVAE framework [60] is applied to the unsupervised SFA of prestack seismic data. Through rigorous statistical analysis, employing techniques, such as Quantile-Quantile (QQ) plot [63], real seismic data are meticulously analyzed. The outcome of this analysis effectively corroborates the hypothesis of non-Gaussian properties inherent in seismic features, thereby enlightening the constraints posed by the GMVAE framework in the context of SFA.
2) In this article, we introduce remarkable technical innovations in the classical GMVAE framework, facilitating the creation of an unsupervised deep clustering approach specifically designed for prestack SFA. By expanding the GMVAE framework with the incorporation of a lognormal mixture-based modeling approach for seismic data features, we achieve effective solutions even in scenarios where the data features deviate from Gaussian distributions. To the best of authors' knowledge, this is the first instance of employing VAE-based methods for unsupervised seismic facies deep clustering.
3) The experimental results on both synthetic and field prestack data demonstrate that our suggested LMVAE approaches can be used to achieve remarkable improvements over the two state-of-the-art (SOTA) approaches.

The rest of this article are organized as follows. Section II provides some background knowledge on the GMVAE framework. In Section III, the problem statement and formulation of SFA are presented. Section IV describes the limitations of the GMVAE framework and the details of the proposed LMVAE method for improvement. The experimental results are reported in Section V. Finally, Section VI concludes this article.

II. Background

In this section, we shall provide a concise overview of some fundamental concepts related to GMVAE framework. In Section II-A, we begin by revisiting the fundamental concepts of the generative and recognition models of GMVAE. Afterward, in Section II-B, we provide a concise presentation of the variational lower bound of GMVAE.

A. Generative and Recognition Model

As an extended version of the VAE, the GMVAE is characterized by both a generative model and a recognition model. As detailed in [60], given the generative model $p_{\beta,\theta}(x,h,w,y) = p(w)p(y)p_\beta(h|y,w)p_\theta(x|h)$, the generative process of the sample x can be described as follows:

$$w \sim \mathcal{N}(0, I) \tag{1}$$

$$y \sim \text{Mult}(\pi) \tag{2}$$

$$h|y,w \sim \prod_{k=1}^{K} \left(\mathcal{N}(\mu_{y_k}(w;\beta), \text{diag}(\sigma^2_{y_k}(w;\beta)))\right)^{y_k} \tag{3}$$

$$x|h \sim \mathcal{N}(\mu(h;\theta), \text{diag}(\sigma^2(h;\theta))) \tag{4}$$

where from the perspective of clustering tasks, K is the predefined number of Gaussian components (i.e., the number of categories). Then, y is a one-hot vector following a multinoulli distribution, with a length of K, sampled from the mixture probabilities π. Specifically, y represents the weights of different mixture components, indicating the probabilities of x belonging to different classes. Meanwhile, h is a continuous latent variable following a Gaussian mixture distribution conditioned on y and w, representing the projection (or features) of x in the latent space. The neural networks parameterized by θ and β give $\mu_{y_k}(\cdot;\beta)$, $\sigma^2_{y_k}(\cdot;\beta)$, $\mu(\cdot;\theta)$, and $\sigma^2(\cdot;\theta)$, respectively.

Subsequently, the recognition model is factorized as follows [60]:

$$q(y,h,w|x) = q_{\phi_h}(h|x)\, q_{\phi_w}(w|x)\, p_\beta(y|h,w) \tag{5}$$

where ϕ_h and ϕ_w denote the parameters of the partial inference network responsible for parameterizing each variational factor and yielding means and diagonal covariances of the posterior variational distributions [60], respectively.

B. Variational Lower Bound

Since the GMVAE is trained using variational inference [61], the model trains the parameter ϕ_h, ϕ_w, β, and θ by maximizing the log-evidence lower bound (ELBO) [61], [62], [64]. As shown in [60], the ELBO used in the training process can be rewritten as follows:

$$\begin{aligned}\mathcal{L}_{\text{ELBO}} &= \mathbb{E}_q\left[\frac{p_{\beta,\theta}(x,h,w,y)}{q(h,y,w|x)}\right] \\ &= \mathbb{E}_{q(x|h)}[\log p_\theta(x|h)] \\ &\quad - \lambda\mathbb{E}_{q(w|x)p(y|h,w)}[KL(q_{\phi_h}(h|x)||p_\beta(h|w,y))] \\ &\quad - \gamma\text{KL}(q_{\phi_w}(w|x)||p(w)) \\ &\quad - \eta\mathbb{E}_{q(h|x)q(w|x)}[\text{KL}(p_\beta(y|h,w)||p(y))] \end{aligned} \tag{6}$$

where $\text{KL}(\cdot)$ stands for the Kullback–Leibler divergence, and the terms in $\mathcal{L}_{\text{ELBO}}$ can be called as the reconstruction term, conditional prior term, w-prior term, and y-prior term (for more details, see [60]), respectively. In addition, the balance parameters λ, γ, and η are introduced to control the weights of the different priors [65], and they are typically set to 1.

III. Problem Statement and Formulation

We assume that the prestack seismic data can be represented by a four-dimensional tensor $\mathcal{X} \in \mathbb{R}^{n_1\times n_2\times n_3\times n_4}$, where n_1, n_2, n_3, and n_4 are the sample numbers of the crossline, inline, timeline, and offset, respectively. The goal of SFA is to devise an automated mechanism to learn the deep features $\mathcal{H} \in \mathbb{R}^{n_1\times n_2\times m_3}$ that can best represent the prestack data $\mathcal{X}$ and then to utilize $\mathcal{H}$ to generate a facies map $Y \in \mathbb{R}^{n_1\times n_2}$ that can provide an insightful interpretation of the subsurface geology, where in accordance with the dimensionality of the features extracted by LMVAE, we assume $\mathcal{H}$ to be in a three-dimensional form, and m_3 represents the length of features corresponding to crossline and inline positions. Mathematically, for a given input $\mathcal{X}$, the corresponding facies map Y can be expressed as follows:

$$\begin{aligned}\mathcal{H} &= f(\mathcal{X},\Psi_1) \\ Y &= g(\mathcal{H},\Psi_2)\end{aligned} \tag{7}$$

where SFA relies on two mapping functions: $f(\cdot): \mathbb{R}^{n_1\times n_2\times n_3\times n_4} \to \mathbb{R}^{n_1\times n_2\times m_3}$, parameterized by Ψ_1, is designed to learn the best representation $\mathcal{H}$ from prestack data $\mathcal{X}$, and $g(\cdot): \mathbb{R}^{n_1\times n_2\times m_3} \to \mathbb{R}^{n_1\times n_2}$, parameterized by Ψ_2, is aimed transforming the incomprehensible seismic latent features $\mathcal{H}$ into a geologic interpretation Y.

Based on the difference in the optimization procedure of parameters Ψ_1 and Ψ_2, SFA can be categorized into two types: isolated learning-based SFA and end-to-end learning-based SFA. In the case of the isolated learning-based SFA, the optimization process of Ψ_1 and Ψ_2 is separated as follows:

$$\begin{aligned}\widehat{\Psi}_1 &= \underset{\Psi_1}{\text{argmin}}\,\mathcal{L}_1(\mathcal{X}, f(\mathcal{X},\Psi_1)) \\ \widehat{\Psi}_2 &= \underset{\Psi_2}{\text{argmin}}\,\mathcal{L}_2(f(\mathcal{X},\widehat{\Psi}_1), g(f(\mathcal{X},\widehat{\Psi}_1),\Psi_2)).\end{aligned} \tag{8}$$

In this category, Ψ_1 is optimized by minimizing the loss function $\mathcal{L}_1(\cdot)$, enabling the mapping function $f(\cdot)$ to estimate the most suitable representation of the prestack seismic data $\mathcal{X}$. Subsequently, the clustering loss $\mathcal{L}_2(\cdot)$ is minimized to learn $g(\cdot)$, parameterized by Ψ_2, thereby facilitating the transformation of the seismic latent features $\mathcal{H}$ into a geologic interpretation Y. Consequently, isolated learning-based SFA achieves the seismic data interpretation by optimizing Ψ_1 and Ψ_2 separately. However, it is worth noting that during the optimization of $\mathcal{L}_2(\cdot)$, $\widehat{\Psi}_1$ obtained from the previous stage may be fixed, which can potentially hinder the loss function from reaching a global optimum. This limitation might result in suboptimal performance of the isolated learning-based SFA approach.

As an alternative, end-to-end learning-based SFA can jointly learn representations of seismic data and perform clustering on learned features. The optimization process of end-to-end learning-based SFA is expressed as follows:

$$\{\widehat{\Psi}_1,\widehat{\Psi}_2\} = \underset{\Psi_1,\Psi_2}{\text{argmin}}\,\mathcal{L}(\mathcal{X}, g(f(\mathcal{X},\Psi_1),\Psi_2)). \tag{9}$$

By the abovementioned simultaneous optimization of both Ψ_1 and Ψ_2, loss $\mathcal{L}(\cdot)$ concerning feature extraction and clustering tries to theoretically reach the global minimum, avoiding the occurrence of suboptimal results that may arise in isolated learning-based SFA.

IV. Methodology

To obtain an efficient solution to Problem (9), the LMVAE, which overcomes the limitations of the traditional GMVAE framework, is presented in this section [60]. The drawbacks of the conventional GMVAE framework are discussed in Section IV-A, and the proposed LMVAE model is introduced in Section IV-B. Lastly, the fundamental structure of the LMVAE model is detailed in Section IV-C.

A. Shortcomings of the GMVAE Framework

As shown in (3), it is commonly assumed that the observed seismic data are generated from a mixture of Gaussians [60], [61], [62], [64]. Hence, during the new seismic data generation process, the input seismic data clustering is accomplished by determining which Gaussian component in the GMVAE's latent space the input data features belong to. Similar to (6), the network training process for this purpose can be formulated as follows:

$$\begin{aligned}\{\widehat{\Psi}_1,\widehat{\Psi}_2\} &= \underset{\Psi_1,\Psi_2}{\text{argmin}}\,-\mathcal{L}_{\text{ELBO}}(\mathcal{X};\Psi_1,\Psi_2) \\ &= \underset{\Psi_1,\Psi_2}{\text{argmin}}\,-\mathbb{E}_q\left[\frac{p_{\Psi_1,\Psi_2}(\mathcal{X},\mathcal{H},Y,w)}{q(\mathcal{H},Y,w|\mathcal{X})}\right]\end{aligned} \tag{10}$$

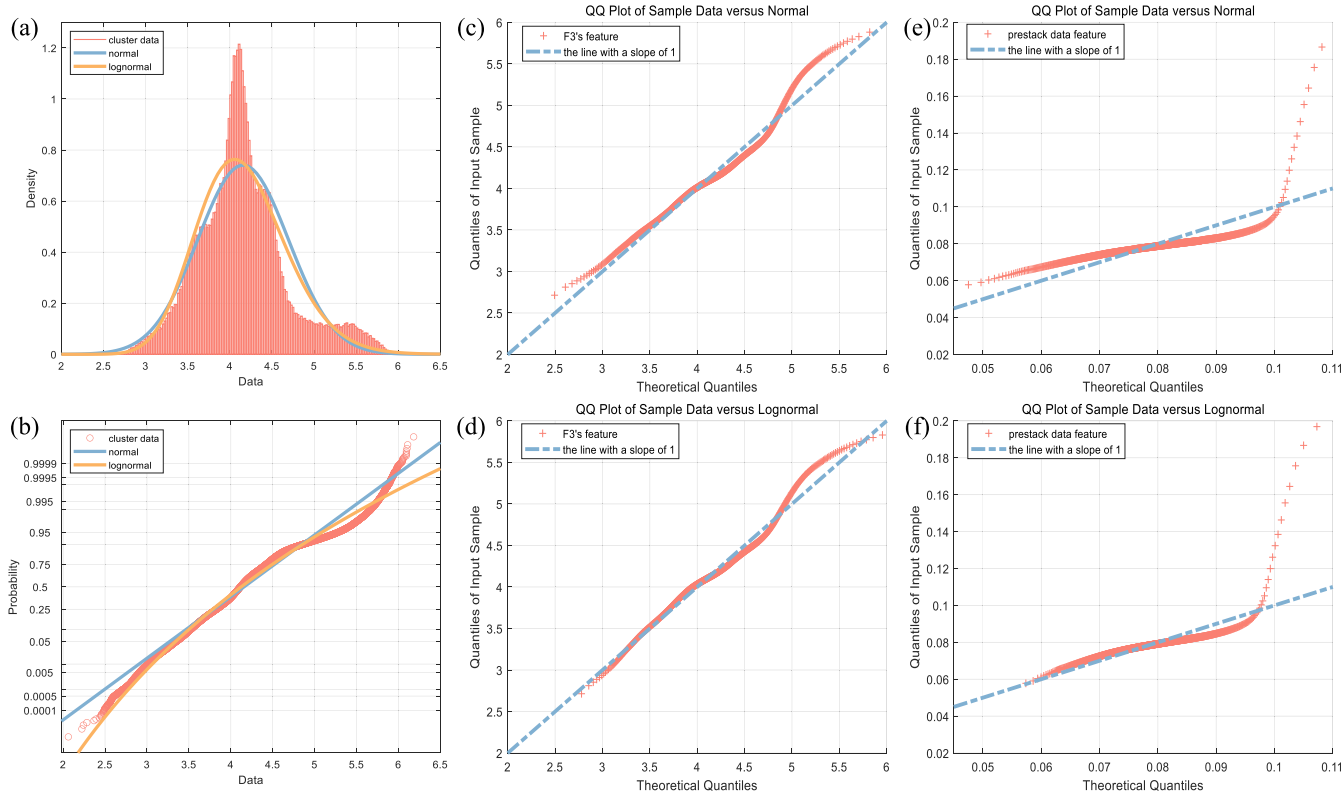


Fig. 1. Statistical plots of seismic data features (belonging to the same cluster). (a) Histogram of F3's features. (b) Probability plot of F3's features. (c) QQ plot between the Gaussian distribution and F3's features. (d) QQ plot between the lognormal distribution and F3's features. (e) QQ plot between the Gaussian distribution and prestack data features. (f) QQ plot between the lognormal distribution and prestack data features.

where the capital letters used here correspond to the lowercase letters introduced in Section II. Ψ_1 and Ψ_2 denote the parameters of the encoder and decoder, respectively.

However, by using the F3 facies map that is skillfully labeled by experts utilizing a precise geological model[1] [66], we may detect the non-Gaussian property of the seismic data, as described in the following steps. As shown in Fig. 1(a) and (b), the features of the F3 seismic samples belonging to the same cluster exhibit an asymmetric distribution, which deviates from the characteristics of a Gaussian distribution. Moreover, we employ QQ plots [63] to analyze the data in Fig. 1(a). In the QQ plot, instances that lie closer to the line with a slope of 1 indicate a stronger similarity between the samples and the theoretical distribution [63]. As depicted in Fig. 1(c) and (d), the QQ plot between the Gaussian distribution and seismic features deviates more from the line with a slope of 1 compared to the QQ plot between the lognormal distribution and seismic features. This observation further supports the non-Gaussian nature of the F3 features. To further corroborate the non-Gaussian nature of seismic data, additional analyzes were conducted. Fig. 1(e) and (f) presents the QQ plot of prestack data from Section V-D, in which the accurate SFA results of Fig. 7(e) are utilized to assign seismic feature categories in the GMVAE. In comparison to Fig. 1(e), the instances shown in Fig. 1(f) are obviously closer to the line with a slope of 1 in the range where the x-axis is less than 0.08. Moreover, in the range where the x-axis is greater than 0.08 but less than 0.1, although both instances exhibit some deviation from the line with a slope of 1, it is discernible that the instances shown in Fig. 1(f) display less deviation. In the range where the x-axis is greater than 0.1, both instances exhibit substantial deviations for a minority of data points. For these points, neither Gaussian nor lognormal distributions can adequately model them, suggesting the presence of a more suitable probability model for this data In summary, Gaussian probability modeling is not suitable for capturing the characteristics of seismic data.

[1][Online]. Available: https://github.com/yalaudah/facies_classification_benchmark

B. Lognormal Mixture-Based Variational Autoencoder

Inspired by the abovementioned discussion, we developed the LMVAE for prestack data SFA, substituting the probability modeling of $\mathcal{H}$ in (3) with the lognormal mixture distribution as follows:

$$\mathcal{H}|\mathrm{Y},\mathrm{w} \sim \prod_{\mathrm{k}=1}^{\mathrm{K}} \ln\mathcal{N}(\mu_{\mathrm{Y_k}}(\mathrm{w};\Psi_1),\mathrm{diag}(\sigma^2_{\mathrm{Y_k}}(\mathrm{w};\Psi_1)))^{\mathrm{Y_k}} \quad (11)$$

where $\ln\mathcal{N}(\cdot)$ denotes the lognormal distribution [67] with respect to the seismic data features belonging to the same cluster. Compared to the limitation of assuming that the observed data are generated from a mixture of Gaussian distributions, the LMVAE may achieve a better performance in handling field seismic data.

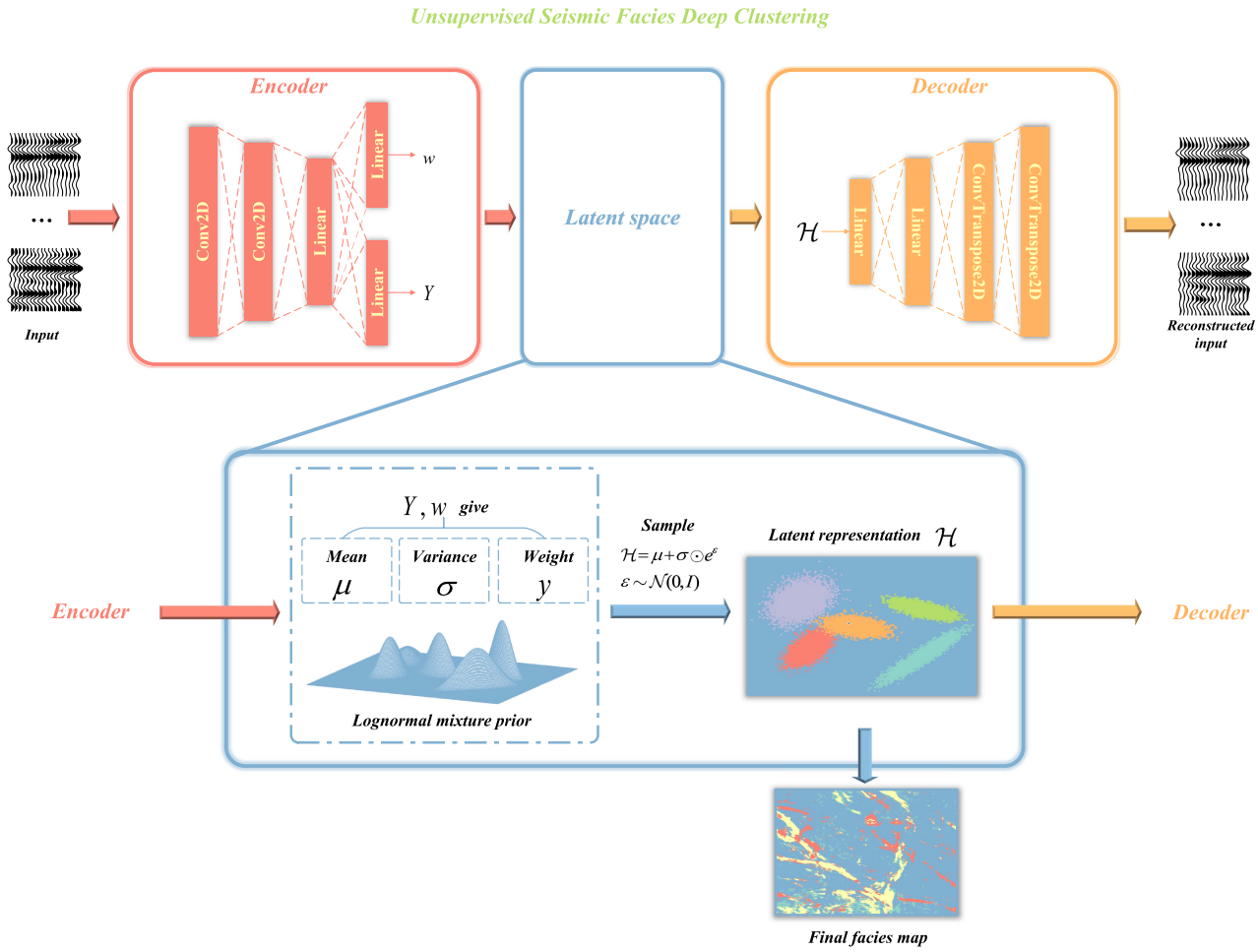


Fig. 2. Schematic of LMVAE.

Furthermore, the $\mathcal{L}_{\text{ELBO}}$ of the LMVAE is akin to the GMVAE. Similarly, the reconstruction term can be formulated utilizing either binary cross-entropy loss or mean squared error loss, while the w-prior term, representing the divergence between two Gaussian variables, can be computed analytically [60]. In the case of the conditional prior term and the y-prior term, their estimations necessitate the utilization of Monte Carlo samples. Drawing upon the conclusions presented in [60], these terms can be expressed as follows:

$$\begin{aligned}&\mathbb{E}_{q(w|\mathcal{X})p(Y|\mathcal{H},w)}\left[KL\left(q_{\phi_{\mathcal{H}}}(\mathcal{H}|\mathcal{X})||p_\beta(\mathcal{H}|w,Y)\right)\right]\\&=\sum_{k=1}^{K}p_\beta\left(Y_k=1|\mathcal{H},w\right)KL\left(q_{\phi_{\mathcal{H}}}(\mathcal{H}|\mathcal{X})||p_\beta\left(\mathcal{H}|w,Y_k=1\right)\right)\end{aligned}\tag{12}$$

$$\begin{aligned}&\mathbb{E}_{q(\mathcal{H}|\mathcal{X})q(w|\mathcal{X})}[KL(p_\beta(Y|\mathcal{H},w)||p(Y))]\\&=\sum_{k=1}^{K}p_\beta\left(Y_k=1|\mathcal{H},w\right)\left(\ln p_\beta\left(Y_k=1|\mathcal{H},w\right)+\ln K\right)\end{aligned}\tag{13}$$

where the computation of the conditional prior term [see (12)] requires the use of the Kullback–Leibler divergence between two lognormal distribution variables (see [67]).

C. Network Structure of the LMVAE

The architecture of the LMVAE closely resembles that of GMVAE [60], leveraging the capabilities of two distinct networks, the inference (encoder) and generative (decoder) networks, to effectively estimate the posterior probability of latent variables and generate seismic data likelihood. As depicted in Fig. 2, the encoder of the LMVAE employs two convolutional layers (kernel size 3×3) and a series of parallel linear layers. Following each convolutional layer, there is a sequential connection of a batch normalization (BN) layer, a leaky rectified linear unit (LeakyReLU) activation function, and a max-pooling layer with dimensions of 2×2. For the sake of simplicity, these components are not explicitly depicted in Fig. 2. Once the convolutional architecture accomplishes the initial feature extraction, a series of parallel linear layers are utilized to generate w and Y. After sampling the specific values of $\mathcal{H}$ based on its mean and variance, the decoder proceeds to reconstruct new data from $\mathcal{H}$. This process involves expanding the dimensions of $\mathcal{H}$ through two connected linear layers, followed by a convolutional structure consisting of upsampling layers, transposed convolutional layers, BN layers, and LeakyReLU layers to reconstruct the features.

In fact, the encoder, employing parallel linear layers, simultaneously output the value of Y, as well as the mean and variance of w and $\mathcal{H}$. The value of Y undergoes a softmax layer to obtain seismic facies results. Meanwhile, by sampling from the mean and variance, the specific values of $\mathcal{H}$ are acquired. Since $\mathcal{H}$ belongs to one of the components in the lognormal mixture distribution, its sampling strategy is as described in [67]

$$\mathcal{H} = \mu + \sigma \odot e^{\varepsilon}$$
$$\varepsilon \sim \mathcal{N}(0, I). \tag{14}$$

Similar with [60], w follows a Gaussian distribution, and its sampling employs a Gaussian variable sampling strategy. After obtaining the specific values of w, a linear layer is employed to generate the prior for $\mathcal{H}$.

V. Experiments

In this section, the performance of the proposed LMVAE is validated using one synthetic and two field prestack seismic datas. All experiments are set to have five seismic facies. Subsequently, the SFA results obtained by the LMVAE are compared with those of three SOTA methods, i.e., Gabor [68], deep convolutional autoencoder (DCAE) [21], and GMVAE [60]. Both the Gabor and DCAE methods are isolation-based methods. Gabor feature extraction is implemented using MATLAB, while DCAE is implemented using PyTorch. After feature extraction, self-organizing maps (SOM) [69] algorithms are employed to cluster the features and generate the final seismic facies maps. The SOM implementation is also carried out in MATLAB. In contrast, the GMVAE and LMVAE are end-to-end learning methods. They directly generate seismic facies maps and are implemented using PyTorch. All experiments are performed on a computer with a 4-core Intel i5-6500 CPU, 32 GB of memory, and an NVIDIA TITAN X (Pascal) GPU. The LMVAE is trained with the Adam optimization algorithm. Specifically, for training on synthetic data, 500 epochs of training are conducted with a learning rate of 5×10^{-7}. When processing the Liziba data, the model undergoes 150 epochs of training with a learning rate of 3×10^{-5}. In the case of the Zhongjiang data, 500 epochs of training are carried out with a learning rate of 7×10^{-5}. The parameters λ, γ, and η in the $\mathcal{L}_{\text{ELBO}}$ are typically set to 1.

A. Performance Metrics

1) RMS Amplitude Attributes: Following the methodology established in prior works [5], [21], the accuracy of facies maps is assessed using the rms amplitude attributes of seismic data, as shown in Figs. 4(a), 5(a), and 7(a).

2) t-Distributed Stochastic Neighbor Embedding (t-SNE): The t-SNE is a nonlinear dimensionality reduction technique employed for visualizing high-dimensional data in a lower dimensional space [70]. In particular, given points b_i in the high-dimensional space and points o_i in the low-dimensional space, t-SNE utilizes joint probability distributions and Student t-distribution to depict the similarity of data in both the high-dimensional and low-dimensional spaces, as illustrated in the following:

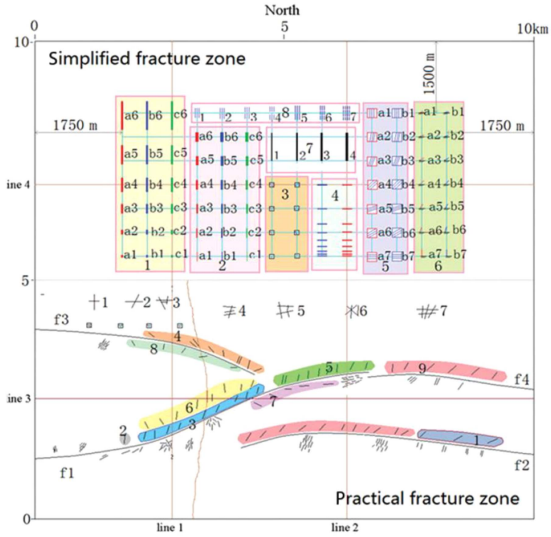
$$p_{ij} = \frac{\exp\left(-\|b_i - b_j\|_2^2 / 2\delta^2\right)}{\sum_{m \neq n} \exp\left(-\|b_m - b_n\|_2^2 / 2\delta^2\right)}$$
$$q_{ij} = \frac{\left(1 + \|o_i - o_j\|_2^2\right)^{-1}}{\sum_{m \neq n} \left(1 + \|o_m - o_n\|_2^2\right)^{-1}} \tag{15}$$

where δ is the variance parameter.

Subsequently, t-SNE determines the precise positions of points o_i by minimizing the following Kullback–Leibler divergence:

$$\text{KL}(P||Q) = \sum_i \sum_j p_{ij} \log \frac{p_{ij}}{q_{ij}}. \tag{16}$$

(a)



(b)

Fig. 3. Schematic of the physical model. (a) Preparation of the physical model. (b) Prototype of the physical model.

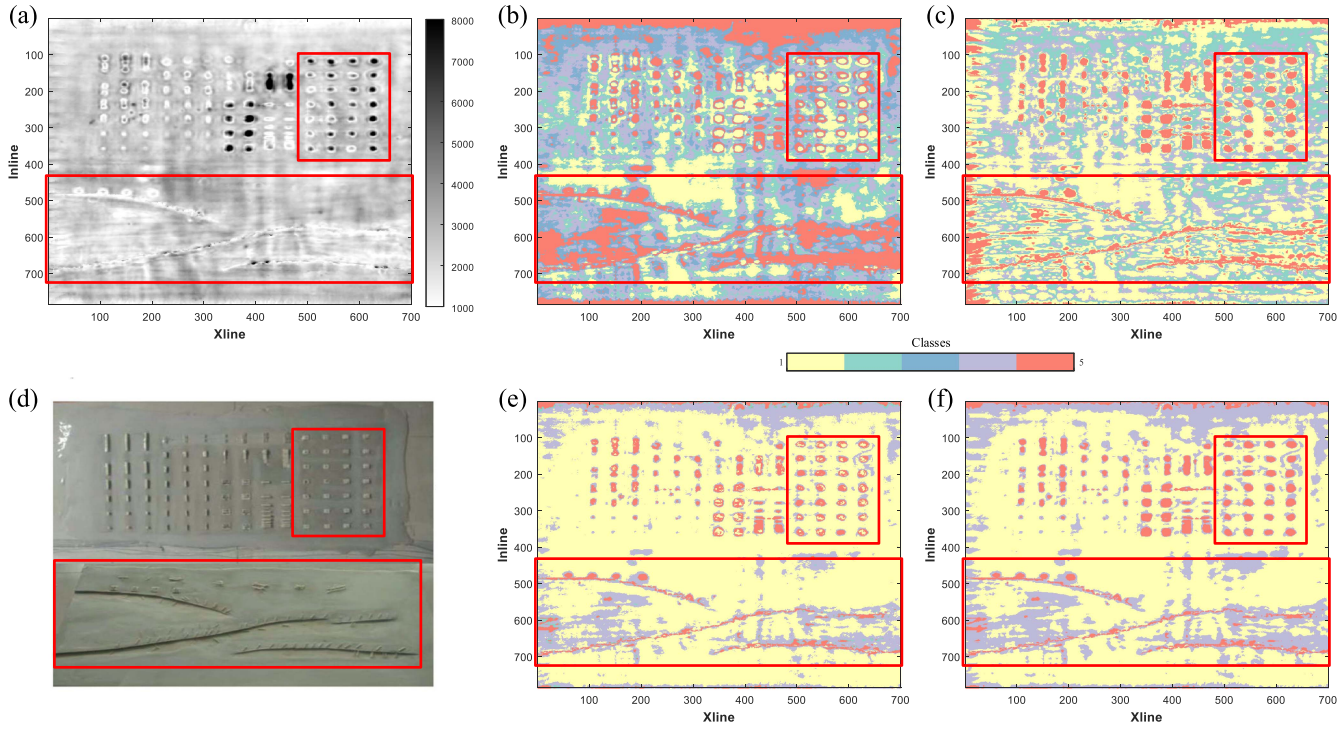


Fig. 4. SFA results of synthetic data. (a) RMS amplitude attribute of synthetic data. (b) Gabor+SOM. (c) DCAE+SOM. (d) Photo of the physical model. (e) GMVAE. (f) LMVAE.

t-SNE finds utility in visualization for both labeled and unlabeled datasets [71], [72], [73]. To facilitate further comparison of the experimental outcomes, we utilized the t-SNE for visualizing the latent representations of data learned by these methods, as shown in Figs. 6(a)–(d) and 8(a)–(d).

3) Davies Bouldin Index: In addition, quantitative metrics are employed to evaluate the algorithm's performance. As SFA involves clustering tasks, when sample labels are known, external indicator like normalized mutual information [74] are commonly used to assess clustering performance. However, in practical seismic data processing, acquiring accurate seismic phase labels is challenging, often relying on limited well-log data as references [53]. Therefore, we resort to internal indicator, such as the Davies Bouldin index (DBI) [75] to quantitatively compare the performance of our LMVAE with other SOTA methods. Specifically, DBI is utilized to compute the ratio of the sum of intracluster distances to the intercluster distances. A smaller DBI value indicates a better clustering performance, as it signifies more compact and well-separated clusters. It is mathematically defined as follows:

$$\mathrm{DBI} = \frac{1}{C}\sum_{i=1}^{C} \max_{j\neq i}\left\{\frac{S_i + S_j}{d_{ij}}\right\} \tag{17}$$

where C denotes the number of clusters, S_i and S_j are present the average distances of each element to their corresponding cluster i and j centroids, and d_{ij} is defined as the distance between centroids of clusters i and j.

B. Validation on Synthetic Data

To demonstrate the effectiveness of our LMVAE, we implement SFA on the synthetic prestack data from an artificial physical model [21] that simulates the Longgang block of China. By controlling various fracture parameters in terms of fracture lengths and orientations [using the preparation method shown in Fig. 3(a)], the physical model can simulate the geological environment of the Longgang fracture zone. The prototype of the physical model is depicted in Fig. 3(b), consisting of a simplified practical fracture zone encompassing various fractures and caves.

The SFA results of the Gabor, DCAE, GMVAE, and LMVAE methods are presented in Fig. 4. To better illustrate the advantages of the LMVAE, we have highlighted selected regions within red boxes for comparison. From Fig. 4(b) and (c), it can be observed that Gabor and DCAE are capable of identifying most of the caves in the facies map, as evident in the upper part of the facies map within the red-boxed regions. However, when it comes to fractures, both methods struggle to recognize them accurately. In comparison to the rms attribute in Fig. 4(a) and the photo of the physical model in Fig. 4(d), within the red-boxed area at the bottom of the facies map, Gabor is almost unable to discern the shape of the fractures, while DCAE, though able to roughly identify the shape, results in a cluttered background in the facies map, leading to blurred fractures boundaries. In contrast, both LMVAE and GMVAE exhibit superior performance in identifying caves and fractures in the facies map and produce clear facies map backgrounds, as illustrated in Fig. 4(e) and (f). However, GMVAE's identification of caves is not flawless,

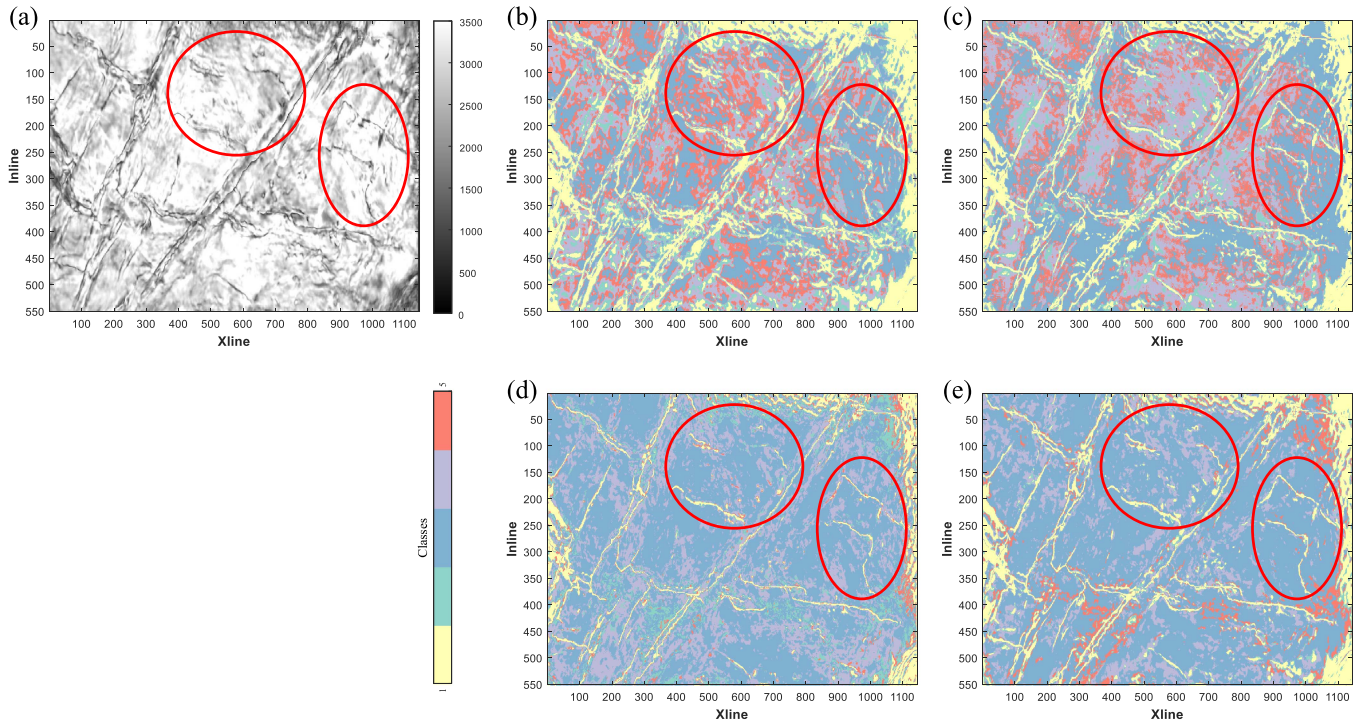


Fig. 5. SFA results of Liziba field data. (a) RMS amplitude attribute of field data. (b) Gabor+SOM. (c) DCAE+SOM. (d) GMVAE. (e) LMVAE.

TABLE I
DBI PERFORMANCE COMPARISON OF THE PROPOSED METHOD WITH EXISTING METHODS

Dataset	Gabor	DCAE	GMVAE	LMVAE
Synthetic data	43.39	33.25	23.46	**11.77**
Liziba field data	27.27	24.66	30.40	**23.62**
Zhongjiang field data	14.57	4.99	6.94	**4.46**

The bold entities are intended to emphasize that our proposed method achieves the best performance metrics.

as there are irregularly distributed other facies classes within the caves, as demonstrated in the red-boxed region. Besides visual comparisons, a quantitative assessment based on DBI between LMVAE and the SOTA methods is presented in Table I. Our LMVAE achieved a lower DBI value (11.77) compared to Gabor, DCAE, and GMVAE, where their respective DBI values were 43.39, 33.25, and 23.46. Therefore, based on the abovementioned comparisons, the results of the LMVAE are more proximate to the artificial physical model.

C. Validation on Liziba Field Data

To verify that the LMVAE can exhibit superior performance in practice, we conduct experiments on a field exploration region, namely, the Liziba region in southwestern Sichuan Province, China [21]. The final facies maps generated by the DCAE, GMVAE, and LMVAE are presented in Fig. 5, where several areas are circled to facilitate the comparison of the results. As shown in Fig. 5(b) and (c), Gabor can only identify a partial number of faults and caves, while DCAE accurately recognizes most of them, but both methods produce facies maps with blurred backgrounds. Moreover, as depicted in Fig. 5(d) and (e), both

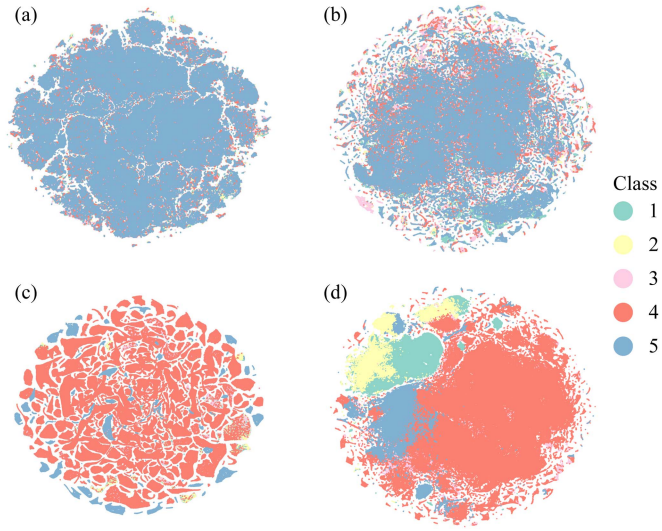


Fig. 6. t-SNE plots for the latent feature space of the (a) Gabor, (b) DCAE, (c) GMVAE, and (d) LMVAE applied to Liziba field data whose colour is labeled according to the corresponding SFA results.

GMVAE and LMVAE effectively identify faults and caves with higher overall resolution in the facies maps. However, within the delineated region, LMVAE achieves even higher resolution (resulting in clearer fault outlines) compared to GMVAE. In Table I, the DBI comparisons are presented, and in DBI comparison to Gabor, DCAE, and GMVAE with DBI values of 27.27, 24.66, and 30.40, respectively, LMVAE exhibits the lowest DBI value of 23.62. Thus, as theoretically anticipated, the LMVAE is capable of better achieving SFA for field data.

Furthermore, the t-SNE results of the features extracted by each method are depicted in Fig. 6(a)–(d). The partition of latent

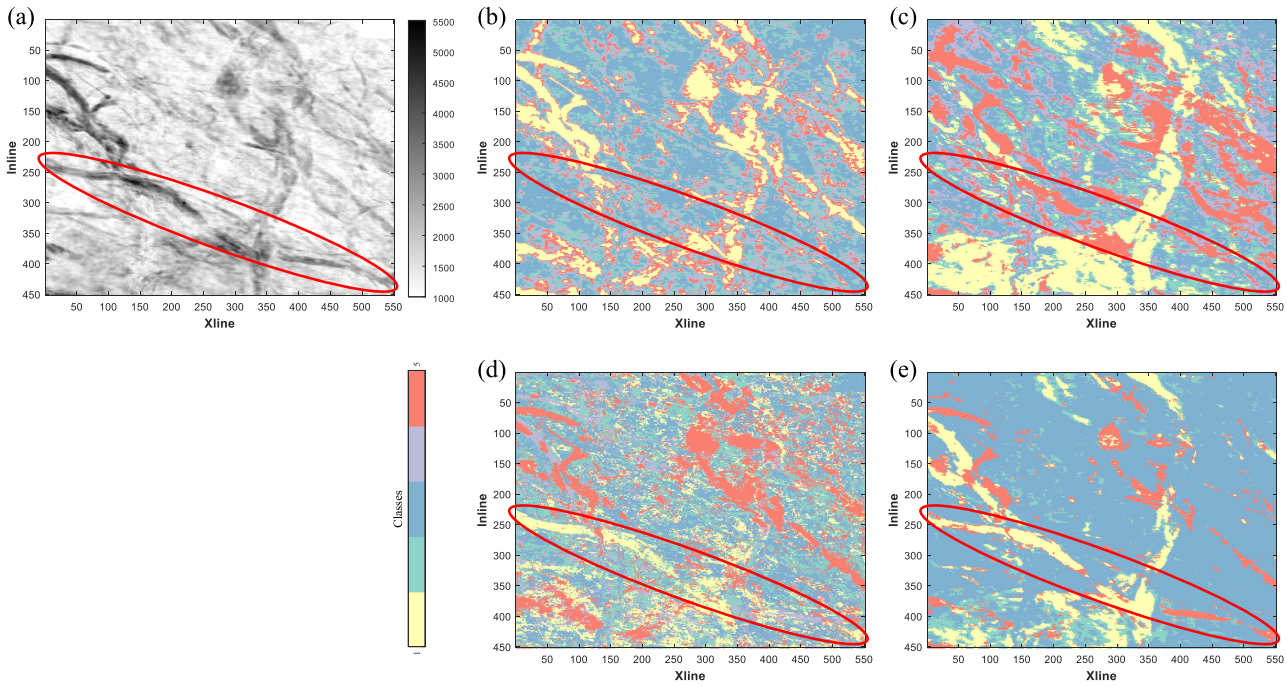


Fig. 7. SFA results of Zhongjiang field data. (a) RMS amplitude attribute of field data. (b) Gabor+SOM. (c) DCAE+SOM. (d) GMVAE. (e) LMVAE.

representations in both Gabor and DCAE are not discernible [see Fig. 6(a) and (b)], while that in the GMVAE is excessively scattered [see Fig. 6(c)]. In contrast, the LMVAE demonstrates significantly more reasonable partitioning in that associated clusters are well-separated and distinguishable [see Fig. 6(d)], revealing that the LMVAE outperforms both Gabor, DCAE, and GMVAE in terms of clustering the features into more distinct and separable clusters.

D. Validation on Zhongjiang Field Data

To further verify the effectiveness of the LMVAE, we apply it to a field exploration region covering the Zhongjiang area and located in northeastern Sichuan Province, China. This region presents unique challenges due to the complex distribution of gas and water, including a narrow distribution width of the underwater channel and numerous overlapping relationships among different geological features. Fig. 7 shows the SFA results of the four methods, and part of the maps is circled for a prominent comparison. From Fig. 7(b) and (c), it can be observed that Gabor accurately identifies most of the features but completely fails to recognize the underwater channels within the circled region. On the other hand, DCAE can only identify a small portion of the channels. Similarly, from Fig. 7(d), it can be seen that GMVAE roughly identifies the wider underwater channels within the circled region, but the overall SFA map appears too blurry and chaotic. In contrast, the LMVAE precisely identifies the underwater channels within the circled region, as shown in Fig. 7(e). The overall results are clearer and more continuous compared to other SOTA algorithms. Moreover, the fact that the LMVAE achieves the lowest DBI value of 4.46 among all methods, as shown in Table I, indicates a significant enhancement in the SFA performance for field seismic data.

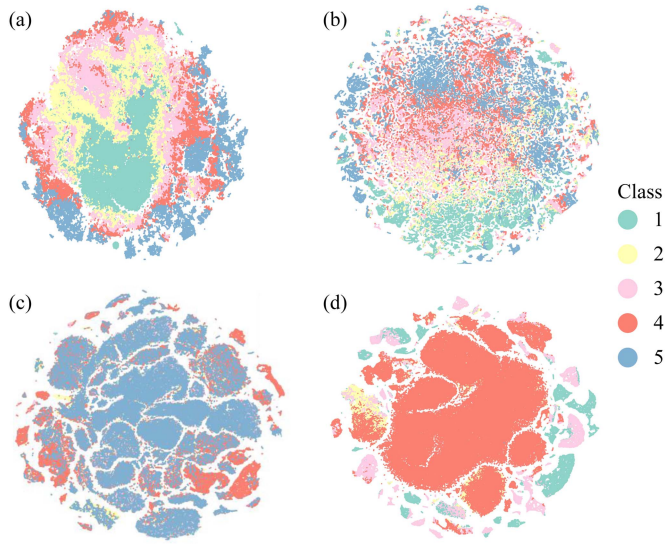


Fig. 8. t-SNE plots for the latent feature space of the (a) Gabor, (b) DCAE, (c) GMVAE, and (d) LMVAE applied to Zhongjiang field data whose colour is labeled according to the corresponding SFA results.

The t-SNE technique was also employed in the experiments of this dataset, as shown in Fig. 8(a)–(d). In the latent representations obtained by DCAE, the data points are scattered too widely without clear clustering [see Fig. 8(b)]. On the other hand, in the latent representations of GMVAE, there are an excessive number of clusters with indistinct categories [see Fig. 8(c)]. In comparison, Gabor exhibits a relatively clear categorization of

TABLE II
COMPUTATION TIME COMPARISON OF THE PROPOSED METHOD WITH THE EXISTING METHODS

Dataset	Gabor	DCAE	GMVAE	LMVAE
Synthetic data	813.34 s	1698.07 s	8266.56 s	7303.02 s
Liziba field data	1282.89 s	979.49 s	1922.73 s	1673.12 s
Zhongjiang field data	863.76 s	159.23 s	5492.37 s	5293.62 s

the latent representations [see Fig. 8(a)], but LMVAE shows even more distinct clustering, with well-separated relevant clusters [see Fig. 8(d)]. This observation indicates that LMVAE produces a well-defined latent space with meaningful clustering.

E. Discussion

Through experimental validation with both physical models and two field datasets, our proposed LMVAE successfully achieves automated end-to-end SFA on seismic data. However, a notable limitation of the LMVAE, in comparison to other benchmarked methods, is its computational efficiency. As shown in Table II, the GMVAE exhibits relatively longer execution times, and the LMVAE, which builds upon GMVAE improvements, shares a similar characteristic. When compared to DCAE, another unsupervised deep learning method, the increased complexity in LMVAE's network architecture and loss functions may necessitate additional training time.

Similar to the GMVAE, the LMVAE is sensitive to its parameters, and achieving optimal results requires meticulous experimentation. Parameters such as the number of training epochs and the learning rate of the optimization algorithm need to be carefully tuned. Furthermore, in addition to these network-specific parameters, achieving superior performance may necessitate further adjustments to the balance parameters λ, γ, and η in the $\mathcal{L}_{\text{ELBO}}$, like in [65]. As shown in [60], it is also crucial to consider phenomena, such as over regularization in the GMVAE framework.

VI. CONCLUSION

In this article, an unsupervised SFA method, the LMVAE, which is based on the GMVAE framework, is proposed to automatically classify prestack seismic data into distinct facies categories. In this LMVAE approach, SFA is formulated as a joint feature extraction and clustering problem. Specifically, in our proposed LMVAE method, feature extraction and clustering are performed simultaneously in an end-to-end manner by modeling the probability distribution of features as a lognormal mixture distribution, thereby assigning class labels while generating the features. Moreover, experimental validations suggest that the lognormal mixture distribution may better suit field seismic data, leading to improved accuracy of the LMVAE method. The experiments conducted on synthetic and real datasets demonstrate the superiority of our proposed SFA method in achieving more effective delineation of interfaces compared to several SOTA methods.

In our future work, we may explore combining the proposed method with other SOTA convolutional neural network approaches (e.g., attention mechanisms). In addition, we plan to investigate the applicability of other GMVAE framework variants in SFA.

ACKNOWLEDGMENT

The authors would like to thank the associate editor, and the three anonymous reviewers, whose constructive suggestions helped to greatly improve and clarify this article.

REFERENCES

[1] J. Dumay and F. Fournier, "Multivariate statistical analyses applied to seismic facies recognition," *Geophysics*, vol. 53, no. 9, pp. 1151–1159, Sep. 1988.
[2] B. P. West, S. R. May, J. E. Eastwood, and C. Rossen, "Interactive seismic facies classification using textural attributes and neural networks," *Leading Edge*, vol. 21, no. 10, pp. 1042–1049, Oct. 2002.
[3] S. Yu and J. Ma, "Deep learning for geophysics: Current and future trends," *Rev. Geophys.*, vol. 59, no. 3, Sep. 2021, Art. no. e2021RG000742.
[4] M. Liu, M. Jervis, W. Li, and P. Nivlet, "Seismic facies classification using supervised convolutional neural networks and semisupervised generative adversarial networks," *Geophysics*, vol. 85, no. 4, pp. O47–O58, Jun. 2020.
[5] Y. Duan, X. Zheng, L. Hu, and L. Sun, "Seismic facies analysis based on deep convolutional embedded clustering," *Geophysics*, vol. 84, no. 6, pp. IM87–IM97, Nov. 2019.
[6] T. Zhao, V. Jayaram, A. Roy, and K. J. Marfurt, "A comparison of classification techniques for seismic facies recognition," in *Interpretation*, vol. 3, no. 4, pp. SAE29–SAE58, Nov. 2015.
[7] F. Qian, Y. He, Y. Yue, Y. Zhou, B. Wu, and G. Hu, "Improved low-rank tensor approximation for seismic random plus footprint noise suppression," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, Feb. 2023, Art. no. 5903519.
[8] F. Qian et al., "Unsupervised seismic footprint removal with physical prior augmented deep autoencoder," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, May 2023, Art. no. 5910920.
[9] Z. Liu, J. Cao, Y. Lu, P. Zhou, and J. Hu, "A hierarchical clustering method of SOM based on DTW distance for variable-length seismic waveform," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, Aug. 2022, Art. no. 7505005.
[10] Z. Hong et al., "A DTW distance-based seismic waveform clustering method for layers of varying thickness," *Appl. Geophys.*, vol. 17, no. 2, pp. 171–181, Jun. 2020.
[11] Z. Hong, K. Li, M. Su, and G. Hu, "Improved spectral clustering approach a new tool for unsupervised seismic facies analysis of variable window length," in *Interpretation*, vol. 9, no. 2, pp. T407–T420, Apr. 2021.
[12] P. Saraswat and M. K. Sen, "Artificial immune-based self-organizing maps for seismic-facies analysis," *Geophysics*, vol. 77, no. 4, pp. O45–O53, Jun. 2012.
[13] M. C. de Matos, P. L. Osorio, and P. R. Johann, "Unsupervised seismic facies analysis using wavelet transform and self-organizing maps," *Geophysics*, vol. 72, no. 1, pp. P9–P21, Jan. 2007.
[14] K. Bauer, B. Norden, A. Ivanova, M. Stiller, and C. M. Krawczyk, "Wavelet transform-based seismic facies classification and modelling: Application to a geothermal target horizon in the NE German basin," *Geophysical Prospecting*, vol. 68, no. 2, pp. 466–482, Jan. 2020.
[15] Y.-J. Xue, J.-X. Cao, R.-F. Tian, H.-K. Du, and Y.-X. Shu, "Application of the empirical mode decomposition and wavelet transform to seismic reflection frequency attenuation analysis," *J. Petroleum Sci. Eng.*, vol. 122, pp. 360–370, Oct. 2014.
[16] H.-K. Du, J.-X. Cao, Y.-J. Xue, and X.-j. Wang, "Seismic facies analysis based on self-organizing map and empirical mode decomposition," *J. Appl. Geophys.*, vol. 112, pp. 52–61, Jan. 2015.
[17] A. Amendola, G. Gabbriellini, P. Dell'Aversana, and A. Marini, "Seismic facies analysis through musical attributes," *Geophys. Prospecting*, vol. 65, no. S1, pp. 49–58, Jan. 2017.
[18] T. Xie, X. Zheng, and Y. Zhang, "Seismic facies analysis based on speech recognition feature parameters," *Geophysics*, vol. 82, no. 3, pp. O23–O35, Mar. 2017.
[19] F. Xu, Z. Li, Y. Luo, Y. Huang, and Y. Wang, "Unsupervised seismic facies pattern recognition based on 3D Gabor transform and two-step cluster analysis," *Solid State Technol.*, vol. 64, no. 2, pp. 3489–3505, 2021.

[20] X. Feng, B. Wen, Q. Yang, and Z. Li, "Seismic facies recognition based on prestack data using two-dimensional Gabor transform and unsupervised clustering," in *Proc. IEEE 2nd Int. Conf. Inf. Commun. Signal Process.*, 2019, pp. 509–513.

[21] F. Qian, M. Yin, X.-Y. Liu, Y.-J. Wang, C. Lu, and G.-M. Hu, "Unsupervised seismic facies analysis via deep convolutional autoencoders," *Geophysics*, vol. 83, no. 3, pp. A39–A43, May 2018.

[22] V. Puzyrev and C. Elders, "Unsupervised seismic facies classification using deep convolutional autoencoder," *Geophysics*, vol. 87, no. 4, pp. 1–39, Jul. 2022.

[23] B. Wen, X. Feng, Q. Yang, and Z. Li, "Seismic facies recognition based on prestack data using two-dimensional convolutional auto-encoder and cluster analysis," in *Proc. IEEE 2nd Int. Conf. Inf. Commun. Signal Process.*, 2019, pp. 430–434.

[24] K. Li, Z. Liu, B. She, G. Hu, and C. Song, "Orthogonal deep autoencoders for unsupervised seismic facies analysis," in *Proc. SEG Int. Expo. Annu. Meeting*, 2019.

[25] X. Liu et al., "Deep classified autoencoder for lithofacies identification," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, Dec. 2022, Art. no. 5909914.

[26] X. Liu, B. Li, J. Li, X. Chen, Q. Li, and Y. Chen, "Semi-supervised deep autoencoder for seismic facies classification," *Geophys. Prospecting*, vol. 69, no. 6, pp. 1295–1315, Apr. 2021.

[27] Y. Zhou and W. Chen, "Recurrent auto-encoder model for unsupervised seismic facies analysis," in *Interpretation*, vol. 10, no. 3, pp. 1–44, Aug. 2022.

[28] M. Liu, P. Nivlet, R. Smith, N. BenHasan, and D. Grana, "Recurrent neural network for seismic reservoir characterization," in *Advances in Subsurface Data Analytics*. Amsterdam, The Netherlands: Elsevier, May 2022, pp. 95–116.

[29] M. Tian and S. Verma, "Recurrent neural network: Application in facies classification," in *Advances in Subsurface Data Analytics*. Amsterdam, The Netherlands: Elsevier, May 2022, pp. 65–94.

[30] Z. Liu, J. Cao, S. Chen, Y. Lu, and F. Tan, "Visualization analysis of seismic facies based on deep embedded SOM," *IEEE Geosci. Remote Sens. Lett.*, vol. 18, no. 8, pp. 1491–1495, Aug. 2021.

[31] S. Chen, Z. Liu, H. Zhou, X. Wen, and Y. Xue, "Seismic facies visualization analysis method of SOM corrected by uniform manifold approximation and projection," *IEEE Geosci. Remote Sens. Lett.*, vol. 20, Mar. 2023, Art. no. 7501805.

[32] H. Sabeti and A. Javaherian, "Seismic facies analysis based on K-means clustering algorithm using 3D seismic attributes," in *Proc. 1st EAGE Int. Petroleum Conf. Exhib.*, May 2009, Art. no. cp–125–00026.

[33] C. Song, L. Li, L. Li, and K. Li, "Robust K-means algorithm with weighted window for seismic facies analysis," *J. Geophys. Eng.*, vol. 18, no. 5, pp. 618–626, Aug. 2021.

[34] E. B. Troccoli, A. G. Cerqueira, J. B. Lemos, and M. Holz, "K-means clustering using principal component analysis to automate label organization in multi-attribute seismic facies analysis," *J. Appl. Geophys.*, vol. 198, Mar. 2022, Art. no. 104555.

[35] O. O. Napoli et al., "Accelerating multi-attribute unsupervised seismic facies analysis with rapids," 2020, *arXiv:2007.15152*.

[36] Z. Zhu et al., "Seismic facies analysis using the multiattribute SOM-K-means clustering," *Comput. Intell. Neurosci.*, vol. 2022, Oct. 2022, Art. no. 1688233.

[37] B. C. Wallet and R. Hardisty, "Unsupervised seismic facies using Gaussian mixture models," in *Interpretation*, vol. 7, no. 3, pp. SE93–SE111, May 2019.

[38] D. Grana, L. de Figueiredo, and K. Mosegaard, "Markov chain Monte Carlo for seismic facies classification," *Geophysics*, vol. 88, no. 3, pp. 1–67, Jan. 2023.

[39] D. Grana, L. Azevedo, and M. Liu, "A comparison of deep machine learning and Monte Carlo methods for facies classification from seismic dataseismic facies classification," *Geophysics*, vol. 85, no. 4, pp. WA41–WA52, Jan. 2020.

[40] K. Sang, N. Lan, and F. Zhang, "Seismic facies analysis method based on spectral clustering machine learning," in *Proc. SEG/AAPG Int. Meeting Appl. Geosci. Energy*, 2022.

[41] W. Yao-Jun, W. Liang-Ji, L. Kun-Hong, L. Yu, L. Xian-Zhe, and X. Kai, "Unsupervised seismic facies analysis using sparse representation spectral clustering," *Appl. Geophys.*, vol. 17, pp. 533–543, May 2020.

[42] K. Noh, D. Kim, and J. Byun, "Intrinsically explainable deep learning for supervised seismic facies classification using intrinsic method," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, Jan. 2023, Art. no. 5901711.

[43] R. Feng, N. Balling, D. Grana, J. S. Dramsch, and T. M. Hansen, "Bayesian convolutional neural networks for seismic facies classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 10, pp. 8933–8940, Oct. 2021.

[44] F. Li, H. Zhou, Z. Wang, and X. Wu, "ADDCNN: An attention-based deep dilated convolutional neural network for seismic facies analysis with interpretable spatial– spectral maps," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 2, pp. 1733–1744, Feb. 2021.

[45] Z. Wang, F. Li, T. R. Taha, and H. R. Arabnia, "Improved automating seismic facies analysis using deep dilated attention autoencoders," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops*, 2019, pp. 511–513.

[46] Z. Wang, Q. Wang, Y. Yang, N. Liu, Y. Chen, and J. Gao, "Seismic facies segmentation via a segformer-based specific encoder-decoder-hypercolumns scheme," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, Feb. 2023, Art. no. 5903411.

[47] X. Chai et al., "An open-source package for deep-learning-based seismic facies classification: Benchmarking experiments on the SEG 2020 open data," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, Jan. 2022, Art. no. 4507719.

[48] B. Abid, B. M. Khan, and R. A. Memon, "Seismic facies segmentation using ensemble of convolutional neural networks," *Wireless Commun. Mobile Comput.*, vol. 2022, pp. 1–13, Mar. 2022.

[49] Y. Zhang, Y. Liu, H. Zhang, and H. Xue, "Seismic facies analysis based on deep learning," *IEEE Geosci. Remote Sens. Lett.*, vol. 17, no. 7, pp. 1119–1123, Jul. 2020.

[50] H. Zhang, T. Chen, Y. Liu, Y. Zhang, and J. Liu, "Automatic seismic facies interpretation using supervised deep learning," *Geophysics*, vol. 86, no. 1, pp. IM15–IM33, Jan. 2021.

[51] X. Chen, Q. Zou, X. Xu, and N. Wang, "A stronger baseline for seismic facies classification with less data," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, May 2022, Art. no. 5914910.

[52] M. Q. Nasim, T. Maiti, A. Srivastava, T. Singh, and J. Mei, "Seismic facies analysis: A deep domain adaptation approach," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, Feb. 2022, Art. no. 4508116.

[53] D. Kim and J. Byun, "Selection of augmented data for overcoming the imbalance problem in facies classification," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, Aug. 2022, Art. no. 8019405.

[54] D. Zhu, J. Cui, Y. Li, Z. Wan, and L. Li, "Adaptive Gaussian mixture model and convolution autoencoder clustering for unsupervised seismic waveform analysis," in *Interpretation*, vol. 10, no. 1, pp. T181–T193, Jan. 2022.

[55] J. Li et al., "Unsupervised contrastive learning for seismic facies characterization," *Geophysics*, vol. 88, no. 1, pp. WA81–WA89, Jan. 2023.

[56] K. Li, S. Chen, and G. Hu, "Seismic labeled data expansion using variational autoencoders," *Artif. Intell. Geosci.*, vol. 1, pp. 24–30, Dec. 2020.

[57] Q. Feng, Y. Li, and H. Wang, "Intelligent random noise modeling by the improved variational autoencoding method and its application to data augmentation," *Geophysics*, vol. 86, no. 1, pp. T19–T31, Jan. 2021.

[58] W. A. McAliley and Y. Li, "Stochastic inversion of geophysical data by a conditional variational autoencoder," *Geophysics*, vol. 89, no. 1, pp. 1–65, Aug. 2023.

[59] A. Tewari, M. J. Giering, and A. Raghunathan, "Parametric characterization of multimodal distributions with non-Gaussian modes," in *Proc. IEEE Int. Conf. Data Mining Workshops*, 2011, pp. 286–292.

[60] N. Dilokthanakul et al., "Deep unsupervised clustering with Gaussian mixture variational autoencoders," 2016, *arXiv:1611.02648*.

[61] Y. Su et al., "Detecting outlier machine instances through Gaussian mixture variational autoencoder with one dimensional CNN," *IEEE Trans. Comput.*, vol. 71, no. 4, pp. 892–905, Apr. 2022.

[62] A. Cao, Y. Luo, and D. Klabjan, "Open-set recognition with Gaussian mixture variational autoencoders," in *Proc. Int. Joint Conf. Artif. Intell.*, 2021, pp. 6877–6884.

[63] J. I. Marden, "Positions and QQ plots," *Statist. Sci.*, vol. 19, no. 4, pp. 606–614, Nov. 2004.

[64] J. A. Figueroa, "Semi-supervised learning using deep generative models and auxiliary tasks," in *Proc. Int. Conf. Neural Inf. Process. Syst.*, 2019, pp. 1–9.

[65] A. Abdulaziz, J. Zhou, A. Di Fulvio, Y. Altmann, and S. McLaughlin, "Semi-supervised Gaussian mixture variational autoencoder for pulse shape discrimination," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, 2022, pp. 3538–3542.

[66] Y. Alaudah, P. Michałowicz, M. Alfarraj, and G. AlRegib, "A machine-learning benchmark for facies classification," in *Interpretation*, vol. 7, no. 3, pp. SE175–SE187, Jul. 2019.

[67] G. Ping, J. Chen, T. Pan, and J. Pan, "Degradation feature extraction using multi-source monitoring data via logarithmic normal distribution based variational auto-encoder," *Comput. Ind.*, vol. 109, pp. 72–82, Aug. 2019.
[68] C. Song, Z. Li, Z. Liu, and G. Hu, "Prestack reflection pattern based seismic facies analysis," in *Proc. SEG Annu. Meeting*, Oct. 2015. pp. 1633–1637.
[69] T. Kohonen, "The self-organizing map," *Proc. IEEE*, vol. 78, no. 9, pp. 1464–1480, Sep. 1990.
[70] L. Van der Maaten and G. Hinton, "Visualizing data using t-SNE," *J. Mach. Learn. Res.*, vol. 9, no. 11, pp. 2579–2605, Nov. 2008.
[71] B. M. Devassy, S. George, and P. Nussbaum, "Unsupervised clustering of hyperspectral paper data using t-SNE," *J. Imag.*, vol. 6, no. 5, May 2020, Art. no. 29.
[72] Y. Wang, A. E. M. Mendez, M. Cartwright, and J. P. Bello, "Active learning for efficient audio annotation and classification with a large amount of unlabeled data," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, 2019, pp. 880–884.
[73] Y. Aytar, C. Vondrick, and A. Torralba, "SoundNet: Learning sound representations from unlabeled video," in *Proc. Int. Conf. Neural Inf. Process. Syst.*, 2016, pp. 892–900.
[74] P. A. Estévez, M. Tesmer, C. A. Perez, and J. M. Zurada, "Normalized mutual information feature selection," *IEEE Trans. Neural Netw.*, vol. 20, no. 2, pp. 189–201, Feb. 2009.
[75] D. L. Davies and D. W. Bouldin, "A cluster separation measure," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-1, no. 2, pp. 224–227, Apr. 1979.

Haowei Hua received the bachelor's degree in electronic information engineering from Hunan University (HNU), Changsha, China, in 2021. He is currently working toward the master's degree in communication engineering with the University of Electronic Science and Technology of China (UESTC), Chengdu, China.

His research interest covers deep learning and low-rank modeling.

Feng Qian (Member, IEEE) received the B.S. and M.S. degrees in optoelectronic science and technology from Sichuan University (SCU), Chengdu, China, in 2001 and 2004, respectively, and the Ph.D. degree in communication and information systems from the University of Electronic Science and Technology of China (UESTC), Chengdu, China, in 2008.

Since 2009, he has been with the School of Information and Communication Engineering and also with the Center for Information Geoscience, University of Electronic Science and Technology of China, Chengdu, China, where he is currently an Associate Professor. From 2014 to 2015, he was also a Visiting Scholar with Columbia University, New York, NY, USA. His current research interests include tensor theory, nonconvex optimization, deep learning, big data analysis, and their applications in seismic data processing.

Gulan Zhang (Member, IEEE) was born in Hubei, China. He received the B.S., M.S., and Ph.D. degrees in geophysical prospecting and information technology from the Chengdu University of Technology, Chengdu, China, in 2004, 2008, and 2016, respectively.

From 2012 to 2016, he was the Geophysicist and the Director of the Institute of Development Seismic (Unconventional oil and gas geophysics), Bureau of Geophysical Prospecting Inc., (BGP), China National Petroleum Corporation (CNPC), Zhuozhou, China. He is currently a Professor with the State Key Laboratory of Oil and Gas Reservoir Geology and Exploitation, School of Geoscience and Technology, and the Dean of the Institute of Geophysics, Southwest Petroleum University, Chengdu, China. His research interests include reservoir characteristics, geophysical inversion, deep learning of seismic data, weak signal detection algorithms, time-frequency analysis algorithms, hydrocarbon detection methods, high-speed-train seismology, and other vibration signal processing algorithms.

Dr. Zhang is currently an Editorial Board Member/Associate Editor of *Petroleum Science, Journal of Geophysics and Engineering, Journal of Applied Geophysics, Acta Geophysica, and Applied Geophysics*.

Yuehua Yue received the bachelor's degree in information and computing science from the North China University of Water Resources and Electric Power, Zhengzhou, China, in 2020 and the master's degree in electronic information from the University of Electronic Science and Technology of China (UESTC), Chengdu, China, in 2023.

He is currently an Engineer with the Satellite Communication Technology Institute, Nanjing Panda Handa Technology Company, Ltd., Nanjing, China. His research interest covers deep learning and prestack seismic data analysis.