



Research paper

MAMCL: Multi-attributes Masking Contrastive Learning for explainable seismic facies analysisLong Han ^a, Xinning Wu ^{a,*}, Zhanxuan Hu ^b, Jintao Li ^a, Huijing Fang ^c^a School of Earth and Space Sciences, University of Science and Technology of China, Hefei, China^b School of Information Science and Technology, Yunnan Normal University, China^c Exploration Research Institute, Anhui Provincial Bureau of Coal Geology, China

ARTICLE INFO

Keywords:

Unsupervised seismic facies classification

Contrastive learning

Interpretability analysis

ABSTRACT

Seismic facies analysis is crucial in hydrocarbon exploration and development. Traditional machine learning approaches typically require manual selection of attributes and lack interpretability analysis. We propose an interpretable framework, multi-attribute masking contrastive learning (MAMCL), designed to adaptively select, explore and aggregate seismic attributes for seismic facies analysis. The MAMCL framework includes a depthwise CNN module for feature extraction and an iTransformer module for feature aggregation. Based on the assumption that different attributes computed on the same seismic sample imply common information associated with the same geologic facies, we formulate an unsupervised strategy of contrastive learning to pre-train the MAMCL framework for refining the attributes. This pre-training method encourages the network to extract and integrate highly correlated attribute features by enhancing the expression of commonalities within the same sample, and implicitly increase the distance between features of different categories by differentiating the expressions of different samples. Ultimately, these refined features only need to be input into a simple clustering algorithm, such as K-Means, to achieve seismic facies classification. MAMCL requires no labels or manual selection of attributes and can utilize the self-attention mechanism of iTransformer to compute adaptive attribute weights, facilitating interpretability analysis. We applied MAMCL framework to both unlogged turbidite channel systems in Canterbury Basin, New Zealand, and logged Chengdu area in Bohai Bay Basin, China, achieving reliable classification results and providing interpretability analysis.

1. Introduction

Seismic facies are groups of different seismic reflections based on the amplitude, frequency, reflection, geometry and reflection continuity, which are different from those of adjacent groups (Roy and Marfurt, 2011). Studying seismic facies enables researchers to identify key stratigraphic and reservoir features from seismic data without well data guidance, which is crucial for hydrocarbon exploration (Marquín, 2014). Traditional seismic facies classification requires experienced interpreters to manually delineate facies boundaries, which is a daunting challenge in areas with complex subsurface structures. In addition, manual interpretation of seismic facies is highly subjective and different interpreters may produce varying results (Tolstaya and Egorov, 2022).

A seismic attribute is any measure of seismic data that enhances visual representation or quantifies features of interpretative interest. Seismic attributes can elucidate features, relationships, and patterns in seismic data that unnoticed (Chopra and Marfurt, 2007). Combining multiple appropriate seismic attributes for seismic facies classification

has become a mainstream method of seismic facies interpretation. Some researchers employ machine learning techniques for automated seismic facies classification, such as the K-means algorithm (Priezzhev and Manral, 2012; Arianfar et al., 2007), SOM (de Matos et al., 2007; Zhao et al., 2017; Du et al., 2015; Kourki and Ali Riahi, 2014), GTM (Chopra and Marfurt, 2014; Roy et al., 2014), GMM (Wallet and Hardisty, 2019), random forest (Kim et al., 2018; Wrona et al., 2018) and their combination methods (da Silva et al., 2023). Although these methods achieve automation to some extent, they struggle to extract and integrate features from multi-attributes, especially when the number of attributes is large. They often require substantial memory and are time-consuming. Researchers often combine the PCA dimensionality reduction algorithm with these methods to improve efficiency (Sabeti and Nadjar, 2011), but this approach can lead to issues of information loss.

Recently, deep learning has advanced rapidly, achieving significant breakthroughs in computer vision (Dosovitskiy et al., 2020; Kirillov

* Corresponding author.

E-mail address: xinmwu@ustc.edu.cn (X. Wu).

et al., 2023; Oquab et al., 2023) and natural language processing (Devlin et al., 2018; Achiam et al., 2023). The Transformer (Vaswani et al., 2017) is a key component of these methods, featuring a large receptive field and robust generalization capabilities, particularly its attention mechanism which flexibly focuses on various parts of the data while providing a basis for interpretability. Its variant, the iTransformer (Liu et al., 2024), explores the use of the attention mechanism to elucidate correlations among different variables, achieving state-of-the-art performance in the domain of time series prediction. The feature extraction and representation capabilities of deep learning have garnered significant favor among many geophysicists. Consequently, some researchers have started applying deep learning to seismic facies analysis, effectively resolving issues related to multi-attribute data feature extraction and dimensionality reduction of high-dimensional data. Zhang et al. (2021) have employed supervised CNN networks for seismic facies classification. However, in most cases, due to the scarcity of well logs, some fields have few or no labels, making supervised networks unsuitable. Consequently, unsupervised methods have attracted significant attention. He et al. (2018) utilized sparse autoencoders to extract features from seismic data and then applied k-means clustering for seismic facies cluster. Zhou and Chen (2022) employed a recurrent autoencoder model to extract temporal features from seismic data and similarly used k-means clustering for seismic facies analysis. Recently, contrastive learning has emerged as a popular paradigm in unsupervised learning (Chen et al., 2020; Radford et al., 2021). Benefiting from abundant unlabeled data, many researchers have adopted this method to train foundational models (Dong et al., 2023; BehnamGhader et al., 2024). Contrastive learning differentiates between similar and dissimilar data points to learn data representations, capturing the fundamental structures and relationships among different data points. Li et al. (2023) applied contrastive learning to extract features from seismic data and attributes, and also conducted seismic facies classification within the network.

However, both machine learning and deep learning approaches face two key issues. The first issue is that attributes used for seismic facies classification require manual selection. Barnes and Laughlin (2002) maintain that the selection of input attributes exerts a greater influence on facies diagrams than the unsupervised learning algorithms used for classification. Many interpreters analyze potential seismic facies by observing rendered images of seismic attributes and select attributes that emphasize these facies based on their experience (Zhao et al., 2017). This approach heavily relies on the interpreters' expertise and is both time-consuming and labor-intensive. Moreover, in areas with complex seismic facies types, this method is prone to subjective judgments. Furthermore, many methods assume that different attributes contribute equally to seismic facies; however, the importance of these attributes can vary with the characteristics of the data (or seismic facies). Zhao et al. (2018) explores various combinations of multiple attributes, using SOM (Self-Organizing Maps) to cluster repeatedly and determine the weights of different attributes, which are then selected in conjunction with expert knowledge. Qi et al. (2020) employs the GMM (Gaussian Mixture Model) clustering algorithm to perform an exhaustive search of candidate attributes to determine the optimal number and combination of attributes to use. However, these approaches are not sufficiently automated and objective. The second issue is that the process of seismic facies classification is difficult to explain, especially when using deep learning methods. Although Lubo-Robles et al. (2022) used SHAP values to interpret seismic facies classification, this method is applicable primarily to labeled data and specific machine learning algorithms.

To tackle these challenges, we introduce the MAMCL framework for explainable seismic facies analysis, employing a workflow of unsupervised pre-training followed by simple clustering. MAMCL primarily consists of a Depthwise CNN1D module for data feature extraction, an iTransformer module for feature aggregation, and a Projector for mapping representations to the space where contrastive loss is applied. We utilize single-trace seismic data along with its multiple attributes

as multi-attribute data training samples, which do not require pre-selection of attributes. Then, the model is pre-trained using an unsupervised contrastive learning approach that involves randomly masking input attributes. This unsupervised pre-training enables the model to effectively uncover commonalities and differences among multiple attributes. After pre-training, the CNN followed by the iTransformer module represents the input samples (characterized by multiple attributes) in a latent space where samples belonging to similar facies are grouped together while those belonging to different facies are separated. Therefore, in this space, we can achieve accurate seismic facies classification by simply performing K-Means clustering. Additionally, we utilize the self-attention mechanism of iTransformer to output data-driven importance scores for different attributes, thereby facilitating interpretability analysis of the algorithm. To the best of our knowledge, MAMCL is the first method that enables interpretability analysis of deep learning in seismic facies analysis. We have successfully applied this method in multiple field datasets, achieving classifications consistent with geologic understanding and well-log ground truth.

This article will first introduce the network architecture of MAMCL. It then illustrates the workflow with a case study of a turbidite system, including pre-training and clustering. Subsequently, MAMCL is applied to another field equipped with well logging. Finally, we summarize our work.

2. Method

In this section, we provide a detailed introduction to our proposed MAMCL framework and its key components, followed by a case study of a turbidite system to illustrate the pre-training and clustering processes.

2.1. Overall architecture

2.1.1. Input data extraction

Seismic amplitude data and its associated attributes are generally stored in the format of three-dimensional cubes. However, for the analysis of seismic facies, the focus is typically on the distribution of facies within a layer of interest. Therefore, it is only necessary to extract data centered at the target stratigraphic horizon. As illustrated in left pink box in Fig. 1a, to fully utilize the temporal characteristics of seismic data, a single trace is extracted from the seismic data volume with a time window L centered at the target horizon. Laterally in inline and crossline directions, we select one for every four traces to save computational cost but still maintain nearly the same performance as using all data. Employing the same method for other attribute volumes, a training sample at some spatial grid point x_i is derived, denoted as $x_i^{C \times L}$, where C represents the number of attributes. Ultimately, we obtain the complete set of samples for the entire stratigraphic horizon, denoted as $\{x_i^{C \times L}\}_N$, where N represents the number of samples. Chen et al. (2020) demonstrated through experiments that unsupervised contrastive learning benefits more from stronger data augmentation compared to supervised learning. In time-series contrastive learning tasks, Eldele et al. (2021) achieved strong data augmentation by adding permutation and noise, resulting in effective representations. Since seismic waveforms are similar to time-series data, we also apply a small amount of noise to the input in each iteration to ensure training stability and robustness to noise.

2.1.2. Depthwise CNN1D for feature extraction

The Depthwise CNN1D module is designed to independently extract fine-grained features from multi-attribute inputs without merging them. We use Depthwise CNN1D instead of traditional CNN1D to prevent the mixing of information among different input attributes, ensuring their independence. This independence is crucial for the next step of attribute-wise masking in contrastive learning. Specifically, for the samples $\{x_i^{C \times L}\}_N$, we employ $M = K \times C$ convolutional kernels for convolution (where K is an integer), and apply appropriate

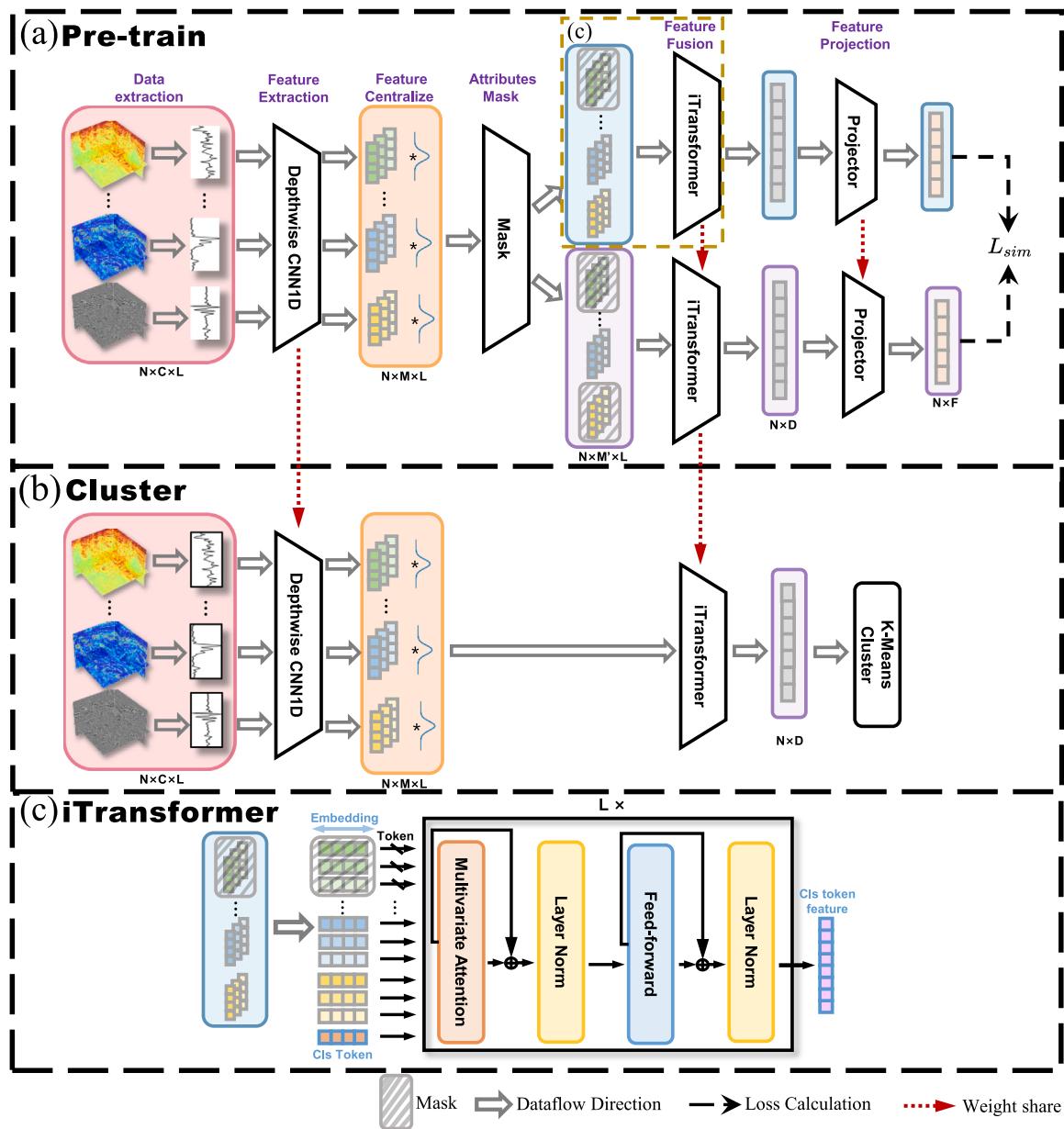


Fig. 1. The overall architecture of MAMCL. (a) The pipeline of unsupervised pre-training with the strategy of masked contrastive learning. (b) The pipeline of seismic facies clustering after pre-training. (c) Detailed structure of iTransformer (yellow box in (a)).

zero-padding to maintain the sequence length unchanged. The output features are then obtained as $\{u_i^{M \times L}\}_N$, where:

$$\{u_i^{M \times L}\}_N = \text{DepthwiseCNN1D}(\{x_i^{C \times L}\}_N) \quad (1)$$

2.1.3. Feature centerization

Since the input seismic attribute sequence is used to characterize the seismic facies category of its central point location, the attribute information at the central location should be given higher weight in seismic facies analysis. We emphasize the information at the center of the feature sequence by multiplying each feature vector u_i extracted by the DCNN1D module with a Gaussian window (as denoted by the orange box in Fig. 1a). Although this Gaussian weighting can be directly applied to the initial input seismic attribute sequence, we have found that applying the Gaussian window to the subsequently extracted feature vectors has a more significant impact on the final seismic facies classification. Specifically, we construct a Gaussian window defined by:

$$G(r) = e^{\frac{-2\sigma^2}{r^2}} \quad (2)$$

where $G(\cdot)$ is the amplitude of the Gaussian function, r denotes the distance from the center of the window, and σ is the standard deviation, which determines the width of the window. The outputs of the Depthwise CNN1D across the time dimension are then multiplied by this Gaussian window, resulting in centralized feature outputs $\{s_i^{M \times L}\}_N$:

$$\{s_i^{M \times L}\}_N = \{u_i^{M \times L}\}_N \times G(r) \quad (3)$$

2.1.4. Mask

The Mask module is pivotal for the operation of our contrastive learning approach. Its function is to generate augmented positive sample pairs for subsequent contrastive learning. Specifically, we mask features corresponding to different attributes with a probability a , thereby obtaining augmented samples that lack varying attributes.

$$\{\tilde{s}_i^{M' \times L}\}_N = \text{MASK}(\{s_i^{M \times L}\}_N) \quad (4)$$

where $M' = K' \times M$, K' is integer. Similarly, we can obtain another sample $\{\tilde{s}_j^{M'' \times L}\}_N = \text{MASK}(\{s_i^{M \times L}\}_N)$. It is important to note that in the depthwise CNN, the number of channels is increased to a multiple of the number of attributes. Therefore, during the masking process, features generated from the same attribute should be considered as a cohesive unit to be masked out or not in this mask module.

2.1.5. iTransformer for feature fusion

The role of the iTransformer is to fuse all high-dimensional features using the self-attention mechanism to produce low-dimensional features. The distinction between it and the traditional Transformer lies in the fact that the traditional Transformer splits attributes along the temporal dimension into timestamps, treating each timestamp as a token that encompasses all attribute features within that timestamp. It learns attention across the temporal dimension, thereby overlooking the intrinsic correlations among attributes. In our seismic facies analysis problem, the temporal length is generally short, and emphasis should be placed on the information at the center position, thus negating the need to learn attention across the temporal dimension. Moreover, different attributes exhibit similar responses within the same type of seismic facies, making the correlation between attributes a critical aspect of seismic facies analysis. Furthermore, different attributes measure seismic facies from various perspectives, presenting distinct characteristics. Aggregating attributes after splitting them based on timestamps would compromise their expressive capabilities. Therefore, the traditional transformer is not suitable for the task of seismic facies analysis. Based on this, we employ the iTransformer network, as depicted in Fig. 1b. Simply put, each channel (i.e., features expanded by depthwise CNN1D corresponding to attributes) is treated as an entire unit and input as an independent token into the vanilla Transformer. The advantages of this approach are manifold: (a) By treating each attribute as a whole, it fully extracts and explores the correlations between attributes. (b) As the time length is generally short, this reduces the computational load. (c) The outcomes are not influenced by the positions of attribute tokens, aligning with the Transformer's assumption of permutation invariance (Engel et al., 2021). (d) Through the self-attention mechanism, the weights of different attributes can be calculated, enhancing the model's interpretability.

Specifically, the masked features $\{\tilde{s}_i^{M' \times L}\}_N$ are split along the channel dimension into M' tokens. A randomly generated $\text{class token} \in R^L$ is defined and concatenated after the M' tokens to aggregate all attribute features. The $M' + 1$ tokens are input into the iTransformer, where linear projections are used to obtain the queries, keys, and values $\mathbf{Q}, \mathbf{K}, \mathbf{V} \in R^{(M'+1) \times D}$, with D being the projection dimension. The representations q_i, k_j serve as the specific query and key for an (attribute) token. The self-attention scores are calculated as $A_{ij} = \text{Softmax}(\mathbf{Q}\mathbf{K}^\top)_{ij} \propto q_i^\top k_j$, and then multiplied with \mathbf{V} to compute the outputs $\text{Output}_i = \sum_j A_{ij} \mathbf{V}_j$, where $\text{Softmax}(\mathbf{z})_i = \frac{e^{\tilde{z}_i}}{\sum_{j=1}^K e^{\tilde{z}_j}}$. Finally,

we take the output corresponding to the *class token* as the aggregated feature $\{\tilde{z}_i^D\}_N = \text{Output}_t$ (where t is the position corresponding to the *class token*). The aggregated features \tilde{z}_j^D of $\{\tilde{s}_j^{M'' \times L}\}$ are derived in the same way. Since each token is pre-normalized along its feature dimension, the attention scores can reveal the weights of the attributes to some extent. Attributes with higher weights are more heavily weighted in subsequent interactions with the values V . This weighting method is data-driven, meaning that different samples are assigned different attribute weights that are suitable for each sample, thereby providing flexibility. Based on this intuition, the proposed mechanism is considered to offer better interpretability for multi-attribute seismic facies analysis problems.

2.1.6. Projector

The role of the Projector is to map features into a contrastive space, facilitating the application of contrastive learning loss. We use an MLP with one hidden layer to obtain $\{\tilde{r}_i^F\}_N = \text{Projector}(\{\tilde{z}_i^D\}_N) =$

$W_2\sigma(W_1\tilde{z}_i)$ and $\{\tilde{r}_j^F\}_N = \text{Projector}(\{\tilde{z}_j^D\}_N) = W_2\sigma(W_1\tilde{z}_j)$, where σ represents the ReLU nonlinearity, and W_1, W_2 represent linear layer. SimCLR (Chen et al., 2020) has demonstrated that a nonlinear projection head enhances the quality of representations from its preceding layers, underscoring the importance of the projector.

2.2. Pre-train

2.2.1. Train data

We illustrate our workflow with an example of seismic facies analysis in a New Zealand field. The work area is located in the Canterbury Basin offshore New Zealand, a turbidite system situated in the transitional zone between the continental slope and uplift. This area was deposited within a single tectonically driven transgressive-regressive cycle, filling the Cretaceous and Tertiary paleo-environments with corals and turbidites. The New Zealand Petroleum and Minerals Company conducted three-dimensional seismic exploration in the area outlined in black in Fig. 2a and has made the data publicly available. Research has confirmed that this region contains rich channel systems, particularly in the area marked by the red frame. Zhao et al. (2016) interpreted the Miocene turbidite systems through the phantom strata shown in Fig. 2b. White arrows indicate two channels potentially filled with mud, which converge downstream to form lobate features. Blue arrows denote older, likely sand-filled channels that predate the mud-rich channels and are incised and overlain by them. Red arrows suggest potential slope fans, while black arrows point to meandering channel complexes, which are not clearly visible in the horizontal slices.

Based on manual interpretation and empirical judgment, Zhao et al. (2016) selected Coherent Energy, Peak Spectral Magnitude, Peak spectral Frequency, and Curvedness as input attributes. Li et al. (2023) and Wallet and Hardisty (2019) have followed the attributes used by Zhao et al. (2016). However, manually selecting attributes for seismic facies classification is subjective and inefficient. It cannot reasonably explain why these attributes are chosen or how to select the optimal ones. Different types of seismic facies classification often require different types of attributes to be selected. To solve these limitations, we input all possible attributes into our network, which automatically optimizes the selection of these attributes for seismic facies classification. The input attributes include seismic data, Coherent Energy, Crossline Dip, GLCM Homogeneity, Inline Dip, Curvedness, Shape Index, Peak Spectral Frequency, and Peak Spectral Magnitude. Treating seismic data as an attribute as well, there are a total of nine attributes. Fig. 3 illustrates their values of these attributes at the target horizon, highlighting their ability to characterize different features. Some are important for the target seismic facies, while others are related to the background and might be redundant. Before inputting them into the network, we first normalize all attribute data using the mean and standard deviation of each dataset, as follows:

$$X_{\text{standardized}} = \frac{X - \mu}{\sigma} \quad (5)$$

where σ and μ represent the standard deviation and mean of each attribute cube.

2.2.2. Loss function

Following the paradigm of contrastive learning, MAMCL employs a contrastive loss

$$L_{i,j} = -\log \left(\frac{\exp(\text{sim}(\tilde{r}_i, \tilde{r}_j)/\tau)}{\sum_{k=1}^{2N} 1_{k \neq i} \exp(\text{sim}(\tilde{r}_i, \tilde{r}_k)/\tau)} \right) \quad (6)$$

N is the number of samples in a batch, $\text{sim}(u, v)$ represents the similarity between u and v , calculated using cosine similarity. τ is a temperature parameter used to scale the similarity scores. $1_{k \neq i}$ is an indicator function that returns 1 if $k \neq i$, and 0 otherwise. We treat features r_i and r_j , generated by masking the same sample, as positive pairs, and consider all other samples in a batch as negative pairs. The contrastive

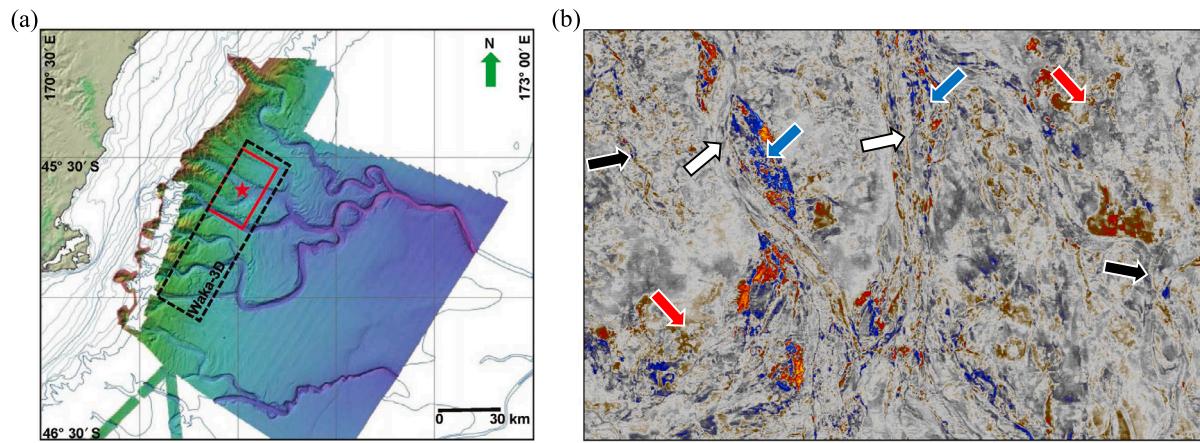


Fig. 2. (a) A map showing the locations of 3D seismic exploration conducted in the Canterbury Basin offshore New Zealand. The black rectangle indicates the extent of the Waka-3D survey, while the smaller red rectangle represents the area studied in this paper. The colors indicate the relative depth of the current seabed(modified from Mitchell and Neil (2012), Zhao et al. (2016)). (b) Seismic slices at the target horizon. Zhao et al. (2016) interpreted the regions indicated by white arrows as stacked channels, blue arrows as sandy channels, red arrows as slope fans, and black arrows as weaker meandering channels.

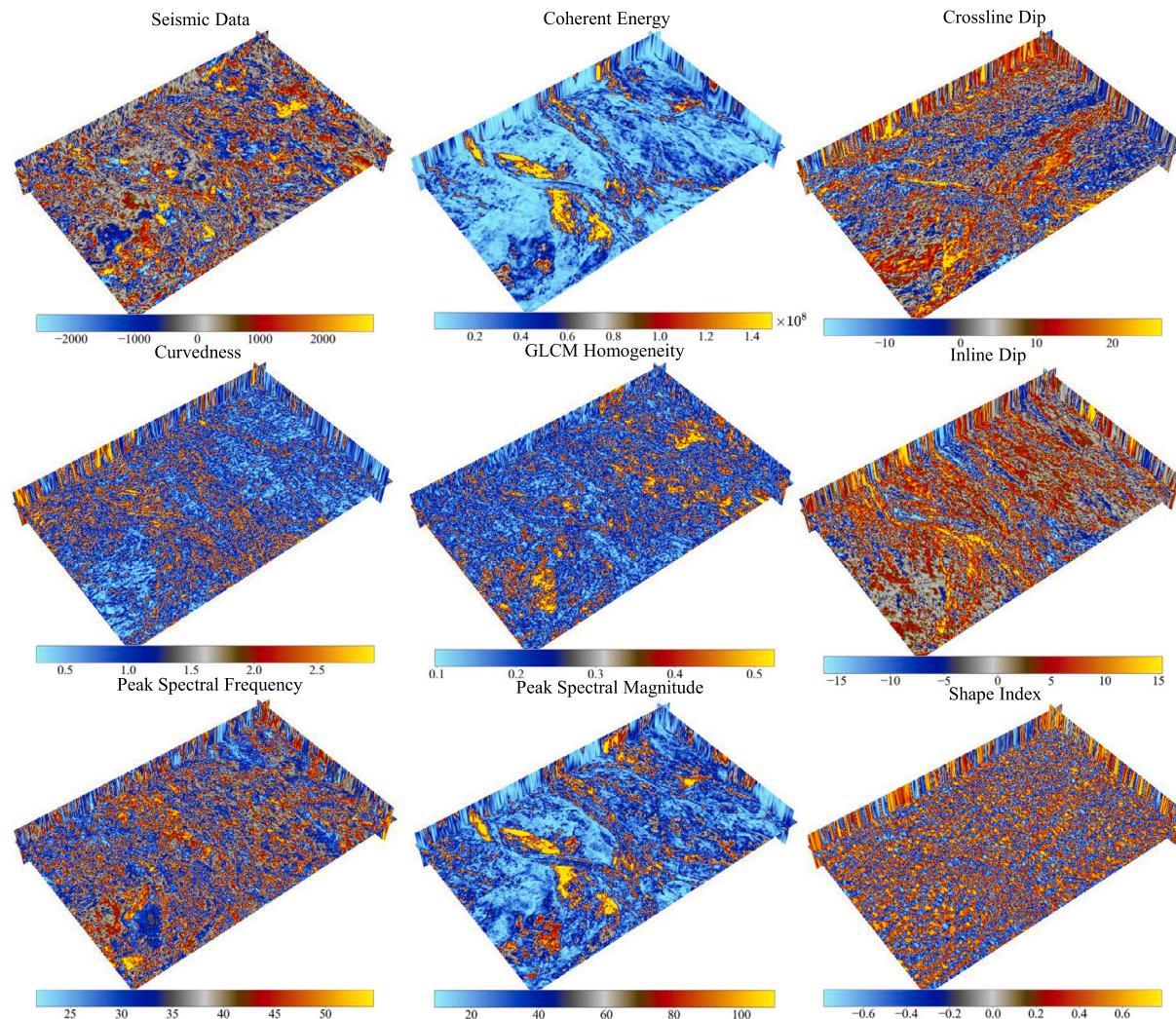


Fig. 3. Attribute slices at the target horizon. These attributes serve as input data for the network, including effective attributes conducive to seismic facies analysis and ineffective attributes filled with noise.

loss aims to maximize the similarity of positive pairs while minimizing the similarity of negative pairs.

This loss function has geological significance and rationality. When we have numerous attributes that are challenging for manual selection,

or when we are in the preliminary stages of exploration and the primary seismic facies of the unknown area are yet to be determined, we consider that the effective attributes calculated at the same position should exhibit common expressions. For example, work areas with channel facies at certain horizon exhibit higher amplitudes in attributes such as coherence energy and peak spectral amplitude. Therefore, when we randomly mask a portion of the attributes, to enhance the similarity of positive pairs, iTransformer will increase the weights of attributes that exhibit similar expressions and reduce the weights of ineffective attributes (e.g. those filled with noise). To enhance the dissimilarity of negative pairs, DepthwiseCNN1D and iTransformer networks will extract more effective features, which promotes uniform distribution of samples on a hypersphere. This process pulls samples of the same category closer and pushes samples of different categories apart. Consequently, this ensures that the clustering centers of different seismic facies are well separated in feature space.

2.2.3. Pre-training process with unsupervised contrastive learning

We set the vertical time or depth window size to 17, which includes the target horizon and 8 sampling points above and below it. We use one layer of Depthwise CNN1D, with the number of convolutional kernels being 8 times the number of input attributes, to adequately extract effective features. The masking rate is 0.5, meaning approximately 50% of the attributes are fed into the iTransformer. The linear layer dimension within the iTransformer is set to 64, thereby fusing and reducing the multi-attribute features down to 64 dimensions. The projector consists of two linear layers, further reducing the 64-dimensional fused features to 16 dimensions to compute the contrastive loss, thereby reducing computational costs. Contrastive learning benefits from larger batch sizes and longer training durations (Chen et al., 2020), so we set the batch size to 1024 and the number of training epochs to 500, using the LARS optimizer (You et al., 2017). LARS adjusts the learning rate for each layer based on the magnitude of the weights and gradients, aiding in faster model convergence and maintaining stability when using large batches.

The pre-training process of MAMCL is outlined as follows: As shown in Fig. 1a, we first input the multi-attribute seismic data into MAMCL, utilizing data cropping to generate training samples. After feature extraction and centralization by Depthwise CNN1D, the samples are input into the Mask module to generate two augmented views, blue and purple, known as positive sample pairs. These positive sample pairs are then fed into the same iTransformer and projector to aggregate features and project them into a contrastive space for computing contrastive loss.

As shown in Fig. 4a, we plotted the variations in self-attention weights for different attributes. Our network exhibited rapid learning during the first 100 epochs, characterized by swift changes in the self-attention weight curves (Fig. 4a). These changes gradually stabilized between 100 to 500 epochs, ultimately achieving good convergence. To demonstrate that our network can effectively separate samples in the feature space, we randomly select samples at each epoch and calculate the similarity matrix between them using cosine similarity. Values closer to 1 indicate higher similarity. We plot the distribution of similarity data in Fig. 4b, where the color gradient from dark to light indicates increasing epochs. At the beginning of training (epoch = 5), the data is distributed near 1, indicating high similarity among data points. As the number of training epochs increases, the feature distances among data points grow, and the similarity distribution gradually shifts to the left, indicating that the features are tending towards a uniform distribution. From a classification perspective, we compute the clustering centers for different categories and reduce them to two dimensions using UMAP (McInnes et al., 2018). The greater the distance between clustering centers, the better the clustering effect. Fig. 4c shows the variation in clustering centers for four categories across training epochs, with colors darkening to lighter indicating an increase in epochs. Initially, the clustering centers for the four categories are close together in space,

particularly between the first and second categories. After training, the four categories gradually separate and eventually distribute evenly across the space, indicating that the features extracted are effective for our followed seismic facies clustering.

2.3. Seismic facies clustering

2.3.1. Cluster result

After pre-training, the model has already acquired the capabilities to effectively extract and fuse features for seismic facies classification. Therefore, we only need to perform simple K-Means clustering on the output features $z \in R^D$ of the iTransformer to obtain the seismic facies classification results for each sample. In the stage of clustering show in Fig. 1b, we extract input data traces from the multiple seismic attribute volumes (centered at the target horizon) without applying random noise or interval sampling. The resulting samples are then fed into the pre-trained Depthwise CNN1D and iTransformer to obtain features for all positions at the target horizon. These features are input into K-Means to classify the seismic facies distribution on the target horizon.

To maintain consistency with the category settings in Fig. 2b, we set the number of classes as four as well in the K-means clustering. The classification results are shown in Fig. 5a. The category covered in sky blue (red arrows) likely represents channels filled with sand bodies. On the left is an older high-amplitude channel, while on the right are multilayered, scattered channels. The yellow category (black arrows) is likely mud-rich channels. The channel on the right stacks with sand body channels, whereas the one on the left cuts through the sand body channels and converges with the right channel. Blue arrows indicate meandering channel complexes. The left side is classified as potential sand body channels, and the right as mud-rich channels. The red category (white arrows) represents slope fans, distributed in the northeast and southwest areas. The navy blue category represents submarine plains, widely distributed across the seismic facies map.

2.3.2. Explainability analysis

The self-attention weights $A_{i,j}$ can reflect the weights of different attributes on the outcome, which facilitates interpretability analysis. We compute the average weights of all attributes across all samples and rank the attributes by weight value, as shown in Fig. 5b. Seismic data, Coherent Energy, Peak Spectral Magnitude, and GLCM Homogeneity show relatively high adaptive weights, while attributes like Curvedness and Peak Spectral Frequency appear lower weights. Attributes with high weights indicate significant geological implications for the sedimentary facies. In the target horizon, there are multistoried channels, slope fans, and levees among other depositional facies, which can be well described by these high-weight attributes. Coherent Energy is used to measure the continuity and regularity of reflection events in seismic data. They are particularly apparent for seismic facies that exhibit continuous and regular reflection events, helping to delineate continuous and uniform channels as well as the edges of different depositional features. Peak Spectral Magnitude can distinguish strata characterized by variations in frequency, making it particularly suitable for depositional facies with rapid sedimentation rates, significant changes in sediment grain size, or high fluid content. GLCM Homogeneity characterizes the texture features of geological structures and quantifies subtle variations within depositional environments. It is helpful in characterizing fine-grained depositional facies, such as deep-sea mud deposits and lake sediments. In Fig. 3, Seismic data, Coherent Energy, Peak Spectral Magnitude, and GLCM Homogeneity clearly display the morphology and contours of channels. However, the smaller the weight, the more the attribute map tends to show noise, which are attributes we do not need.

Our network can learn adaptive attribute weights for each input data, which differs from traditional clustering methods (e.g., K-Means, SOM, GMM) and regular neural networks (e.g., MLP). Since each attribute can express different geological features, the network focuses

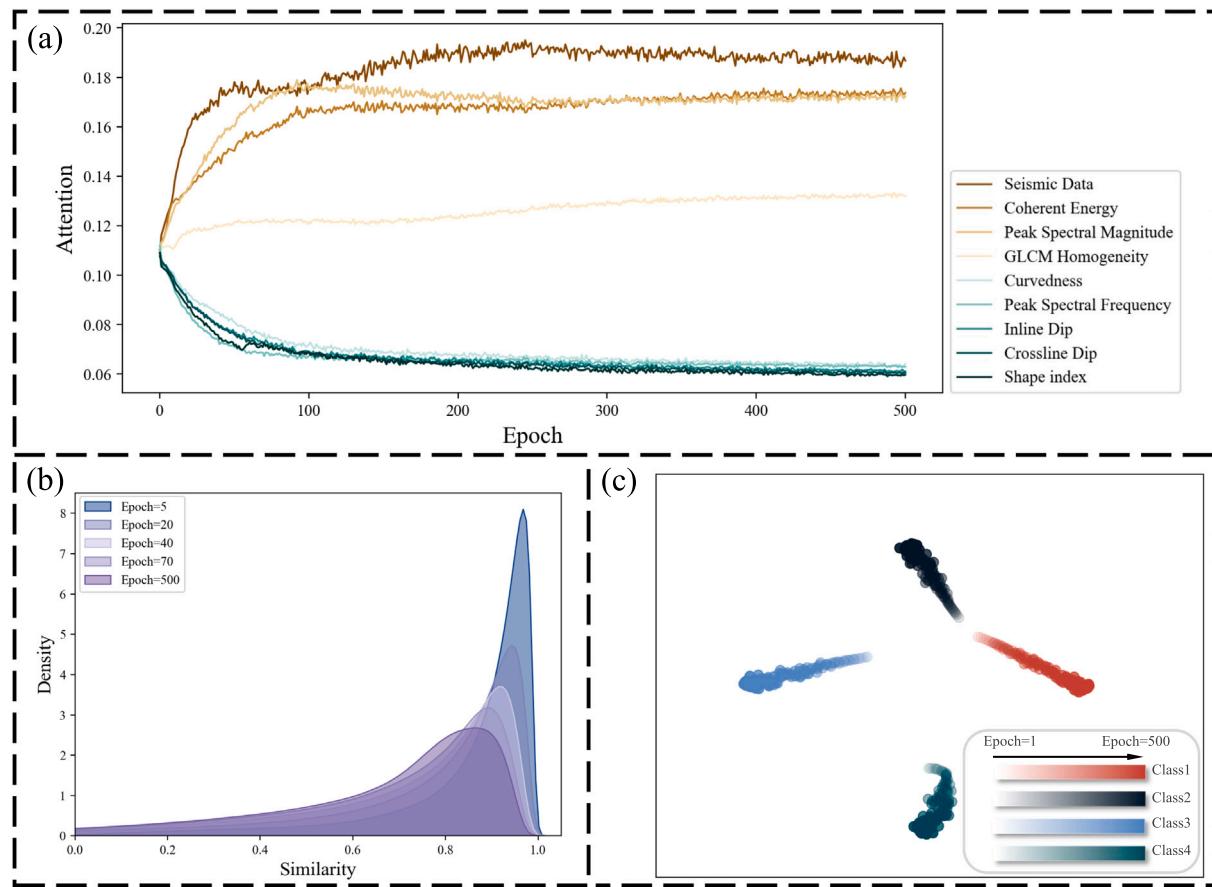


Fig. 4. Changes in (a) attribute weights, (b) distribution of similarity among samples during training and (c) clustering centers during training.

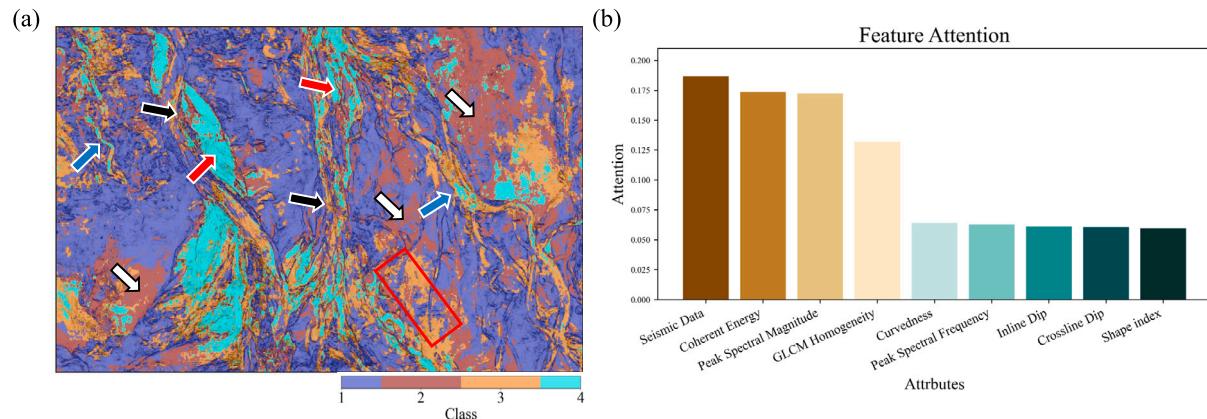


Fig. 5. (a) Results of seismic facies classification. Different seismic facies categories are represented by various colors. The Sobel-filter similarity (Chopra and Marfurt, 2014) is overlaid in black to show the boundaries. (b) Attribute weights for the entire work area, sorted by weight value.

on different regions within each attribute, helping us capture the target seismic facies. Fig. 6 shows the weight distribution for each attribute, with lighter colors indicating higher weights. In coherent energy, the network focuses more on the sandy channel area. In GLCM homogeneity, the network focuses on the submarine plain area. In Peak spectral magnitude, despite the sandy channel area having strong amplitude in the attribute slice, the network still focuses on the muddy channel and slope fan area. Additionally, in attributes filled with invalid features, such as shape index, the network can still extract features resembling channel edges. Thus, the network's self-attention mechanism can mimic the process of the human eye searching for target seismic facies in complex attribute maps, significantly reducing labor costs.

Our self-attention weights can effectively reflect seismic facies characteristics. For example, in coherent energy, regions with continuous and homogeneous seismic facies tend to receive higher weights, which share characteristics with sandy channels. In GLCM homogeneity, regions with complex texture features receive higher weights, potentially indicating the characteristics of submarine plains with complex sediment accumulation. Seismic facies are inferred sedimentary facies based on seismic data, and sedimentary facies are direct tools for geological analysis. Therefore, our attention-weighted maps can reflect geological features indirectly through their relation to seismic facies. However, due to inherent uncertainties in the data, such as noise and resolution limitations, as well as model randomness during the training,

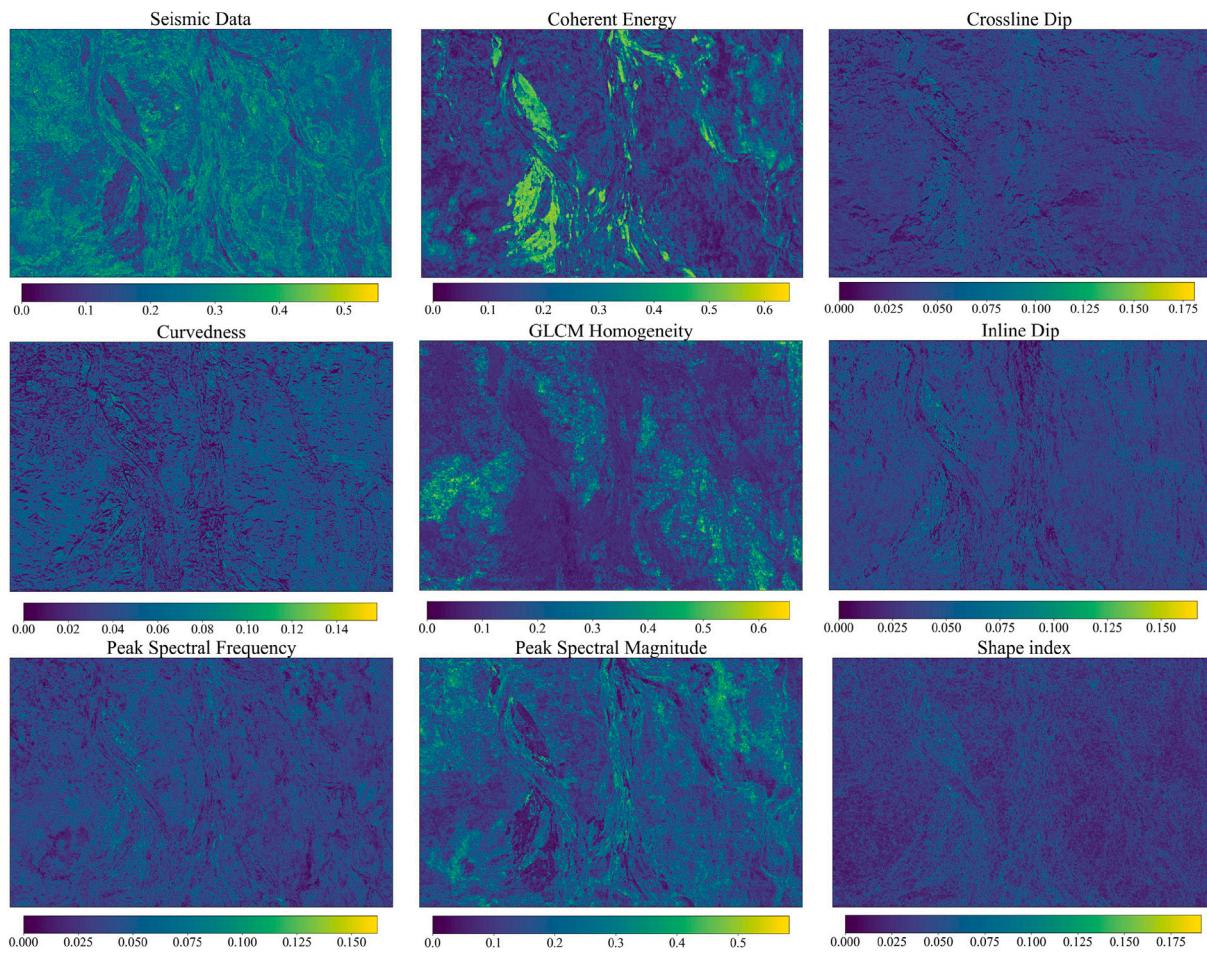


Fig. 6. Self-attention weight distribution for each attribute.

our interpretation results may not always fully align with geological interpretations.

2.3.3. Comparative analysis

Fig. 7 displays the classification results by Zhao et al. (2016), Wallet and Hardisty (2019) and Li et al. (2023). Zhao et al. (2016) used SOM to obtain the seismic facies map shown in Fig. 7a, plotting each point in the facies on a 2D color bar. Consequently, their method does not explicitly specify the exact number of facies classes, but rather visually indicates implicitly whether two points belong to the same class based on color differences. Wallet and Hardisty (2019) employed a combination of SOM and GMM for classification, dividing the seismic facies into seven classes (Fig. 7b). Li et al. (2023) developed a contrastive learning-based classification approach, constructing small cubes as input samples, using ResNet to extract image features, calculating contrastive losses between different attributes and seismic data, and using a classification head to output the classification results (Fig. 7c). Compared to their work, our method yields generally consistent results as shown in Fig. 7d. However, our approach is more competitive in identifying composite channels and certain minor channels (two blue arrows). Moreover, all of their methods require manual pre-selection of proper attributes and the selection of attributes is highly goal-oriented (Zhao et al., 2016). When the basic seismic facies types of the work area are unknown, redundant attributes can impact their computational results. On the contrary, our approach is unaffected by the selection of attributes. The iTransformer can adaptively select attributes, providing us with reliable results. Furthermore, the features output by the network can not only be utilized for seismic facies classification using K-Means but also combined with well logging data

for lithology prediction and other downstream tasks, demonstrating enhanced portability.

2.4. Ablation experiments

2.4.1. Network architecture

To demonstrate the effectiveness of our multi-attribute masking contrastive learning strategy and the significance of the iTransformer feature integrator, we designed three ablation experiments. Fig. 8a shows the results of K-Means clustering using all attributes directly. Due to the lack of feature extraction and attribute selection, the seismic facies classification results are noticeably noisy, and it fails to correctly distinguish between muddy channels, slope fans, and submarine plains. Fig. 8b employs a generic contrastive learning framework, but instead of using attribute masking for data augmentation, it applies varying degrees of noise to generate augmented data. The classification results remain heavily noisy. This is because, in the absence of attribute masking for generating augmented data, the network is unable to be well-trained for extracting effective attributes, leading to noisy attributes impacting the classification outcomes. Fig. 8c utilizes a contrastive learning framework based on attribute masking, but does not employ the iTransformer for feature integration. Instead, it directly averages the features extracted by a Depthwise CNN. Compared to Fig. 8b, the classification results show that noise has been almost entirely eliminated, and important seismic facies can be roughly identified, indicating that the network has the capability to extract effective attributes. However, due to the lack of iTransformer for data-driven attribute-weighted integration, some seismic facies are misclassified, such as the upper part of the sand body channel on the left being

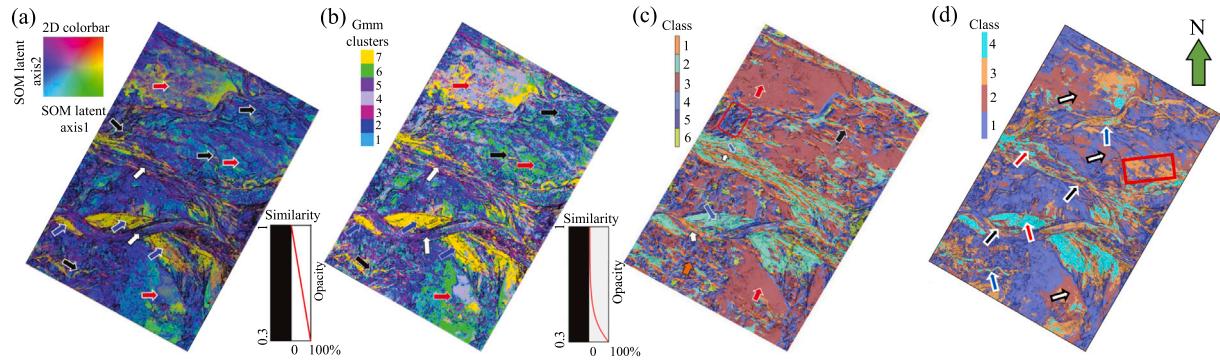


Fig. 7. Previous work using (a) SOM (Zhao et al., 2016), (b) GMM (Wallet and Hardisty, 2019) and (c) Contrastive Learning (Li et al., 2023). (d) Our work using MAMCL.

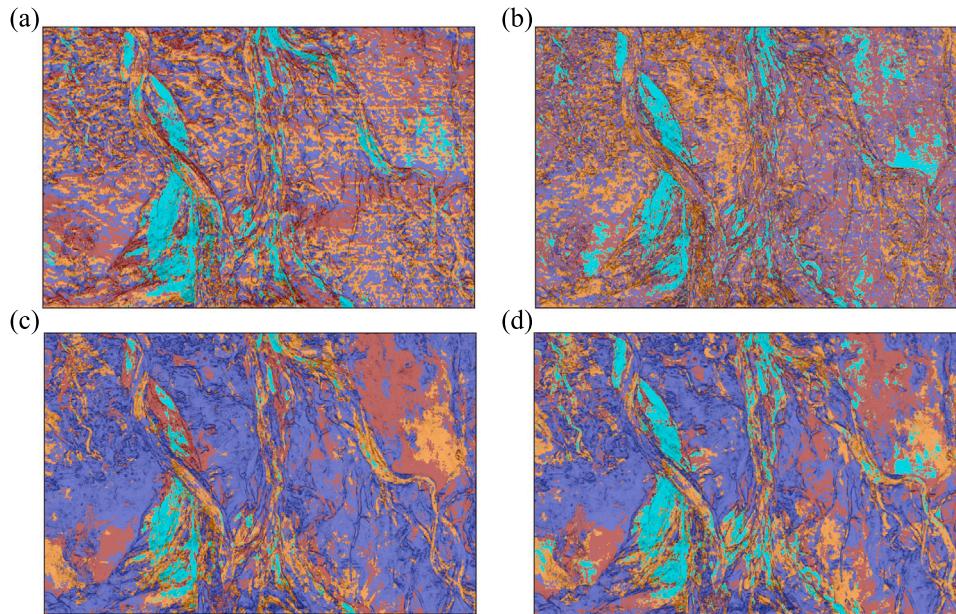


Fig. 8. Results of three ablation experiments compared to our results. (a) Directly employing K-Means clustering. (b) Enhancing data with various noises instead of using attribute masks. (c) Aggregating features using average aggregation rather than employing iTransformer. (d) Using all components.

classified as a slope fan, and the contiguous slope fans in the lower right corner being classified as a submarine plain. Fig. 8d displays the classification results using the complete components. It achieves the best classification performance in terms of both noise reduction and classification accuracy.

2.4.2. Window size

In addition to discussing the role of network architecture, we also investigated the impact of the information content on the classification results. We achieved this by varying σ of the Gaussian window. Larger σ values indicate that information from further away from the strata is input into the network, while smaller σ mean that only data closer to the horizon are inputted. Fig. 9a shows the results with $\sigma = 0.5$. Due to the smaller Gaussian window, the information input is insufficient, leading to the inability to differentiate certain seismic facies, such as the mixing of muddy channel facies with slope fan facies on the right. Fig. 9c depicts the results for $\sigma = 2.5$. With a larger Gaussian window, excessive information is input, potentially leading to interference from facies of other horizon in the classification, thereby resulting in numerous errors. Fig. 9b shows the results for a moderate size of Gaussian window with $\sigma = 1.5$, resulting in the best classification outcomes. In our tests on multiple examples, $\sigma = 1.5$ is generally a good choice to include important data features for the seismic facies classification on the target horizon.

2.4.3. Attribute test

Our method adaptively learns to select and weigh attributes, minimizing the dependency of the results on specific input attributes. This is particularly important for practical applications, as we do not always have full knowledge of the geological features in the study area. To evaluate the robustness of our method, we conducted experiments by removing different attributes to assess their impact on classification performance, with results shown in Fig. 10. Upon removing the most critical attributes, the model failed to produce accurate classifications due to the absence of key information reflecting seismic facies. When moderately important attributes were removed, minor misclassifications appeared in the red box area north of the channel, indicating that incorporating more effective attributes can enhance classification performance. However, when less important attributes were excluded, the classification results remained largely unchanged, as our method adaptively selects relevant attributes.

In practical applications, when we have an initial geological understanding of the study area, we can select potentially sensitive attributes accordingly. In cases where the geological conditions are not well understood, we can include as many attributes as possible. Since our method adaptively adjusts attribute weights, consistent and reliable classification results can still be achieved. Following the initial classification results, the attribute selection can be refined to obtain more detailed classification outcomes.

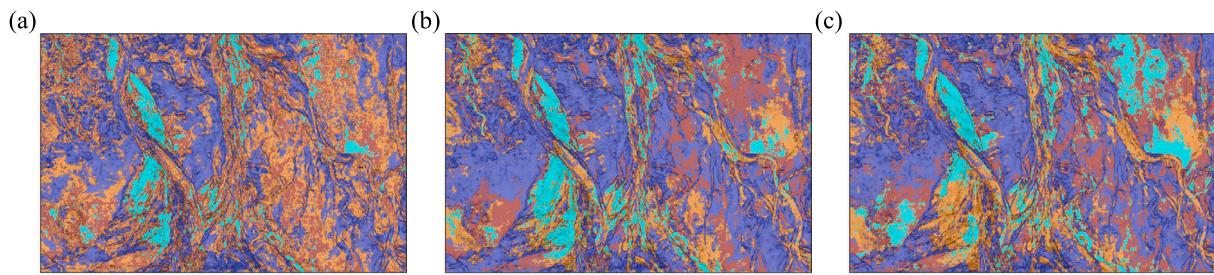


Fig. 9. Clustering results with different sizes of Gaussian window. (a) $\sigma = 0.5$, (b) $\sigma = 1.5$ (generally good choice), (c) $\sigma = 2.5$.

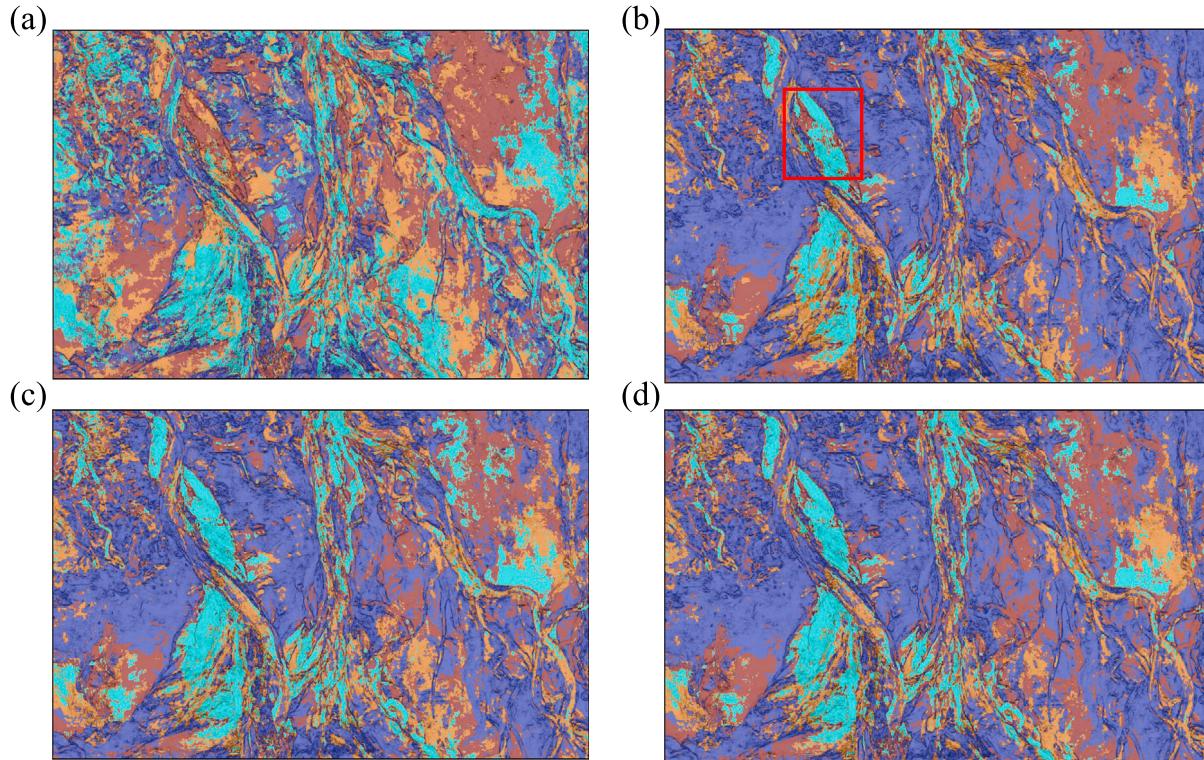


Fig. 10. Clustering results with the lack of different attributes. (a) Lack of seismic data and coherent energy, (b) Lack of peak spectral magnitude and GLCM homogeneity, (c) Lack of curvedness and peak spectral frequency, (d) Lack of crossline dip and shape index.

2.4.4. Training cost

We conducted the training using an NVIDIA Tesla V100 32 GB GPU. The memory usage was 7800 MB, which is dependent on the number of attributes. The training speed was 2100 samples per second.

Self-Supervised pre-training typically leverages large-scale datasets to capture richer representations and improve model generalization (El-Nouby et al., 2021). To explore the impact of training data volume on clustering performance and results, we extract seismic traces at different intervals for network training and subsequent clustering, and simultaneously record the runtime. The results are shown in Fig. 11.

We found that even when selecting one sample for every 64 traces, the clustering results remained stable, demonstrating the flexibility of our method with respect to training data volume. However, as the interval increases, clustering performance declines to some extent, as indicated in the red box. This occurs because fewer data are available during pre-training, causing the model to be less familiar with the full data distribution, which can reduce generalization when geological conditions are complex. Nevertheless, higher interval significantly reduce training time due to fewer training samples. Therefore, when selecting the sampling interval, it is essential to balance geological complexity, clustering performance, and computational cost.

3. Another example

3.1. Overview of the work field

The work area is located on the structural ridge of the eastern slope belt of the Chengdao in the Bohai Bay Basin. The target horizon is primarily developed in fluvial facies, making it a favorable location for oil and gas accumulation. Fig. 12 displays a seismic slice at the target horizon, which contains a sand-filled channel, a primary area for hydrocarbon accumulation. Thus, identifying the morphology and boundaries of this channel is crucial for hydrocarbon exploration. Within the work area, there are 12 well logs. At the target horizon, the majority of well logs indicate sandstone, which is concentrated in and around the channel, while wells showing mudstone and other lithologies are distributed outside the channel.

3.2. Train data

To unlock the potential of the model, we utilized a broader set of attribute data as inputs, including Seismic Data, Instantaneous Phase, Generalized Spectral Decomposition, Dip Deviation, Variance, Quadrature Amplitude, Sweetness, RMS Amplitude, Edge Detection, among

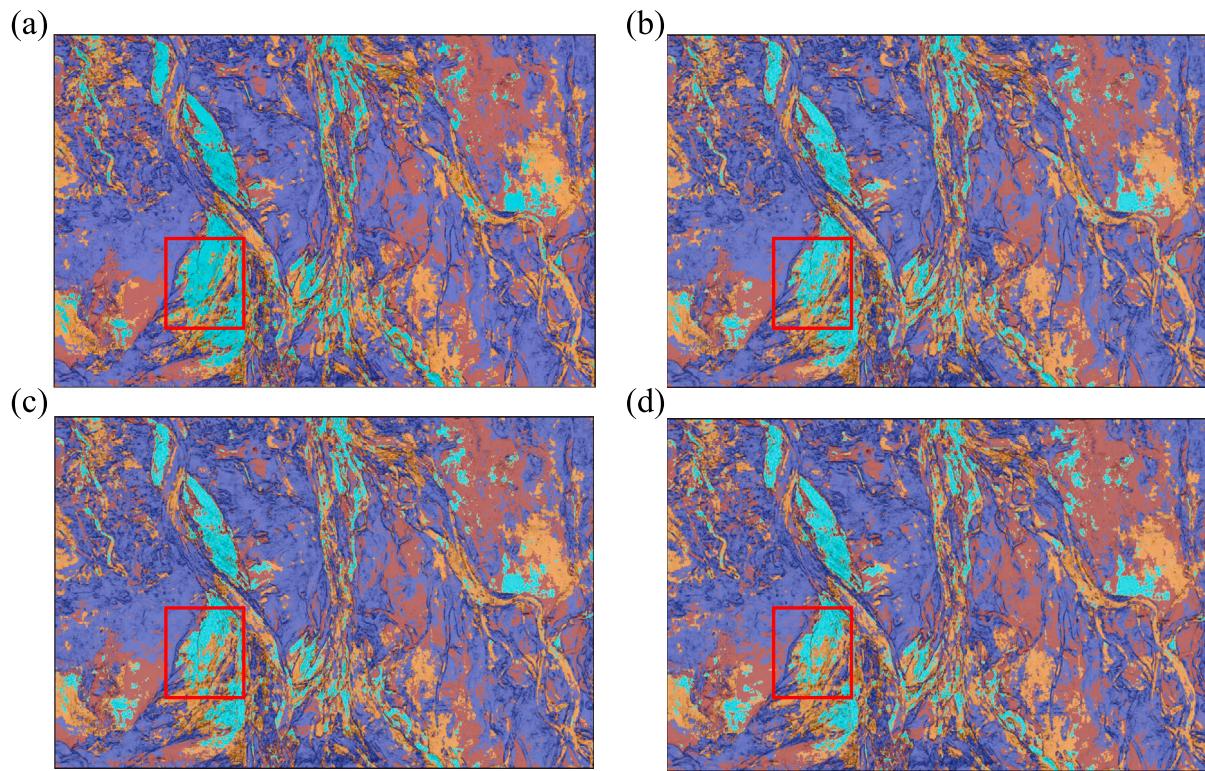


Fig. 11. Clustering results and runtime corresponding to models trained with seismic traces extracted at different intervals. (a) Select one for every 4 traces, runtime = 17.5 h, (b) Select one for every 16 traces, runtime = 4.1 h, (c) Select one for every 36 traces, runtime = 1.8 h, (d) Select one for every 64 traces, runtime = 0.9 h.

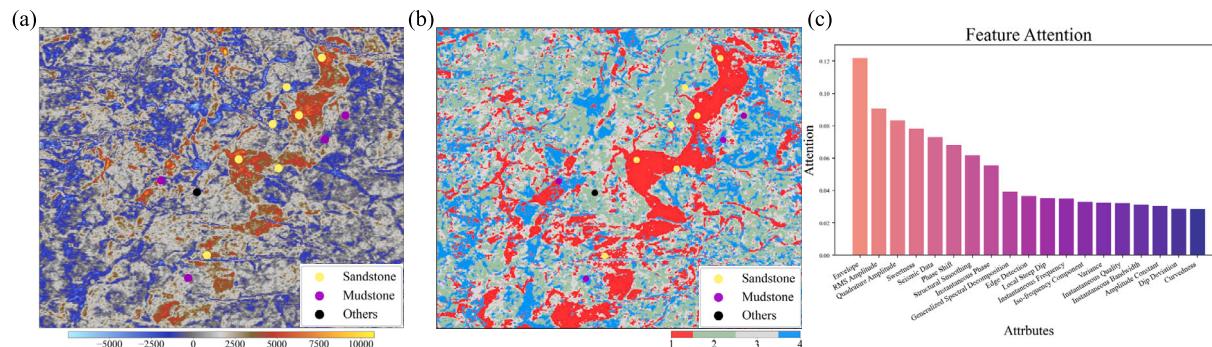


Fig. 12. (a) Seismic slice at target horizon. Solid circles represent well logs, with different colors indicating different lithologies. (b) Results of seismic facies classification. Different seismic facies categories are represented by various colors. Solid circles in different colors represent various lithologies of well logs. (c) Attribute weights for the entire work area, sorted by weight value.

a total of 19 attributes. Similar to Section 2.2.1, we also applied mean-variance normalization to the data.

3.3. Results and analysis

Given that the target horizon is located at the base of the channels, we extracted a time window of 13 sample points above horizon. All other parameters are consistent with those described in Section 2.2.3. The results of the seismic facies classification are shown in Fig. 12b. The first category (red) represents channel facies, including a main channel and several smaller, irregular channels. The remaining categories (blue, green, gray) may be floodplain facies and do not have the conditions for oil and gas storage. Even with an increased number of input attributes, our network achieves stable classification results, demonstrating its high robustness. Fig. 12c illustrates the distribution of weights across 19 attributes, with the Envelope attribute having the highest weight. This is because the channel facies is the main

seismic facies at the target horizon, characterized by a widespread and continuous distribution. Due to the strong amplitude, the channel is likely filled with sand bodies. Attributes such as RMS Amplitude, Quadrature Amplitude, and Sweetness, which are sensitivity to fluids in the rocks, also carry significant weights. This suggests that the study area is likely characterized by meandering river deposits, with potential fluid presence in the reservoirs. These findings are consistent with the insights discussed in Section 3.1.

When well log data are available, we can use these data to validate the results of the seismic facies classification. Since seismic facies themselves do not directly indicate geological significance, the rationality of the classification results can be demonstrated to some extent as long as the same lithology is categorized under the same seismic facies class. In Fig. 12b, sandstone wells (yellow) are primarily concentrated in the channel facies, suggesting that the channel may be sandy. Mudstone wells (purple) are clustered in the areas filled in blue, indicating that the blue-filled seismic facies may be mud-rich. Other lithologies

(black) are located in the light green areas. This indicates that our results are rational and accurate. These well log data facilitate linking seismic facies classification results with lithology, thereby providing convenience for subsequent oil and gas exploration and development.

4. Conclusion

In this paper, we propose a contrastive learning architecture for unsupervised seismic facies classification, termed MAMCL, which comprises two stages: pre-training and clustering. In the first stage, we employ attribute masking to obtain augmented samples, pre-training the network by maximizing the similarity within the same sample's augmented views and minimizing the similarity between different samples. This pre-training approach facilitates the extraction and integration of highly correlated attribute features, thereby eliminating the need for prior attribute selection. In the second stage, we use simple K-means clustering to achieve seismic facies classification results, and utilize the self-attention mechanism of the iTransformer to compute adaptive attribute weights, enabling interpretability analysis of the model. Finally, we applied MAMCL to the turbidite systems off the coast of New Zealand and a well-logged work area in the Chengdao region of the Bohai Bay Basin, China, achieving seismic facies classification and model interpretation. We designed four ablation experiments to demonstrate the efficacy of this method.

Our method still has some limitations. Firstly, there are some unreasonable classifications in our results, such as the slope fans at the base of the main channels in the turbidite system. Secondly, we are currently unable to impose artificial constraints on our model, so the training process may be uncontrollable. Thirdly, our input data consists of single-trace data, which prevents us from utilizing the spatial information. In the future, we plan to incorporate artificial constraints into the self-attention mechanism. Additionally, we aim to integrate spatial information into the input data to enhance the continuity of the classification results. Finally, we plan to employ more advanced clustering algorithms or co-training methods to improve clustering performance.

CRediT authorship contribution statement

Long Han: Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Xinming Wu:** Writing – review & editing, Supervision, Resources, Project administration, Methodology, Funding acquisition, Formal analysis, Conceptualization. **Zhanxuan Hu:** Software, Methodology, Formal analysis. **Jintao Li:** Visualization, Methodology, Investigation. **Huijing Fang:** Investigation, Formal analysis.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

This research was financially supported by National Science Foundation of China under grant no 42374127. We thank the USTC supercomputing center for providing computational resources for this work. We would like to thank New Zealand Petroleum and Minerals for generously providing the Waka-3D seismic survey to the public which can be accessed at <https://wiki.seg.org/wiki/Waka-3D>. We would like to thank Sinopec for providing the seismic data and well logging data for the Chengdao work area.

Code availability section

Name of the code/library: Seismic-facies
 Contact: Longhzzz@mail.ustc.edu.cn
 Hardware requirements: Python 3.8
 Program language: Python
 Software required: NVIDIA Tesla V100, GPU memory 24G.
 Program size: 31.8 M
 The source codes are available for downloading at the link: <https://github.com/Longhzzz/Seismic-facies>

References

- Achiam, J., Adler, S., Agarwal, S., Ahmad, L., Akkaya, I., Aleman, F.L., Almeida, D., Altenschmidt, J., Altman, S., Anadkat, S., et al., 2023. Gpt-4 technical report. arXiv preprint arXiv:2303.08774.
- Arianfar, A., Khedri, B., Haghghi, M., Golalzadeh, A., Poladzadeh, M., Mehdiipour, Z., 2007. Case history: Seismic facies analysis based on 3D multiattribute volume classification in Shadegan Oilfield–Asmari reservoir, Iran. In: SPE Reservoir Characterisation and Simulation Conference and Exhibition?. SPE, pp. SPE-111078.
- Barnes, A.E., Laughlin, K.J., 2002. Investigation of methods for unsupervised classification of seismic data. In: SEG Technical Program Expanded Abstracts 2002. Society of Exploration Geophysicists, pp. 2221–2224.
- BehnamGhader, P., Adlakha, V., Mosbach, M., Bahdanau, D., Chapados, N., Reddy, S., 2024. LLM2vec: Large language models are secretly powerful text encoders. arXiv preprint arXiv:2404.05961.
- Chen, T., Kornblith, S., Norouzi, M., Hinton, G., 2020. A simple framework for contrastive learning of visual representations. In: International Conference on Machine Learning. PMLR, pp. 1597–1607.
- Chopra, S., Marfurt, K.J., 2007. Seismic attributes for prospect identification and reservoir characterization. Society of Exploration Geophysicists and European Association of
- Chopra, S., Marfurt, K.J., 2014. Seismic facies analysis using generative topographic mapping. In: SEG Technical Program Expanded Abstracts 2014. Society of Exploration Geophysicists, pp. 1390–1394.
- de Matos, M.C., Osorio, P.L., Johann, P.R., 2007. Unsupervised seismic facies analysis using wavelet transform and self-organizing maps. Geophysics 72 (1), P9–P21.
- Devlin, J., Chang, M.-W., Lee, K., Toutanova, K., 2018. Bert: Pre-training of deep bidirectional transformers for language understanding. arXiv preprint arXiv:1810.04805.
- Dong, X., Bao, J., Zheng, Y., Zhang, T., Chen, D., Yang, H., Zeng, M., Zhang, W., Yuan, L., Chen, D., et al., 2023. Maskclip: Masked self-distillation advances contrastive language-image pretraining. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 10995–11005.
- Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., et al., 2020. An image is worth 16x16 words: Transformers for image recognition at scale. arXiv preprint arXiv:2010.11929.
- 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. J. Appl. Geophys. 112, 52–61.
- El-Nouby, A., Izacard, G., Touvron, H., Laptev, I., Jegou, H., Grave, E., 2021. Are large-scale datasets necessary for self-supervised pre-training? arXiv preprint arXiv: 2112.10740.
- Eldele, E., Ragab, M., Chen, Z., Wu, M., Kwoh, C., Li, X., Guan, C., 2021. Time-series representation learning via temporal and contextual contrasting. arXiv abs/2106.14112, URL: <https://api.semanticscholar.org/CorpusID:235658361>.
- Engel, N., Belagiannis, V., Dietmayer, K., 2021. Point transformer. IEEE Access 9, 134826–134840.
- He, Y., Cao, J., Lu, Y., Gan, Y., Lv, S., 2018. Shale seismic facies recognition technology based on sparse autoencoder. In: International Geophysical Conference, Beijing, China, 24–27 April 2018. Society of Exploration Geophysicists and Chinese Petroleum Society, pp. 1744–1748.
- Kim, Y., Hardisty, R., Torres, E., Marfurt, K.J., 2018. Seismic facies classification using random forest algorithm. In: SEG International Exposition and Annual Meeting. SEG, pp. SEG–2018.
- Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., Xiao, T., Whitehead, S., Berg, A.C., Lo, W.-Y., et al., 2023. Segment anything. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 4015–4026.
- Kourki, M., Ali Riahi, M., 2014. Seismic facies analysis from pre-stack data using self-organizing maps. J. Geophys. Eng. 11 (6), 065005.
- Li, J., Wu, X., Ye, Y., Yang, C., Hu, Z., Sun, X., Zhao, T., 2023. Unsupervised contrastive learning for seismic facies characterization. Geophysics 88 (1), WA81–WA89.
- Liu, Y., Hu, T., Zhang, H., Wu, H., Wang, S., Ma, L., Long, M., 2024. Itransformer: Inverted transformers are effective for time series forecasting. In: The Twelfth International Conference on Learning Representations. URL: <https://openreview.net/forum?id=JePfAI8fah>.

- Lubo-Robles, D., Devegowda, D., Jayaram, V., Bedle, H., Marfurt, K.J., Pranter, M.J., 2022. Quantifying the sensitivity of seismic facies classification to seismic attribute selection: An explainable machine-learning study. *Interpretation* 10 (3), SE41–SE69.
- Marroquín, I.D., 2014. A knowledge-integration framework for interpreting seismic facies. *Interpretation* 2 (1), SA1–SA9.
- McInnes, L., Healy, J., Melville, J., 2018. Umap: Uniform manifold approximation and projection for dimension reduction. arXiv preprint [arXiv:1802.03426](https://arxiv.org/abs/1802.03426).
- Mitchell, J., Neil, H.L., 2012. OS20/20 Canterbury — Great South Basin TAN1209 Voyage Report. *Voyage Report*, National Institute of Water and Atmospheric Research Ltd. (NIWA).
- Oquab, M., Darcret, T., Moutakanni, T., Vo, H., Szafraniec, M., Khalidov, V., Fernandez, P., Haziza, D., Massa, F., El-Nouby, A., et al., 2023. Dinov2: Learning robust visual features without supervision. arXiv preprint [arXiv:2304.07193](https://arxiv.org/abs/2304.07193).
- Priesszhev, I., Manral, S., 2012. 3D seismic waveform classification. In: Istanbul 2012-International Geophysical Conference and Oil & Gas Exhibition. Society of Exploration Geophysicists and The Chamber of Geophysical ..., pp. 1–4.
- Qi, J., Zhang, B., Lyu, B., Marfurt, K., 2020. Seismic attribute selection for machine-learning-based facies analysis. *Geophysics* 85 (2), O17–O35.
- Radford, A., Kim, J.W., Hallacy, C., Ramesh, A., Goh, G., Agarwal, S., Sastry, G., Askell, A., Mishkin, P., Clark, J., et al., 2021. Learning transferable visual models from natural language supervision. In: International Conference on Machine Learning. PMLR, pp. 8748–8763.
- Roy, A., Marfurt, K.J., 2011. Cluster assisted 3D and 2D unsupervised seismic facies analysis, an example from the Barnett shale formation in the Fort Worth basin, Texas. In: SEG International Exposition and Annual Meeting. SEG, pp. SEG-2011.
- Roy, A., Romero-Peláez, A.S., Kwiatkowski, T.J., Marfurt, K.J., 2014. Generative topographic mapping for seismic facies estimation of a carbonate wash, Veracruz basin, southern Mexico. *Interpretation* 2 (1), SA31–SA47.
- Sabeti, H., Nadjar, B., 2011. Seismic facies classification using 2-D and 3-D multi-attribute hierarchical clustering algorithms. In: SEG Technical Program Expanded Abstracts 2011. Society of Exploration Geophysicists, pp. 1160–1164.
- da Silva, A.L.S., Carneiro, V., Cerqueira, A.G., 2023. Cluster analysis to detect reflectors associates with salt dome structures.
- Tolstaya, E., Egorov, A., 2022. Deep learning for automated seismic facies classification. *Interpretation* 10 (2), SC31–SC40.
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, L., Polosukhin, I., 2017. Attention is all you need. *Adv. Neural Inf. Process. Syst.* 30.
- Wallet, B.C., Hardisty, R., 2019. Unsupervised seismic facies using Gaussian mixture models. *Interpretation* 7 (3), SE93–SE111.
- Wrona, T., Pan, I., Gathorpe, R.L., Fossen, H., 2018. Seismic facies analysis using machine learning. *Geophysics* 83 (5), O83–O95.
- You, Y., Gitman, I., Ginsburg, B., 2017. Large batch training of convolutional networks. arXiv preprint [arXiv:1708.03888](https://arxiv.org/abs/1708.03888).
- Zhang, H., Chen, T., Liu, Y., Zhang, Y., Liu, J., 2021. Automatic seismic facies interpretation using supervised deep learning. *Geophysics* 86 (1), IM15–IM33.
- Zhao, T., Li, F., Marfurt, K.J., 2017. Constraining self-organizing map facies analysis with stratigraphy: An approach to increase the credibility in automatic seismic facies classification. *Interpretation* 5 (2), T163–T171.
- Zhao, T., Li, F., Marfurt, K.J., 2018. Seismic attribute selection for unsupervised seismic facies analysis using user-guided data-adaptive weights. *Geophysics* 83 (2), O31–O44.
- Zhao, T., Zhang, J., Li, F., Marfurt, K.J., 2016. Characterizing a turbidite system in Canterbury basin, New Zealand, using seismic attributes and distance-preserving self-organizing maps. *Interpretation* 4 (1), SB79–SB89.
- Zhou, Y., Chen, W., 2022. Recurrent autoencoder model for unsupervised seismic facies analysis. *Interpretation* 10 (3), T451–T460.