

A seismic facies classification method based on the convolutional neural network and the probabilistic framework for seismic attributes and spatial classification

Zhege Liu¹, Junxing Cao², Yujia Lu¹, Shuna Chen¹, and Jianli Liu¹

Abstract

In the early stage of oil and gas exploration, due to the lack of available drilling data, the automatic seismic facies classification technology mainly relies on the unsupervised clustering method combined with the seismic multiattribute. However, the clustering results are unstable and have no clear geologic significance. The supervised classification method based on manual interpretation can provide corresponding geologic significance, but there are still some problems such as the discrete classification results and low accuracy. To solve these problems, inspired by hyperspectral and spatial probability distribution technology, we have developed a classification framework called the probabilistic framework for seismic attributes and spatial classification (PFSSC). It can improve the continuity of the classification results by combining the classification probability and the spatial partial probability of the classifier output. In addition, the convolutional neural network (CNN) is a typical classification algorithm in deep learning. By convolution and pooling, we could use it to extract features of complex nonlinear objects for classification. Taking advantage of the combination of PFSSC and CNN, we could effectively solve the existing problems mentioned above in seismic facies classification. It is worth mentioning that we select seismic the multiattribute by maximal information coefficient (MIC) before the seismic facies classification. Finally, using the CNN-PFSSC and MIC methods, we can obtain high accuracy in the test set, reasonable continuity within the same seismic facies, clear boundaries between different seismic facies, and seismic facies classification results consistent with sedimentological laws.

Introduction

Seismic facies analysis means that seismic analyses respond to characteristics of underlying geologic bodies according to their seismic reflection parameters. It can serve in sedimentary microfacies research and reservoir characteristic description. In recent years, the seismic facies analysis method has been developed to be automated and intelligent, and the results are more objective, have higher work efficiency, and are more conducive to the quantitative description of the reservoir. Based on image segmentation, Lomask et al. (2007) provide a modified version of the normalized cuts image segmentation method to partition seismic images along salt boundaries. Also, Halpert et al. (2014) apply an efficient graph-based image segmentation technology on 3D seismic images to delineate salt bodies. Liu et al. (2017) introduce a spatial constraint during seismic facies analysis, which modifies the seed region growing algorithm to avoid the over-fitting problem in the support vector machine (SVM). At the same time, some of the

literature uses texture analysis to quantitatively map seismic facies (West et al., 2002; Gao, 2007). There are two major categories of seismic facies analysis methods using pattern recognition: supervised and unsupervised. Taking unsupervised methods as an example, K-means clustering, principal component analysis, projection pursuit, and neural networks are widely used in seismic facies analysis (Coléou et al., 2003; de Matos et al., 2006; Saraswat and Sen, 2012; Roy et al., 2013; Zhao et al., 2018). For supervised seismic facies analysis, SVM artificial neural networks (West et al., 2002; Bagheri and Riahi, 2013; Roy et al., 2013) are most popular. In addition, Qi et al. (2016) propose a semisupervised seismic facies analysis method to delineate the salt, which introduces a degree of supervision by computing histograms for each candidate attribute for a suite of user-defined facies to calculate the probability of a given facies at each voxel.

In this paper, we mainly focus on the early stage of exploration, due to the limited well-logging data, and

¹Chengdu University of Technology, Geophysics Institute, Chengdu 610059, China. E-mail: liuzhege@163.com; 1033415284@qq.com; chen_shuna@163.com; 903641224@qq.com.

²State Key Laboratory of Oil and Gas Reservoir Geology and Exploitation, Chengdu 610059, China and Chengdu University of Technology, Geophysics Institute, Chengdu 610059, China. E-mail: caojx@cdut.edu.cn (corresponding author).

Manuscript received by the Editor 21 December 2018; revised manuscript received 24 April 2019; published ahead of production 28 May 2019; published online 30 July 2019. This paper appears in *Interpretation*, Vol. 7, No. 3 (August 2019); p. SE225–SE236, 14 FIGS., 5 TABLES.

<http://dx.doi.org/10.1190/INT-2018-0238.1>. © 2019 Society of Exploration Geophysicists and American Association of Petroleum Geologists. All rights reserved.

the anisotropy of the carbonate reservoir in the Sichuan Basin, it is difficult to classify the seismic facies through the unsupervised method or to straightly use a group of seismic attributes, and the distribution of some seismic facies may not be consistent with the discipline of sedimentology. We try to use the supervised classification method to classify seismic facies, but there is not a large number of well logs in this area, so we need to add some training samples only relying on the results of manual interpretation. These sample profiles also account for only a small part of the whole work area (approximately 1%). In the process of supervised learning, there will inevitably be over-fitting, and the classification results will show discrete and discontinuous errors. In view of the above problems, referring to the hyperspectral classification method (Liu and Lu, 2016), this paper proposes a new seismic classification method especially for the oil fields with few drillings, which is based on a probabilistic framework for seismic attributes and spatial classification (PFSSC), and we classify seismic facies through the convolutional neural network (CNN) (LeCun et al., 1989). CNN has achieved good results in various fields of geophysics, such as identifying geologic features, earthquake detection and location, and first-break wave pickup (Huang et al., 2017; Perol et al., 2018; Yuan et al., 2018). In the case of complex data, most cases can achieve better results than shallow learning. Here, we combine the two methods and apply them to seismic facies classification. Using characteristics analysis of seismic attributes and the CNN-PFSSC method, we can get a reasonable seismic facies distribution with clear boundaries between the two seismic facies and a more continuous distribution inside of one seismic facie. The reasonability evaluation of the seismic facies classification is depended on its consistency with artificial interpretation results. In this paper, we evaluate it through comparison with the manually interpreted seismic profiles and sedimentary facies (depositional environment interpretation) plane-distribution map. In the “Application” section, we take the seismic facies of a reef-shoal reservoir in a carbonate platform in the Sichuan Basin as a research target.

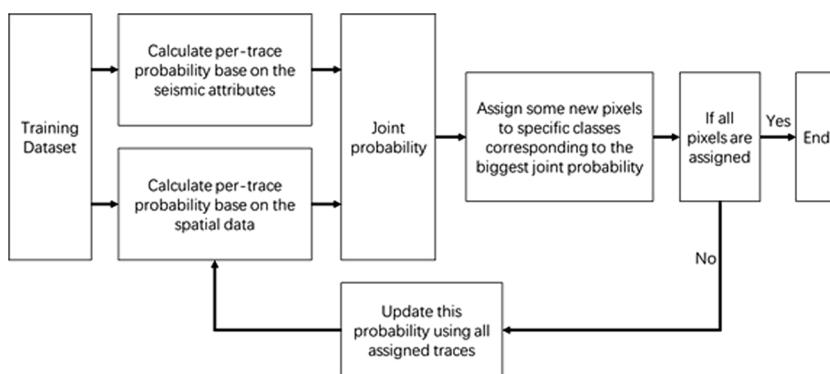


Figure 1. A probabilistic framework based on seismic attributes and spatial classification.

Methods

PFSSC

Seismic facies is a comprehensive representation on seismic profiles of the interfaces between different sedimentary systems, lithology, and ensemble characteristics of underground geologic bodies. In this paper, we do seismic facies classification to serve for the sedimentary facies analysis. It can also be used for lithologic analysis in future. Referring to Walther's law of correlation of facies, only those facies and facies areas can be superimposed primarily, which can be observed beside each other at the present time. Considering that seismic facies have a corresponding relationship with sedimentary facies, the distribution of seismic facies should also have some similarities to that of sedimentary facies. For example, the distribution of seismic facies should also have a certain degree of continuity within a specific geologic scale, in the vertical and horizontal directions, such as in the same facies area. However, it should be noted that this continuity scale is dependent on our research duty, if we want to use seismic facies classification for reservoir description, then we should adjust the continuity scale to a new one, such as the same lithologic facies or some special geologic bodies. This kind of continuity in spatial classification is similar to the classification of ground objects in spectral classification. If the result of the classifier aims at the output of each seismic trace, it is easy to produce discrete classification results in the cluttered seismic reflection, lacking continuity. The framework of classification probability combined with spatial distribution information has been discussed in the classification of spectral features (Zhang et al., 2011; Fauvel et al., 2013; Camps-Valls et al., 2014). In this paper, based on the framework proposed by Liu et al. (2017), we use seismic attribute data instead of hyperspectral data and we use CNN as a classifier to supervise the classification of seismic facies. The classification probability framework based on seismic attributes and spatial information is shown in Figure 1.

In the whole process of classification optimization, the output of the classifier can be regarded as a planar 2D image, and the classification result of each seismic trace corresponds to one pixel. The class probability value of a

single pixel can be obtained through the classifier, then we can estimate the spatial probability information by combining the spatial distribution, and finally we can get the classification results.

First, the first- or second-order neighborhood system is needed to define the neighborhood space. This means that a pixel class will be affected by the classification results of four or eight neighboring pixels in the neighborhood. In this paper, we take the first-order neighborhood system as an example. In Figure 2, the neighborhood system of r pixel includes four blue pixels. Here, we define all pixels as follows:

$$S = \{1 \leq t \leq N | \text{Cor}(t) = (i, j), 1 \leq i \leq N_x, 1 \leq j \leq N_y\}, \quad (1)$$

$$X = \{x_s | s \in S, 1 \leq x_s \leq L\}, \quad (2)$$

where S is a collection of all pixels, X is the class label corresponding to the pixel, N_x is the width of the image, N_y is the height of the image, $N = N_x N_y$ is the number of all pixels, L represents the number of classes, and $\text{Cor}(t)$ represents the coordinates of the t pixel. Suppose we already know that the class of the pixel r is l , and $P(x_s = l) = 1$, then we can define that the conditional probability of the adjacent pixels have the same classification as follows:

$$P(x_s = l | x_r = l) = \beta. \quad (3)$$

Thus, we can derive that the conditional probability that pixel s belongs to a category but different from the pixel as follows:

$$P(x_s = l' | x_r = l) = \frac{1 - \beta}{L - 1}, \quad l' \neq l. \quad (4)$$

As shown in Figure 3, the pixel position relationship shows that the pixel m is not adjacent to r but adjacent to s . With the help of intermediate pixel nodes, it is still possible to write conditional probability formulas for the classification of pixels and pixels when the category of pixels is given:

$$\begin{aligned} & P(x_m = l | x_r = l) \\ &= P(x_m = l, x_s = l | x_r = l) + P(x_m = l, x_s \neq l | x_r = l) \\ &= P(x_m = l | x_s = l, x_r = l) \cdot P(x_s = l | x_r = l) \\ &\quad + P(x_m = l | x_s \neq l, x_r = l) \cdot P(x_s \neq l | x_r = l) \\ &= P(x_m = l | x_r = l) \cdot P(x_s = l | x_r = l) \\ &\quad + \sum_{l' \neq l}^L P(x_m = l | x_s = l') \cdot P(x_s = l' | x_r = l) \\ &= \beta^2 + \frac{(1 - \beta)^2}{L - 1}. \end{aligned} \quad (5)$$

By analogy, it is possible to calculate the probability that any pixel has the same classification condition as a pixel in a given class. The conditional probability of a pixel representing a distance from a known class of pixels having the same classification is expressed as follows:

$$P_n = \left(\frac{L\beta - 1}{L - 1} \right)^{n-1} \left(\beta - \frac{1}{L} \right) + \frac{1}{L}. \quad (6)$$

After obtaining the conditional probability of pixel-pixel classification, we

need to determine the class of a certain pixel. The spatial classification probability in the window range is calculated by taking the specific pixel as the center and the radius as the window:

$$P_{\omega}^{\text{spa}} = \{p_{\omega_1}^{\text{spa}}, \dots, p_{\omega_L}^{\text{spa}}\}, \quad (7)$$

$$p(x_w = l) = \begin{cases} \frac{1}{z} \sum_{i \in W}^l p_{w_{y_i}}^i I(y_i = l), & l \in [1, L], f(W) \neq 0, \\ \frac{1}{L}, & l \in [1, L], f(W) = 0, \end{cases} \quad (8)$$

where P_{ω}^{spa} is the spatial distribution probability of each category of pixels, W represents the set of pixels contained in a pixel-centered window, and the function of $f(W)$ is to determine the number of classified pixels in the window.

Finally, it is necessary to estimate β to calculate the final classification conditional probability for each category. Here, the conditional probabilities of adjacent pixels are counted according to the posterior results of the classification results. When the posterior probability is counted, only the marked pixels are counted, whereas the unmarked pixels are not processed:

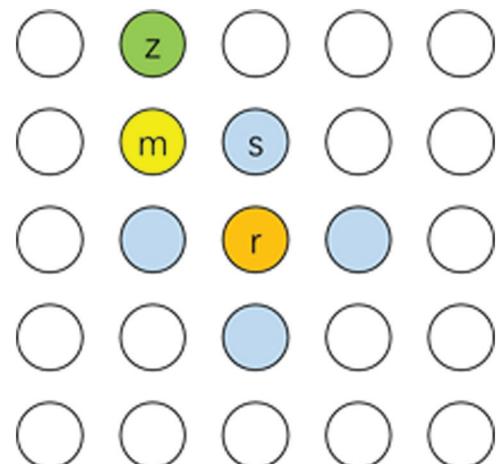


Figure 2. A schematic diagram of the neighborhood relation.

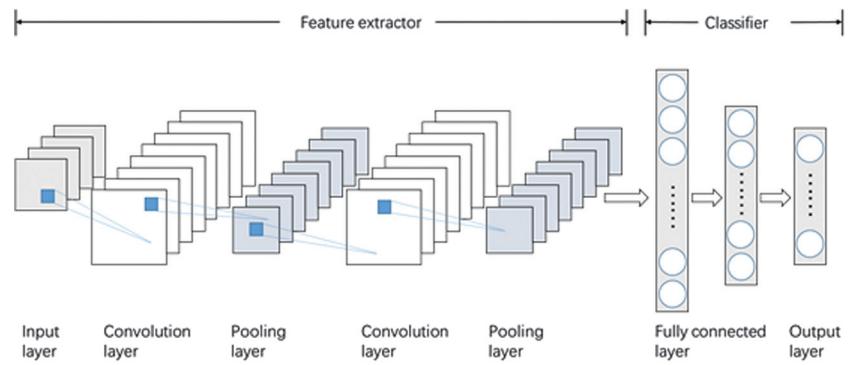


Figure 3. Convolution neural network structure.

$$n_{\text{same}} = \sum_{i=1}^N \sum_{j \in \text{Nei}(i)} I(x_i = x_j), \quad (9)$$

$$n_{\text{diff}} = \sum_{i=1}^N \sum_{j \in \text{Nei}(i)} I(x_i \neq x_j), \quad (10)$$

$$\beta = \frac{n_{\text{same}}}{n_{\text{same}} + n_{\text{diff}}}, \quad (11)$$

where N is the total number of pixels. The function I is the indication function. The function Nei is the neighborhood function, which returns the first-order neighborhood pixels of the current position.

After calculating the probability of seismic attribute classification P_1 and the corresponding spatial distribution probability P_2 , the whole classification area can be optimized and the result of the attribute classification can be reclassified. First, according to the classification results of seismic attributes, only the classification results whose probability value is greater than a certain threshold are retained, and other pixels are treated as unlabeled. Then, when an unmarked pixel is recalculated, we set a window around the pixel and reevaluate its classification results by using the classification probability of the pixel and the categories of other allocated pixels in the window. The above processes are carried out on the whole layer plane in turn, until all of the pixels are allocated to the classification result aborted. The probability that pixel ω belongs to class l is

$$p_{\omega_l} = w p_{\omega_l}^{\text{attr}} + (1 - w) p_{\omega_l}^{\text{spa}}, \quad (12)$$

where w is the proportion between the attribute classification probability and the spatial distribution probability. The final classification result $\text{Label}(\omega)$ is

$$\text{Label}(\omega) = \arg \max(p_{\omega_1}, \dots, p_{\omega_L}). \quad (13)$$

Finally, we display the flowchart of the process in Table 1.

CNN

The CNN was proposed for handwritten image recognition by professor LeCun et al. (1989) from the University of Toronto, Canada (LeCun et al., 1989). Hinton and Salakhutdinov (2006) put forward the concept of deep learning based on it. When CNN has a multilayer hidden layer structure, it can reduce the dimension of high-dimensional nonlinear data (Hinton and Salakhutdinov, 2006). Finally, Hinton's student Krizhevsky et al. (2012) implemented the ImageNet model on GPU in 2012, and they achieve outstanding results in the application of image recognition. Because CNN has good performance for high-dimensional nonlinear classification problems, we use CNN as a PFFSC classifier in this paper.

The basic structure of CNN consists of input layer, convolution layer, pooling layer, fully connected layer, and output layer. The structure of the classic CNN is shown in Figure 3.

Assume that CNN has L layers, the feature of the (l) th layer is represented by x^l , $l = 1, \dots, L - 1$. In the convolution layer and the pooling layer, x^l consists of multiple feature maps x_j^l , which is $x^l = \{x_1^l, \dots, x_{N_l}^l\}$. In the fully connected layer, feature x^l is a vector, which is $x^l = (x_1^l, \dots, x_{N_l}^l)^T$, and N_l is the number of the feature maps or the features of the (l) th layer. Now, let us briefly introduce the five structures and role of the network in the below.

First, the role of the input layer is to accept the input data, which can be a tensor of multiple dimensions. In the case of an image, it is a 2D vector.

Second, the function of the convolutional layer is to extract features using a convolution operation. If the (l) th layer is a convolutional layer, the feature map of the layer can be expressed as

Table 1. Processing of framework for seismic attributes and spatial classification.

Probabilistic framework for seismic attributes and spatial classification

- 1) Calculate the per-trace seismic attributes distribution $p_{\omega_l}^{\text{attr}}$;
- 2) Calculate the initial per-trace spatial distribution $p_{\omega_l}^{\text{spa}}(0)$;
- 3) Mark all samples with $\text{Max}(p_{\omega_l}^{\text{attr}})$ below the threshold as unlabeled;
- 4) **for** iter = 0:N **do**
 - 1) For per-trace is not assigned, calculate $p_{\omega_l} = w p_{\omega_l}^{\text{attr}} + (1 - w) p_{\omega_l}^{\text{spa}}(\text{iter})$ and assign trace in window with maximum joint probability $p_{\omega_l}^{\text{max}}$ to specific classes;
 - 2) Add these reassigned traces into training samples.**end**
- 5) Update the spatial distribution to $p_{\omega_l}^{\text{spa}}(\text{iter})$.
- 6) End this process if all pixels have been assigned. Else, return to step (4).

$$x_j^l = f \left(\sum_{i=1}^{N^{l-1}} G_{i,j}^l \left(k_{i,j}^l \otimes x_i^{l-1} \right) + b_j^l \right), \quad j = 1, \dots, N^l, \quad (14)$$

where $k_{i,j}^l$ and b_j^l are the offsets of the convolution kernel and the convolutional layer, respectively. The operation symbol \otimes represents the convolution operation, and G^l is the connection matrix between the current convolutional layer and the previous feature map. When $G_{i,j}^l$ equals one, it means that the feature map x_i^{l-1} are associated with x_j^l ; whereas when $G_{i,j}^l$ equals zero, it means that they are not associated. The function $f(x)$ represents an activation function.

Third, the pooling layer is connected behind the convolutional layer. The function of down-sampling is realized by pooling the local area of the feature map. If the (l) th layer is a pooling layer, the feature map of the layer can be expressed as

$$x_j^{l+1} = p(x_j^l), \quad (15)$$

where $p(x)$ is the pooling operation function. The pooling layer has a function similar to feature selection. And this means, according to the rules, that the important features can be calculated from the local regions of the feature map. Normally, the pooling layer can select local regions without overlapping. Therefore, it can reduce the feature dimension while ensuring the robustness of the features. It can be seen from equation 2 that the feature maps of the convolutional layer and the pooling layer correspond to each other; that is, $N^{l+1} = N^l$.

Fourth, the fully connected layer is located behind the convolutional or pooling layer and is connected with the feature maps' output from the previous layer. If the (l) th layer and the previous layer are fully connected layers, the feature vector of (l) th layer can be expressed as

$$x^l = f(w^l x^{l-1} + b^l), \quad (16)$$

where w^l and b^l represent the weight and offset of the fully connected layer, respectively. The function $f(x)$ represents the activation function. If the $(l-1)$ th layer is a convolution layer or a pooling layer, then all of the feature maps' output from the convolutional layer or the pooling layer need to be arranged into a vector, and then we can calculate the feature vector of the (l) th layer according to equation 3.

Fifth, whereas the CNN is used as a classifier, the output layer could use Softmax regression. Then, the prediction vector of the output classification result will be $y = (y_1, \dots, y_C)^T$, where C represents the number of categories. Each component of the prediction vector can be calculated from

$$y_i = \frac{e^{-w_i^l x^{l-1}}}{\sum_{j=1}^C e^{-w_j^l x^{l-1}}}, \quad (17)$$

where w_i^l is the weight of the Softmax regression and $i = 1, \dots, C$. The output of the Softmax layer shows the probability expectation of each category corresponding to each input sample. When calculating the classification result, we should take category i as the output result, and this i is from the maximal output probability value y_i .

Workflow

The overall flow of the CNN-PFFSC algorithm is shown in Figure 4, which is mainly divided into four parts.

In the first part, we first extracted the various seismic attributes from the 3D seismic data, and then we pre-process the seismic attributes, including removing outliers and normalization. This is the data preparation for the later training. Then, we need to optimize the pre-processed seismic attributes. Because the redundancy is likely to increase the complexity of the classification algorithm and make the influence of different types of seismic attributes on the classification result becomes unbalanced, it is necessary to select the appropriate attributes combination. In this paper, for the nonlinear classification algorithm, we choose the maximal information coefficient (MIC) (Reshef et al., 2011; Zhang et al., 2014) as a standard. The MIC of each pair of attributes is calculated to evaluate the correlation between the attributes. MIC is a maximization and normalization method based on mutual information, whose fairness is better than other methods, such as the mutual information estimation, distance correlation, Spearman correlation coefficient, master curve-based method, and maximum correlation.

In the second part, we first selected several representative profiles from the 3D seismic data, and then we obtained the seismic facies classification for those profiles depending on experienced seismic interpreters. Then, we mapped the manually classified seismic facies categories to its corresponding digital labels. And, the seismic data corresponding to the labels can be used as training and test samples.

In the third part, CNN is trained with training samples until the loss function converges steadily. The classification results of the test data set are output, and the accuracy, recall and F1-score of the corresponding tags are compared to evaluate the classification performance of CNN. In addition, the classification results of the whole horizon data are predicted, and the geologic structure is inspected artificially. If the requirements of the interpreters are met, follow-up steps will be carried out. If not, the profiles will be reselected or we will increase the new samples for interpretation, new training and testing samples will be generated, and training and evaluation will be reconducted. To meet the training requirements, we suggest that at least 80% of the correctness of the sample and manual interpretation results be verified on the interpretation profile of the connecting wells.

In the fourth part, based on the classification results from CNN output, combined with the framework of spatial distribution probability classification, we optimized

the classification result through iteration, and then we output the optimal result finally.

Geologic setting

The study area is located in the Northeast Sichuan Basin, with a range of approximately 400 km² and a target interval depth of 5000–6000 m. In addition, we take the formation Fei1 in the target interval as an application example. The seismic sampling interval is 1 ms, the main frequency is 30 Hz, and there are three wells X1, Y1, and Z1 in this area, which are located in three different sedimentary facies in the Fei1 sedimentary period, which are the platform marginal shoals, inner shoals, open platform, and lower slope (Figure 5).

Figure 6 shows the artificial interpretation of the seismic facies of the seismic profile across three wells. The Fei1 formation we are studying is from horizon F1 to horizon F2 in Figure 6. According to the seismic reflection characteristics, the seismic facies in this profile is divided into four types. The seismic facies characteristic of the first type is continuous reflection, middle-high amplitude, and low frequency, and its corresponding sedimentary facies is platform marginal shoals. The characteristic of the second type is semi-continuous reflection, low frequency, and middle-high amplitude, and its corresponding sedimentary facies is inner shoals. The third type is chaotic reflection and low amplitude, and its corresponding sedimentary facies is the transition zone, which is the transitional area between the marginal shoals and the inner shoals. The fourth type is discontinuous, chaotic reflection, low amplitude, and low frequency, and its corresponding

sedimentary facies is the lower slope in the Fei1 sedimentary period.

Application

We apply the algorithm to the research area. Before that, we need to complete two steps: (1) determine the data sample and label and (2) choose the best classifier.

Preparation

First, we need to prepare data samples and labels. We divided the labeled samples into two parts: training and test samples. The position of the sample labeling is shown in Figure 7. The red line is the training sample position, and the blue line is the test sample position. The information of sample classifications is shown in Table 2.

Second, we extracted 28 types of seismic attributes as the data source of the optimized attributes. We used MIC as the reference index, which is illustrated in the third section in this paper. Then, we took one attribute as the representative, which is selected from the attribute group with the MIC value exceeding 0.75. Finally, we optimized six types of attributes as the training sample data (MIC values are shown in Table 3, and horizontal slices are shown in Figure 8). In Table 3, we abbreviate the entropy of the gray-level co-occurrence matrix (GLCM), relative impedance, attenuation of weak-energy frequency, low-frequency partial energy, root-mean-square (rms) amplitude differential, rms amplitude to entropy of gray-level co-occurrence matrix (EGLCM), relative impedance (RI), attenuation of weak-energy frequency (AWEF), low-frequency partial energy (LFPE), root-mean-square of amplitude differential (RMSAD), and root-mean-square of amplitude (RMSA), respectively. At this point, we have the sample data and labels ready.

Classifier selection

In this paper, we have compared the classification results of the Gaussian Bayesian, K-nearest neighbor, SVM, multilayer perceptron, CNN, and the combination

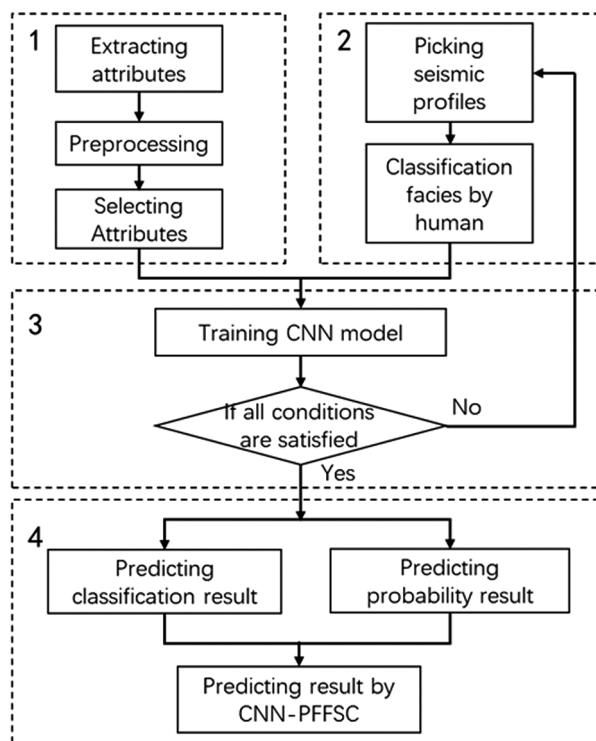


Figure 4. The CNN-PFFSC workflow.

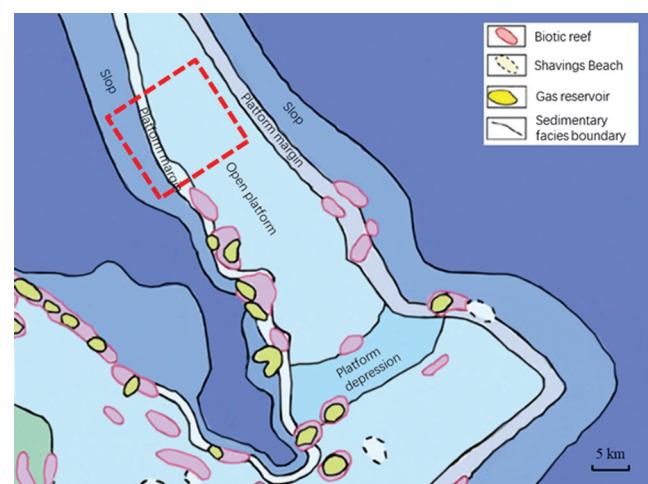


Figure 5. Research area.

of CNN and spatial information classification methods. CNN uses two layers of a convolution-pooling structure, which is implemented in Python with the tensorflow framework. The key codes and parameters are shown in Table 4. During classifying, the attributes of each seismic trace selected are input as N -dimensional

vectors. For shallow networks, it is necessary to flatten the N -dimensional vectors to a 1D vector. For CNN, the input layer uses a 1D convolution core, and each seismic attribute corresponds to a channel of the convolution core.

Based on the same test set, the classification results from the algorithms mentioned above are shown in Figure 9. In this figure, we show the labelled samples with different seismic waveforms and part of the manual labeling samples. The color strips from left to right are the seismic facies corresponding to the lower slope, the platform marginal shoals, the transition zone, and the inner shoals. In addition, the classification color bar shows the classification results' distribution from various algorithms based on the test set, the error color bar highlights the distribution of the wrong classification samples, the green color bar is the correct classification result, and the pink color bar is the wrong classification result. Obviously, the results of the CNN classification are better than other traditional classification algorithms, and CNN-PFFSC further improves the classification accuracy based

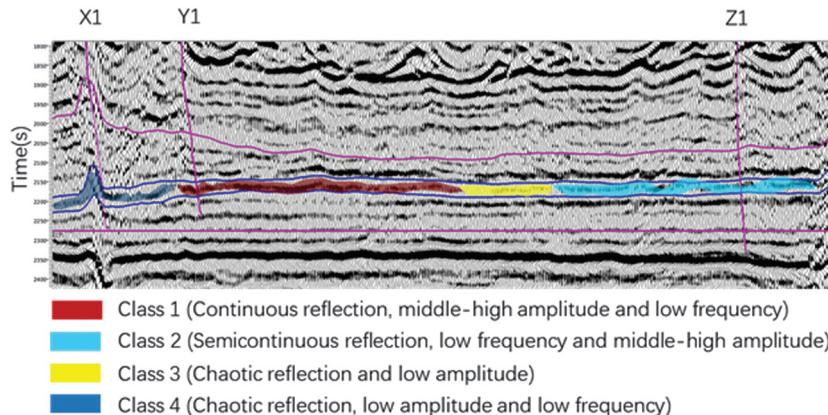


Figure 6. Artificial interpretation for the well-tie seismic cross section.

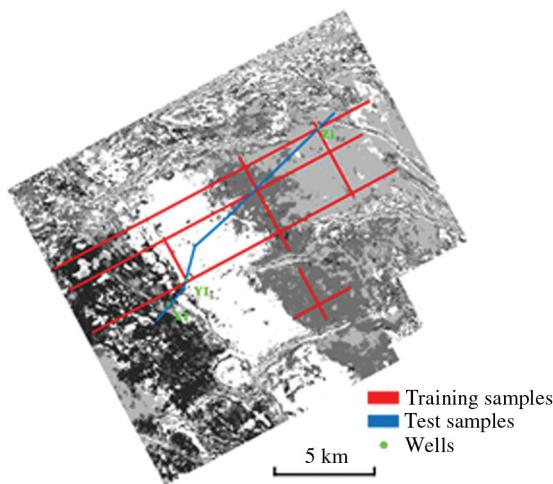


Figure 7. The training-sample and test-sample distribution.

Table 3. Maximum information coefficient of optimum attributes.

	EGLCM	RI	AWEF	LFPE	RMSAD	RMSA
EGLCM	1	0.45	0.48	0.57	0.45	0.49
RI	0.45	1	0.59	0.7	0.49	0.52
AWEF	0.48	0.59	1	0.7	0.51	0.58
LFPE	0.57	0.7	0.7	1	0.65	0.73
RMSAD	0.45	0.49	0.51	0.65	1	0.41
RMSA	0.49	0.52	0.58	0.73	0.41	1

Table 2. Training samples and test samples.

Class of seismic facies	Number of training samples	Number of test samples	Type of sedimentary facies	Lithology
Class 1	1614	188	Marginal shoals	Oolite dolomite and lime-bearing dolomite
Class 2	1720	233	Inner shoals	Micrite dolomite with oolitic dolomite
Class 3	1222	160	Transition zone	Unknown
Class 4	1397	178	Lower slope	Limestone

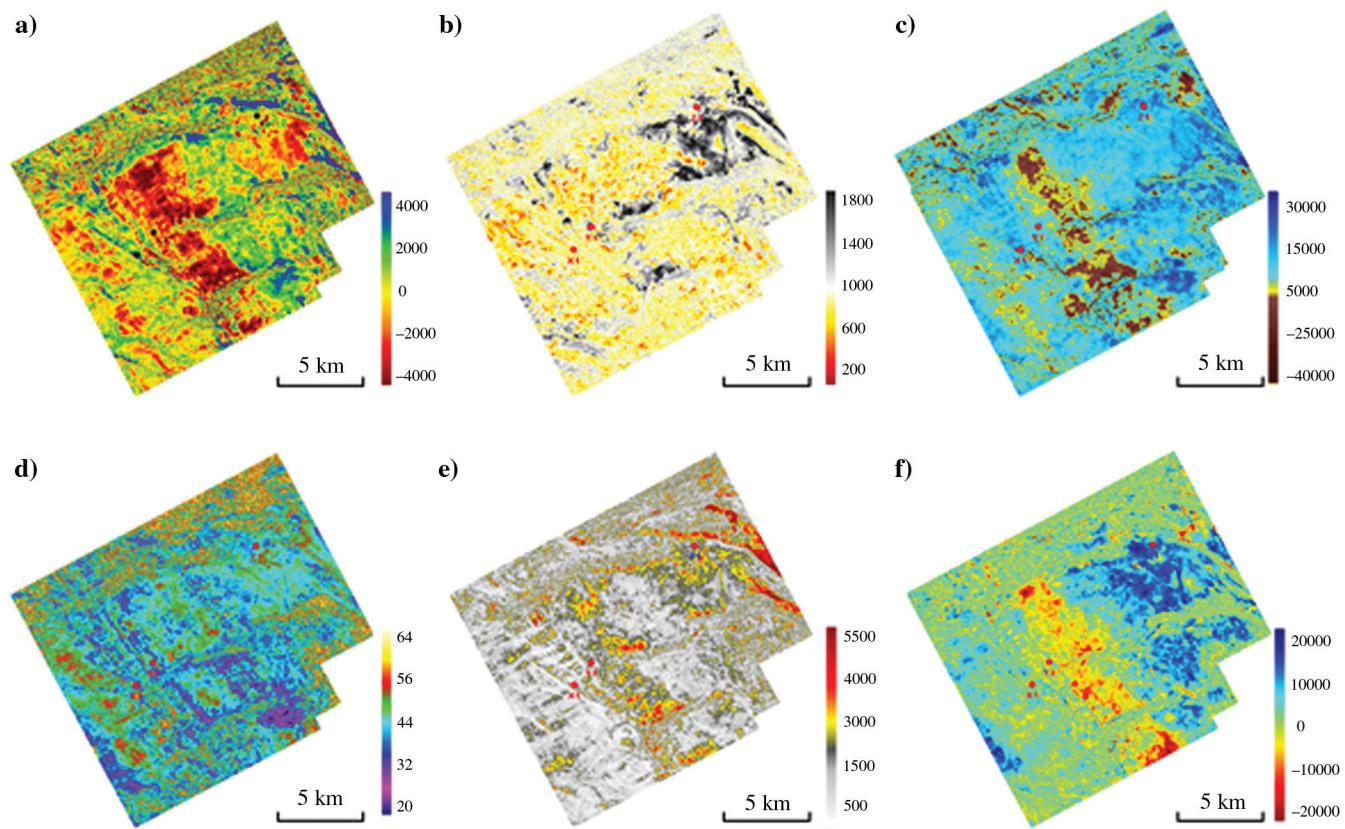


Figure 8. Selected attributes: (a) entropy of GLCM, (b) relative impedance, (c) attenuation of weak energy frequency, (d) low-frequency partial energy, (e) rms amplitude differential, and (f) rms amplitude.

Table 4. The CNN graph-construction code.

CNN graph construction code in tensorflow

```
# Input/output layer: tf_x/tf_y
tf_x = tf.placeholder(tf.float32, [None, sample_size, attributes_num])
tf_y = tf.placeholder(tf.int32, [None, label_size])
# Convolution layer/Pooling layer: conv_layer1-pool_layer2-conv_layer1-pool_layer2
conv_layer1 = tf.layers.conv1d(inputs= tf_x, filters=8, kernel_size=3, strides=1, padding='SAME', activation=tf.nn.relu)
pool_layer1 = tf.layers.max_pooling1d(conv1, pool_size=2, strides=2, padding='SAME')
conv_layer2 = tf.layers.conv1d(inputs=pool1, filters=4, kernel_size=3, strides=1, padding='SAME', activation=tf.nn.relu)
pool_layer2 = tf.layers.max_pooling1d(conv2, pool_size=2, strides=2, padding='SAME')
# Flat layer
flat_layer = tf.reshape(pool_layer2, [-1, flat_size])
# Fully connection layer
fully_connection_layer = tf.layers.dense(network.flat, label_size)
# Adding Dropout Regularization Processing to Full Connection Layer and Output
output = tf.nn.dropout(fully_connection_layer, keep_prob=0.8)
# Calculating Softmax Cross Entropy as Loss Function
loss = tf.losses.softmax_cross_entropy(onehot_labels=tf_y, logits=output)
# Minimizing Loss Function by Using Adam Optimal Algorithms
train_operate = tf.train.AdamOptimizer(learning_rate=0.01).minimize(loss)
```

on the CNN results. The following is an evaluation of various classification algorithms depending on numerical indicators, as shown in Table 5.

From the classification indexes of the test set in Table 4, the accuracy and recall rate of the CNN classification algorithm on the test set are higher than other algorithms. Therefore, we chose CNN as the classifier of PFSSC. The results of the CNN classification will directly affect the accuracy of the results processed in the later stage under PFSSC. Figure 10 shows the

probability distribution of various seismic zones in the output interval after CNN classification training.

Results

Putting the CNN classification results into PFSSC optimization can further improve the accuracy of the classification and enhance the boundaries. Although CNN-PFSSC cannot fully achieve the correct rate of manual interpretation (such as an accuracy rate greater than 95%), it can be considered that in the manual labeling sample, some waveforms themselves are also near the fault zone, which will cause some wrong labeling, and finally the classification result is incorrect. From the output of the entire classification results of the various classification algorithms in Figure 11, CNN-PFSSC still achieved the best results.

Figure 12 shows a step-by-step iterative process for the classification results of CNN based on PFSSC. For the initial condition, we first take the points, of which the classification probability is greater than 0.95 in the CNN classification, as the point with assigned category, and then, we start the iteration for these points. The weight of the classifier probability and the spatial distribution probability is set to 0.6.

Table 5. Numerical results of classification algorithms on test data sets.

Algorithm	Accuracy	Recall	F1-score
Gaussian Bayesian	64.01%	64.72%	59.43%
KNN	78.73%	76.12%	76.50%
SVM	81.23%	81.21%	81.20%
ML	80.01%	79.10%	79.46%
CNN	84.23%	83.65%	83.41%
CNN-PFSSC	88.10%	84.33%	86.75%

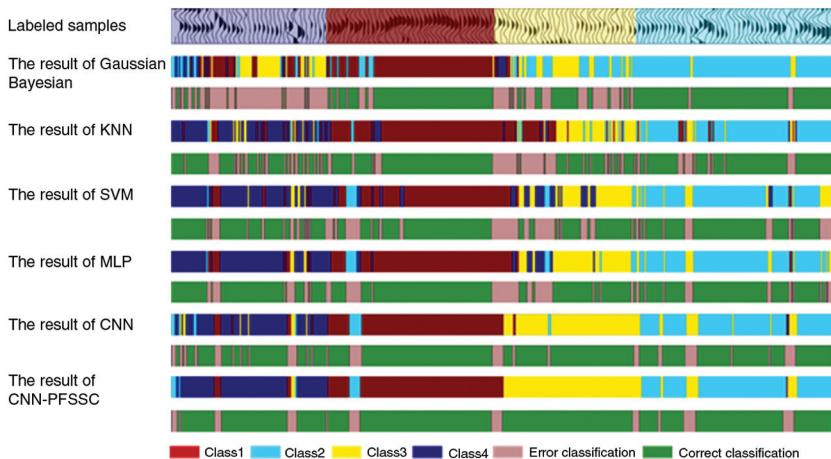


Figure 9. Comparisons of classification algorithms on test data set.

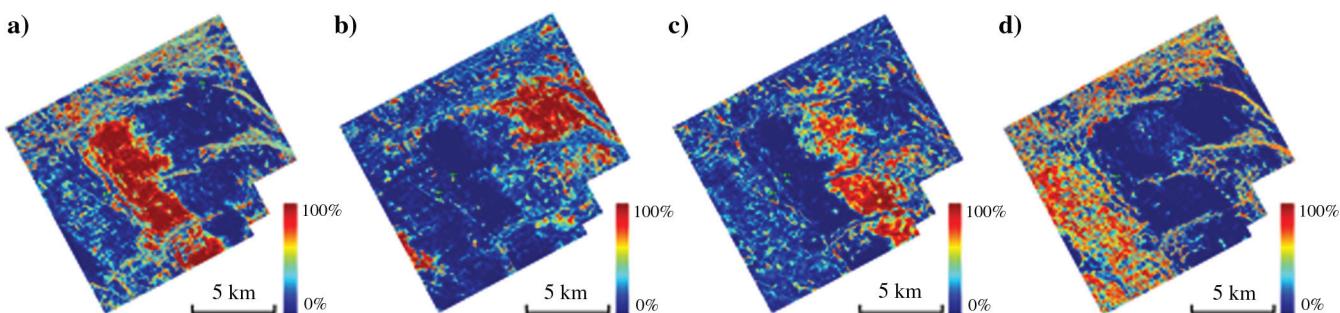


Figure 10. Results of various probability distributions: (a) class 1, (b) class 2, (c) class 3, and (4) class 4.

One goal of the seismic facies division is to help to determine the boundary of sedimentary facies consistent with geology. Figure 13 shows the results of the sedimentary facies interpretation based on the CNN-PFSSC results. The results are consistent with the actual geologic structure. The classification results are consistent with the logging interpretation at the three drillings.

Last, we compared the unsupervised classification results based on K-means with the supervised classification results based on CNN-PFFSC in this paper. This unsupervised classification algorithm is uncertain on the number of clusters, and some trials are needed to select them through geologic understanding. Here, we divided the seismic waveforms and seismic attributes into four and seven classes, respectively, and the clustering results are shown in Figure 14. Due to the lack of geologic understanding of artificial interpretation, and the uncertainty of classification categories, the unsupervised clustering re-

sults have difficulty displaying the geologic features of the sedimentary facies, and the boundaries are not clear.

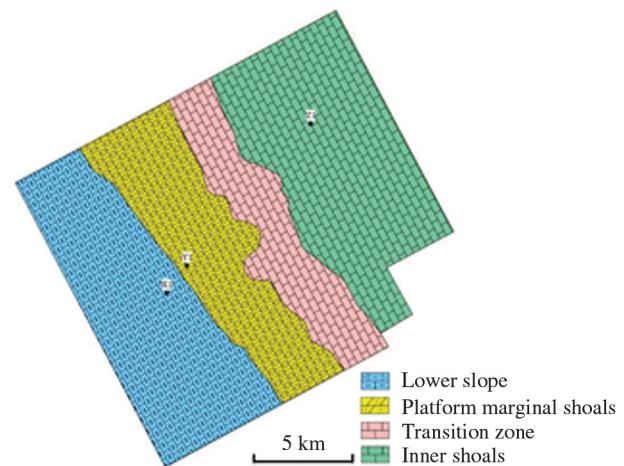


Figure 13. Artificial interpretation results of sedimentary facies.

Figure 11. (a) Gaussian Bayesian, (b) KNN, (c) SVM, (d) MLP, (e) CNN, and (f) CNN-PFSSC.

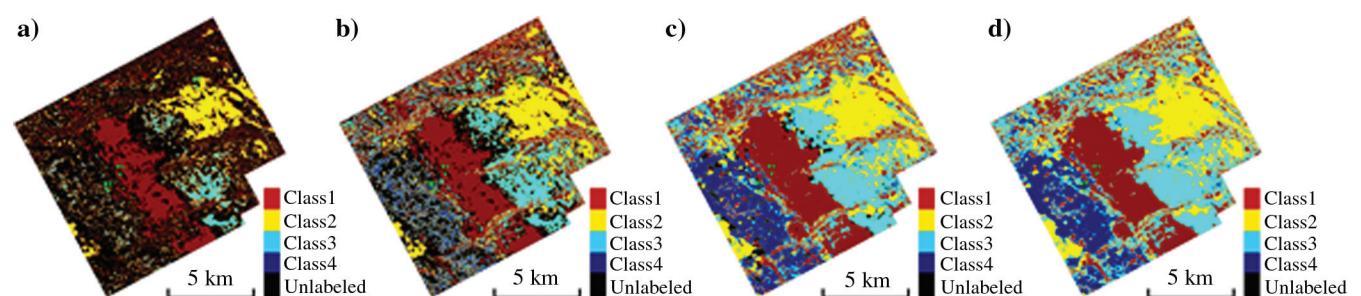
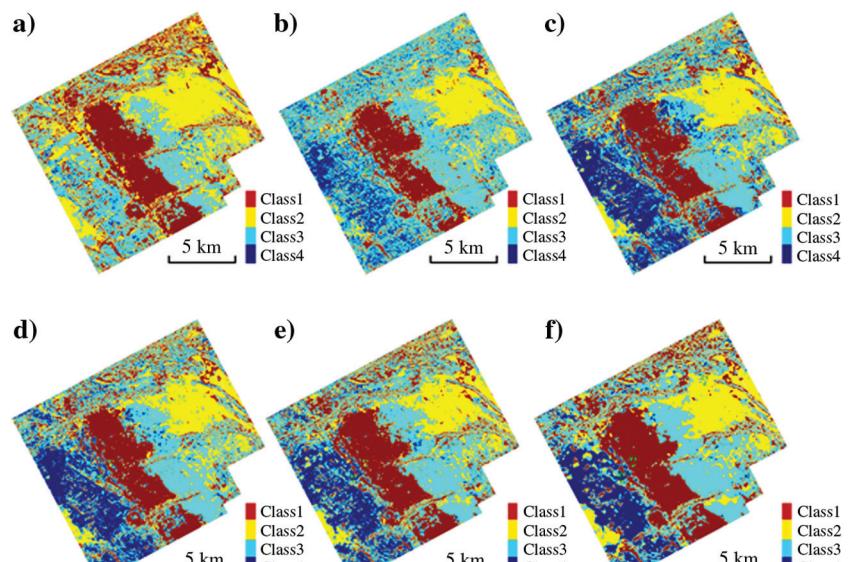


Figure 12. The classification results from (a-d) are separately iteration 1–5–10–15 times.

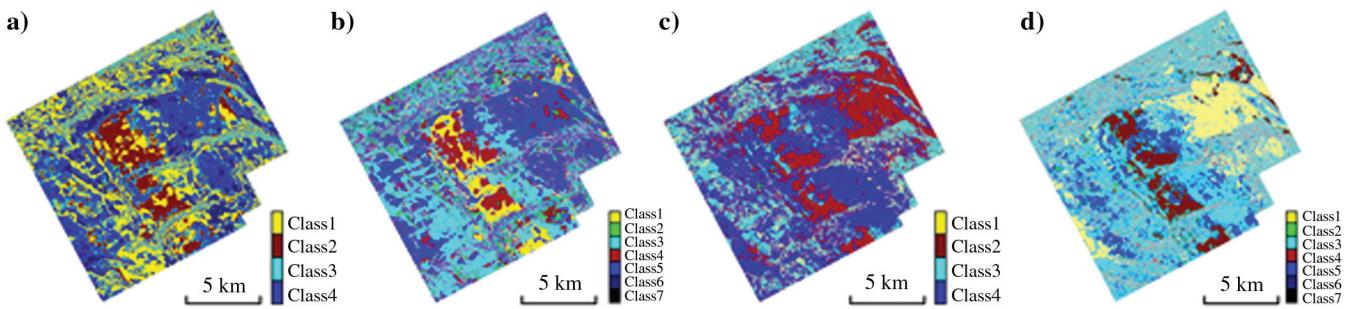


Figure 14. (a) K-means clustering results based on waveform in four categories. (b) K-means clustering results based on the waveform in seven categories. (c) K-means clustering results based on attributes in four categories. (d) K-means clustering results based on attributes in seven categories.

Conclusion

In this paper, a method of supervised classification of seismic facies has been proposed. This seismic facies classification can be realized through the division of optimized seismic attributes based on the framework of the spatial probability classification and the CNN algorithm. Here are some conclusions that we obtained for this method.

First, results of the artificial seismic facies analysis can be obtained depending on how experienced are the seismic interpreters. During this interpretation, we can establish the one-to-one corresponding relationship between the seismic facies and the sedimentary facies. Thus, we can get the seismic facies classification with a clearer sedimentary significance.

Second, comparing with some other classification algorithm mentioned in this paper, we demonstrated that CNN has better performance as a classifier in seismic multiattribute classification. The correct rate in the test data set reached approximately 84% in the application section. Moreover, by use of the maximum information coefficient to evaluate seismic attributes, we can minimize the redundancy of data while keep maximizing the information from the selected seismic attributes at the same time. Considering that the selected attributes still have a strong nonlinear relationship, using CNN's convolution, pooling, and dropout features can increase the algorithm's nonlinear expression ability and generalization effect.

Third, because there are many wrong classification results based on the output of single seismic trace, it may result in discrete classification for a continuous seismic facies, we use the probability information of the spatial distribution to improve it. When the probability information of the spatial distribution is used to restrict the classification result, the result will become more reasonable in a geologic view; e.g., the continuity within a seismic facies can be enhanced, and the boundary between two sedimentary facies becomes clear. The “Application” section demonstrates that this method is very effective for seismic facies classification in the study area with few drillings.

Acknowledgments

This research is supported by the key project of the National Natural Science Foundation of China (grant no. 41430323) and the National Natural Science Foundation of China Joint Fund Project (grant no. U1562219). We thank J. You and assistant professor Y. Xue for their help and suggestions. We appreciate the reviewers for their constructive suggestions.

Data and materials availability

Data associated with this research are available and can be obtained by contacting the corresponding author.

References

- Bagheri, M., and M. A. Riahi, 2013, Support vector machine based facies classification using seismic attributes in an oil field of Iran: Iranian Journal of Oil and Gas Science and Technology, **2**, 1–10.
- Camps-Valls, G., D. Tuia, L. Bruzzone, and J. A. Benediktsson, 2014, Advances in hyperspectral image classification: Earth monitoring with statistical learning methods: IEEE Signal Processing Magazine, **31**, 45–54, doi: [10.1109/MSP.2013.2279179](https://doi.org/10.1109/MSP.2013.2279179).
- Coléou, T., M. Poupon, and K. Azbel, 2003, Unsupervised seismic facies classification: A review and comparison of techniques and implementation: The Leading Edge, **22**, 942–953, doi: [10.1190/1.1623635](https://doi.org/10.1190/1.1623635).
- de Matos, M. C., P. L. Osorio, and P. R. Johann, 2006, Unsupervised seismic facies analysis using wavelet transform and self-organizing maps: Geophysics, **72**, no. 1, P9–P21, doi: [10.1190/1.2392789](https://doi.org/10.1190/1.2392789).
- Fauvel, M., Y. Tarabalka, J. A. Benediktsson, J. Chanussot, and J. C. Tilton, 2013, Advances in spectral-spatial classification of hyperspectral images: Proceedings of the IEEE, **101**, 652–675, doi: [10.1109/JPROC.2012.2197589](https://doi.org/10.1109/JPROC.2012.2197589).
- Gao, D., 2007, Application of three-dimensional seismic texture analysis with special reference to deep-marine facies discrimination and interpretation: An example from offshore Angola, West Africa: AAPG Bulletin, **91**, 1665–1683, doi: [10.1306/08020706101](https://doi.org/10.1306/08020706101).

- Halpert, A. D., R. G. Clapp, and B. Biondi, 2014, Salt delineation via interpreter-guided 3D seismic image segmentation: *Interpretation*, **2**, no. 2, T79–T88, doi: [10.1190/INT-2013-0159.1](https://doi.org/10.1190/INT-2013-0159.1).
- Hinton, G. E., and R. R. Salakhutdinov, 2006, Reducing the dimensionality of data with neural networks: *Science*, **313**, 504–507, doi: [10.1126/science.1127647](https://doi.org/10.1126/science.1127647).
- Huang, L., X. Dong, and T. E. Clee, 2017, A scalable deep learning platform for identifying geologic features from seismic attributes: *The Leading Edge*, **36**, 249–256, doi: [10.1190/tle36030249.1](https://doi.org/10.1190/tle36030249.1).
- Krizhevsky, A., I. Sutskever, and G. E. Hinton, 2012, Image-net classification with deep convolutional neural networks: *Advances in Neural Information Processing Systems*, 1097–1105.
- LeCun, Y., B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, and L. D. Jackel, 1989, Backpropagation applied to handwritten zip code recognition: *Neural Computation*, **1**, 541–551, doi: [10.1162/neco.1989.1.4.541](https://doi.org/10.1162/neco.1989.1.4.541).
- Liu, J., X. Dai, L. Gan, L. Liu, and W. Lu, 2017, Supervised seismic facies analysis based on image segmentation: *Geophysics*, **83**, no. 2, O25–O30, doi: [10.1190/geo2015-0539.1](https://doi.org/10.1190/geo2015-0539.1).
- Liu, J., and W. Lu, 2016, A probabilistic framework for spectral-spatial classification of hyperspectral images: *IEEE Transactions on Geoscience and Remote Sensing*, **54**, 5375–5384, doi: [10.1109/TGRS.2016.2562018](https://doi.org/10.1109/TGRS.2016.2562018).
- Lomask, J., R. G. Clapp, and B. Biondi, 2007, Application of image segmentation to tracking 3D salt boundaries: *Geophysics*, **72**, no. 4, P47–P56, doi: [10.1190/1.2732553](https://doi.org/10.1190/1.2732553).
- Perol, T., M. Gharbi, and M. Denolle, 2018, Convolutional neural network for earthquake detection and location: *Science Advances*, **4**, e1700578, doi: [10.1126/sciadv.1700578](https://doi.org/10.1126/sciadv.1700578).
- Qi, J., T. Lin, T. Zhao, F. Li, and K. Marfurt, 2016, Semisupervised multiattribute seismic facies analysis: *Interpretation*, **4**, no. 1, SB91–SB106, doi: [10.1190/INT-2015-0098.1](https://doi.org/10.1190/INT-2015-0098.1).
- Reshef, D. N., Y. A. Reshef, H. K. Finucane, S. R. Grossman, G. McVean, P. J. Turnbaugh, and P. C. Sabeti, 2011, Detecting novel associations in large data sets: *Science*, **334**, 1518–1524, doi: [10.1126/science.1205438](https://doi.org/10.1126/science.1205438).
- Roy, A., B. L. Dowdell, and K. J. Marfurt, 2013, Characterizing a Mississippian tripolitic chert reservoir using 3D unsupervised and supervised multiattribute seismic facies analysis: An example from Osage County, Oklahoma: *Interpretation*, **1**, no. 2, SB109–SB124, doi: [10.1190/INT-2013-0023.1](https://doi.org/10.1190/INT-2013-0023.1).
- Saraswat, P., and M. K. Sen, 2012, Artificial immune-based self-organizing maps for seismic-facies analysis: *Geophysics*, **77**, no. 4, O45–O53, doi: [10.1190/geo2011-0203.1](https://doi.org/10.1190/geo2011-0203.1).
- West, B. P., S. R. May, J. E. Eastwood, and C. Rossen, 2002, Interactive seismic facies classification using textural attributes and neural networks: *The Leading Edge*, **21**, 1042–1049, doi: [10.1190/1.1518444](https://doi.org/10.1190/1.1518444).
- Yuan, S., J. Liu, S. Wang, T. Wang, and P. Shi, 2018, Seismic waveform classification and first-break picking using convolution neural networks: *IEEE Geoscience and Remote Sensing Letters*, **15**, 272–276, doi: [10.1109/LGRS.2017.2785834](https://doi.org/10.1109/LGRS.2017.2785834).
- Zhang, B., S. Li, X. Jia, L. Gao, and M. Peng, 2011, Adaptive Markov random field approach for classification of hyperspectral imagery: *IEEE Geoscience and Remote Sensing Letters*, **8**, 973–977, doi: [10.1109/LGRS.2011.2145353](https://doi.org/10.1109/LGRS.2011.2145353).
- Zhang, Y., S. Jia, H. Huang, J. Qiu, and C. Zhou, 2014, A novel algorithm for the precise calculation of the maximal information coefficient: *Scientific Reports*, **4**, 6662, doi: [10.1038/srep06662](https://doi.org/10.1038/srep06662).
- Zhao, T., F. Li, and K. J. Marfurt, 2018, Seismic attribute selection for unsupervised seismic facies analysis using user-guided data-adaptive weights: *Geophysics*, **83**, no. 2, O31–O44, doi: [10.1190/geo2017-0192.1](https://doi.org/10.1190/geo2017-0192.1).

Biographies and photographs of the authors are not available.