Petroleum Science
https://doi.org/10.1007/s12182-020-00474-6

ORIGINAL PAPER

Lithofacies identification using support vector machine based on local deep multi-kernel learning

Xing-Ye Liu[1] · Lin Zhou[2] · Xiao-Hong Chen[3] · Jing-Ye Li[3]

Received: 1 November 2019

Abstract
Lithofacies identification is a crucial work in reservoir characterization and modeling. The vast inter-well area can be supplemented by facies identification of seismic data. However, the relationship between lithofacies and seismic information that is affected by many factors is complicated. Machine learning has received extensive attention in recent years, among which support vector machine (SVM) is a potential method for lithofacies classification. Lithofacies classification involves identifying various types of lithofacies and is generally a nonlinear problem, which needs to be solved by means of the kernel function. Multi-kernel learning SVM is one of the main tools for solving the nonlinear problem about multi-classification. However, it is very difficult to determine the kernel function and the parameters, which is restricted by human factors. Besides, its computational efficiency is low. A lithofacies classification method based on local deep multi-kernel learning support vector machine (LDMKL-SVM) that can consider low-dimensional global features and high-dimensional local features is developed. The method can automatically learn parameters of kernel function and SVM to build a relationship between lithofacies and seismic elastic information. The calculation speed will be expedited at no cost with respect to discriminant accuracy for multi-class lithofacies identification. Both the model data test results and the field data application results certify advantages of the method. This contribution offers an effective method for lithofacies recognition and reservoir prediction by using SVM.

Keywords Lithofacies discriminant · Support vector machine · Multi-kernel learning · Reservoir prediction · Machine learning

1 Introduction

Lithofacies identification is a critical content in stratigraphic correlation and sedimentary facies analysis. Lithofacies is an indispensable component of sedimentary facies, which can represent different lithologies or same lithology containing different types of fluids. Lithofacies prediction has a magnificent guiding significance for reservoir prediction and subsequent reservoir properties prediction. Accurate identification of lithofacies is conducive to the process of exploration, development and the stable production of resources (Xiong et al. 2010; Li et al. 2012). Delving into the abundant information about lithology and fluid contained in seismic data is helpful to improve lateral resolution and accuracy in the inter-well area. However, the relationship between lithofacies and seismic information is extremely complex and difficult to construct because multitudinous factors affect each other. As a result, it brings great challenges on seismic lithofacies identification (Liu et al. 2017; Huang et al. 2017a, b).

Traditional methods use rock physics and geostatistics to achieve reservoir characterization and prediction (Jalalalhosseini et al. 2014, 2015; Liu et al. 2018). Nowadays, machine learning has attracted wide attention in geoscience because of its advantages in addressing big data issues (e.g., Huang et al. 2016; Chen 2017, 2018; Chen et al. 2019). Neural network

Edited by Jie Hao

✉ Lin Zhou
linzhou6308@sina.com

[1] Shaanxi Provincial Key Laboratory of Geological Support for Coal Green Exploitation, College of Geology and Environment, Xi'an University of Science and Technology, Xi'an 710054, Shaanxi, China

[2] Hunan Provincial Key Laboratory of Shale Gas Resource Utilization, Hunan University of Science and Technology, Xiangtan 411201, Hunan, China

[3] State Key Laboratory of Petroleum Resources and Prospecting, National Engineering Laboratory for Offshore Oil Exploration, China University of Petroleum-Beijing, Beijing 102249, China

Published online: 30 June 2020

is one of the most used machine learning algorithms in geophysics (Kobrunov and Priezzhev 2016). However, it depends on the network topology, initial weights and thresholds and is sensitive to learning rate. Improper parameters will result in a slow convergence, falling to a local minimum, and subsequently outputting unsatisfactory predictions (Caruana and Niculescu-Mizil 2006). Support vector machine (SVM) is an effective machine learning method that aims at maximizing the distance from support vector to separating hyperplanes (Li et al. 2004; Zhang et al. 2005, 2018; Liu et al. 2020). It shows a powerful performance during the dispose of small-samples, nonlinearity and high-dimensional problems. SVM is guided by the maximum margin decision boundary and does not require the data to follow a specific distribution (Li et al. 2004; Vapnik 1999; Mou et al. 2015). Besides, it has relatively simple mathematical form with strong generalization ability (Suykens and Vandewalle 1999). The SVM method can be introduced to lithofacies identification (Abedi et al. 2012; Wang et al. 2016). Using SVM to identified lithofacies includes two steps. The first is training by employing training attributes (input) and known lithofacies (output) and obtaining the relationship between attributes and lithofacies. The second step is predicting lithofacies by inputting the attributes in target areas into decision function. Li et al. (2004) used SVM to recognize and predict reservoirs from seismic data, demonstrating the feasibility of SVM. Torres and Reveron (2013) integrated rock physics and simultaneous seismic inversion and successfully identified the reservoir zones by utilizing SVM in Orinoco Oil Belt, Venezuela. Zhao et al. (2015) compared the artificial neural network with SVM for lithofacies recognition and proved that SVM was mathematically more robust and easier to train. Besides, Zhao et al. (2014) introduced the proximal support vector machines (PSVM, see Fung and Mangasarian 2001, 2005; Mangasarian and Wild 2005) into lithofacies classification in Barnett Shale for saving computational cost, demonstrating the validity of PSVM classifier in binary classification between shale and limestone.

However, SVM is originally used to solve binary classification problems. Lithofacies identification is a multi-classification, high-dimensional and nonlinear problem. When solving the problem of nonlinear classification or prediction, using kernel functions maps the original data into a higher-dimensional feature space (Qu et al. 2019). Then, according to the training data, a classification hyperplane is established as a decision surface to separate the data belonging to different categories in the high-dimensional space (Zhu et al. 2015). Unfortunately, it is inadequate to distinguish the characteristics of different lithofacies only by using a single kernel under this situation. Otherwise, the identified accuracy will be impacted. The multi-kernel learning (MKL) is an available alternative with more flexibility than single-kernel function (Crammer and Singer 2001). Introducing the MKL method can improve the accuracy and stability of lithofacies discrimination based on SVM (Li et al. 2014). Through multi-kernel mapping, the high-dimensional space is divided into a combined space composed of several feature spaces. Then, each of the characteristic components can be embedded in the corresponding kernel function. MKL-SVM is less explored in lithofacies identification and reservoir prediction applications. Qin (2017) analyzed the methods and principles of commonly used techniques for lithofacies identification. Based on logging data, it was proved that MKL-SVM could enhance the accuracy of lithofacies classification. Cheng et al. (2018) successfully accomplished the lithofacies classification by the MKL-SVM method, but the calculation was time-consuming and the operation was cumbersome. The MKL algorithm is confronted with the dilemmas in selecting the appropriate of kernel function, determining the combination way of kernel functions and calculating the coefficients of each kernel function. In general, substantial kernel matrices jointly participate in the computation. This is the main reason why the computational dimension is large, and the spatial complexity and memory occupancy are high. Then, the computational time increases dramatically (Lin et al. 2007; Li et al. 2016). As a result, application of lithofacies identification based on MKL-SVM is severely limited in practice. GöNen and Alpaydin (2008) proposed a local multi-kernel learning (LMKL) SVM method that selected the appropriate kernel function locally to reduce the computational complexity. Although it effectively improves the sparsity of the kernels and reduces spatial complexity, it weakens the "complementarity" between the kernels and has serious parameter redundancy (GöNen and Alpaydın 2013). It also ignores the global characteristics of data. Generally, selecting proper kernel functions and determining their weights are difficult and depend heavily on experience. In addition, the computational efficiency of LMKL-SVM is inadequate (Ding 2014).

Jose et al. (2013) generalized LMKL to learn a tree-based primal feature that was high-dimensional and sparse and put forward a local deep multi-kernel learning (LDMKL) SVM method (Bengio et al. 2010). It could take both global and local features of data into account and facilitated the efficiency of multi-kernel learning while ensuring accuracy. This method focused on learning the best decision boundary in a sparse, high-dimensional representation, which could jointly learn both kernel and SVM parameters. Nevertheless, only a single-kernel function was used for global features and it was mainly aimed at the problem of binary classification. Following this research, we introduced LDKML-SVM into lithofacies classification and improved it for multi-class. Several low- and high-dimensional kernel functions are combined to distinguish the attributes of different lithofacies more accurately, so as to effectively build the relationship between lithofacies and training attributes. Using the local deep kernel function with tree structure can promote the computational efficiency. Taking the global features into account maintains the recognition accuracy. Automatic learning of kernel function parameters and decision parameters

avoids the deviation caused by artificial selection of parameters. The goals are to supply a promising new method for lithofacies identification and reservoir prediction, to overcome the weaknesses of the existing SVM-based lithofacies identification method and to promote the practicability of lithofacies classification based on SVM. Model data test and field data application verify validity of the proposed method.

2 Methodology

Compared with the LMKL-SVM, LDMKL-SVM is an improved method that can take into account the global and local characteristics of data at the same time. It concentrates on learning the best decision boundary in a sparse and high-dimensional representation. We introduced this method to lithofacies classification and extended it for multi-class facies. Multiple global kernel functions that are used to learn global features are set to be low-dimensional. The local kernel functions are composed of mapping functions that are tree-structured, high-dimensional and sparse. For that reason, the LDMKL-SVM method is conducive to improving the efficiency and accuracy when classifying lithofacies. The number of global and local kernel functions can be adjusted according to the sophistication of the problem. Another advantage of LDMKL-SVM is that it can learn parameters of kernel function and decision parameters of SVM at the same time. By learning the training data, we can establish relationship between training attributes and lithofacies using LDMKL algorithm. Inputting the measured data into the decision function of SVM realizes lithofacies recognition in other wells or in inter-well area. In the present method, the decision function of SVM can be expressed as (Jose et al. 2013):

$$\begin{aligned} y(\mathbf{x}) &= sign\left(\sum_i \alpha_i y_i K\left(\mathbf{x}, \mathbf{x}_i\right)\right) \\ &= sign\left(\sum_{ijk} \alpha_i y_i \phi_{G_j}\left(\mathbf{x}_i\right)\phi_{G_j}(\mathbf{x})\phi_{L_k}\left(\mathbf{x}_i\right)\phi_{L_k}(\mathbf{x})\right) \\ &= sign\left(\mathbf{w}^T\left(\boldsymbol{\Phi}_G(\mathbf{x}) \otimes \boldsymbol{\Phi}_L(\mathbf{x})\right)\right) \\ &= sign\left(\boldsymbol{\Phi}_L^T(\mathbf{x})\mathbf{W}^T\boldsymbol{\Phi}_G(\mathbf{x})\right) \\ &= sign\left(\mathbf{W}^T(\mathbf{x})\boldsymbol{\Phi}_G(\mathbf{x})\right) \end{aligned} \tag{1}$$

where $K\left(\mathbf{x}, \mathbf{x}_i\right) = \sum_{j,k} K_{Gj} K_{Lk}$ represents multi-kernel learning function, $\mathbf{x}_i$ represents the data, subscript $j = 1, \ldots, J$ denotes the jth global kernel function $K_G = \boldsymbol{\Phi}_G \otimes \boldsymbol{\Phi}_G$ and $K_L = \boldsymbol{\Phi}_L \otimes \boldsymbol{\Phi}_L$ are the global and local kernel functions, respectively. K_L consists of sparse and tree-structured mapping functions $\boldsymbol{\Phi}_L$ that contains high-dimensional local features. $\boldsymbol{\Phi}_G$ represents global mapping relations that contains low-dimensional global features. $\mathbf{w}_k = \sum_i \alpha_i y_i \phi_{L_k}\left(\mathbf{x}_i\right)\boldsymbol{\Phi}_G\left(\mathbf{x}_i\right)$, $k = 1, \ldots, M$ denotes the kth dimension of $\boldsymbol{\Phi}_L$, y_i represents the type of lithofacies, α_i is coefficient, $\mathbf{W} = [\mathbf{w}_1, \ldots, \mathbf{w}_k, \ldots, \mathbf{w}_M]$ and $\mathbf{W}(\mathbf{x}) = \mathbf{W}\boldsymbol{\Phi}_L(\mathbf{x})$.

In order to ensure computational efficiency, global features are usually low-dimensional. In this case, global mapping kernels can be set in linear and quadratic. To make prediction efficient, $\boldsymbol{\Phi}_L$ is tree-structured. Each dimension of $\boldsymbol{\Phi}_L$ corresponds to a node in the tree, and the dimension of $\boldsymbol{\Phi}_L(\mathbf{x})$ is nonzero only in the case that the corresponding node lies on the path traversed from the root to one of the leaves. Otherwise, it is equal to 0. Thus, for any locations, $\boldsymbol{\Phi}_L(\mathbf{x})$ has only log M nonzero dimensions, accelerating the computation (Jose et al. 2013). Figure 1 displays a four-layer tree structure schematic diagram. Only those dimensions of $\boldsymbol{\Phi}_L(\mathbf{x})$ are nonzero which correspond to the path traversed by $\mathbf{x}$ from the root to a leaf (as the black node shown in Fig. 1), which reduces the number of features to be calculated.

For the kth-dimensional local feature, its state value can be controlled by the indicator variable $I_k(\mathbf{x})$ (Jose et al. 2013),

$$I_k(\mathbf{x}) = \prod_{a \in ancestors(k)} \frac{1}{2}\left(\tanh\left(s_I \theta_a^T \mathbf{x}\right) + (-1)^{C(a)}\right) \tag{2}$$

where s_I represents a contraction factor. Jose et al. (2013) introduced a scale parameter θ_a^T to make tree learning amenable to sub-gradient descent and keep the sparsity. $C(a)$ is 0 if a node is its parent's left child and 1 if it is its parent's right child. The high nonlinear features are mainly embodied in the $\boldsymbol{\Phi}_L$. Jose et al. (2013) had tested the performance of local deep mapping functions with different forms and drew a conclusion that using the hyperbolic tangent function with a scale parameter $\theta_k'^T$ could output an excellent result. Following the conclusion, we utilized the hyperbolic tangent function to construct the local kernel function. The k^{th}-dimensional local feature mapping function is as follows:

$$\phi_{L_k} = \tanh\left(\sigma \theta_k'^T \mathbf{x}\right) I_k(\mathbf{x}) \tag{3}$$

where σ is a contraction factor and $\theta_k'^T$ is a scale parameter. $\boldsymbol{\Theta} = [\theta_1, \ldots, \theta_M]$ and $\boldsymbol{\Theta}' = [\theta_1', \ldots, \theta_M']$ denote the learning parameters that can be gained by solving the following objective function (Jose et al. 2013):

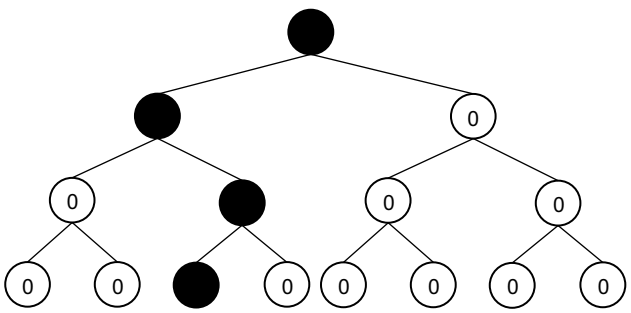

Fig. 1 Sketch Map of a four-dimensional local feature with tree-structured

$$\min_{\mathbf{W},\Theta,\Theta'} P(\mathbf{W},\Theta,\Theta') = \frac{\lambda_W}{2}\sum_k \|\mathbf{w}_k\|_2^2 + \frac{\lambda_\theta}{2}\sum_k \|\theta_k\|_2^2 + \frac{\lambda_{\theta'}}{2}\sum_k \|\theta'_k\|_2^2 + \sum_i L\left(y_i, \boldsymbol{\Phi}_L^T(\mathbf{x}_i)\mathbf{W}^T\mathbf{x}_i\right) \tag{4}$$

where λ_W , λ_θ , $\lambda_{\theta'}$ represent the coefficients, and L denotes the loss function. The loss function is broadly used to estimate the degree of inconsistency between the predicted value $y(\mathbf{x})$ and the real value y at $\mathbf{x}$. In order to extend the binary SVM to multi-class classification, the common practice is using one-vs-all or one-vs-one strategy (Duan and Keerthi 2005). In contrast, we introduced the multi-class loss function to solve the multi-class problem. There are several variations in multi-class loss functions. We use the loss function proposed by Crammer and Singer (2001):

$$L = \max\left(0, 1 + \max_{\widehat{y}_i \neq y_i}\left(\boldsymbol{\Phi}_L^T(\mathbf{x}_i)\mathbf{W}_{\left(\widehat{y}_i\right)}^T\mathbf{x}_i\right) - \boldsymbol{\Phi}_L^T(\mathbf{x}_i)\mathbf{W}_{(y_i)}^T\mathbf{x}_i\right) \tag{5}$$

where $\max_{\widehat{y}_i \neq y_i}\left(\boldsymbol{\Phi}_L^T(\mathbf{x}_i)\mathbf{W}_{\left(\widehat{y}_i\right)}^T\mathbf{x}_i\right)$ represents the estimated score of the real facies y_i misclassified into other lithofacies. The minimum of objective function (Eq. 4) can be found by the primal stochastic sub-gradient descent method, neglect of keeping dual variables or dual sparsity (see Jose et al. 2013; Orabona et al. 2010). Equation 4 contains terms about learning parameters of kernel functions and decision parameters of SVM. As a result, the method can jointly learn both kernel learning and SVM decision parameters. For the tth iteration of training point $\mathbf{x}_i$, the corresponding iteration update formula is:

$$\mathbf{W}^{(t+1)} = \mathbf{W}^{(t)} - \beta_t \nabla_{\mathbf{W}} P\left(\mathbf{W}^{(t)}, \Theta^{(t)}, \Theta'^{(t)}, \mathbf{x}_i\right) \tag{6}$$

$$\Theta^{(t+1)} = \Theta^{(t)} - \beta_t \nabla_{\Theta} P\left(W^{(t)}, \Theta^{(t)}, \Theta'^{(t)}, \mathbf{x}_i\right) \tag{7}$$

$$\Theta'^{(t+1)} = \Theta'^{(t)} - \beta_t \nabla_{\Theta'} P\left(W^{(t)}, \Theta^{(t)}, \Theta'^{(t)}, \mathbf{x}_i\right) \tag{8}$$

where β_t denotes iteration step,

$$\nabla_{\mathbf{w}_k} P(\mathbf{x}_i) = \lambda_W \mathbf{w}_k + \nabla_{\mathbf{w}_k}\left(\max_{\widehat{y}_i \neq y_i} \boldsymbol{\Phi}_L^T(\mathbf{x}_i)\mathbf{W}_{\left(\widehat{y}_i\right)}^T\mathbf{x}_i\right) - \phi_{L_k}(\mathbf{x}_i)\mathbf{x}_i \tag{9}$$

$$\nabla_{\theta_k} P(\mathbf{x}_i) = \lambda_\Theta \theta_k - \sum_a \tanh\left(\sigma\theta_a'^T\mathbf{x}_i\right)\nabla_{\theta_k} I_a(\mathbf{x}_i)\left(\mathbf{w}_{\left(\max_{\widehat{y}_i \neq y_i}\right)_a}^T\mathbf{x}_i - \mathbf{w}_{(y_i)a}^T\mathbf{x}_i\right) \tag{10}$$

$$\nabla_{\theta'_k} P(\mathbf{x}_i) = \lambda_{\Theta'}\theta'_k - \sigma\left[1 - \tanh^2\left(\sigma\theta_k'^T\mathbf{x}_i\right)\right] I_k(\mathbf{x}_i)\left(\mathbf{w}_{\left(\max_{\widehat{y}_i \neq y_i}\right)_k}^T - \mathbf{w}_{(y_i)k}^T\right)\mathbf{x}_i\mathbf{x}_i \tag{11}$$

Solving the objective function iteratively, we can acquire parameters of local deep multi-kernel learning and SVM decision parameters. For the seismic lithofacies classification problem with large amount of data, it can also obtain discriminant lithofacies with appropriate accuracy while ensuring the computational efficiency.

3 Test on model data

First, we tested the present method on a modified fluvial channel system from the Stanford V reservoir model (Mao and Journel 1999), which had 120 CDPs and 100 CDPs in the horizontal direction of X and Y, with interval of 25 meters. The curved channel model includes channel, point bar, natural dike and flood plain subfacies. The lithofacies developed in different sedimentary facies are discrepant. The floodplain, channel, natural dike and point bar are sandstone, mudstone, sandy mudstone and siltstone deposit, respectively. Extract some traces from the original model as training data. Test is done on the whole model, and here, we show a slice of it to contrast clearly. Figure 2 shows lithofacies of one slice in the model and locations of the training data. Training attributes include density and P-wave velocity (displayed in Fig. 3). Figure 3c, d shows the probability density distribution of elastic attributes in different facies. It should be mentioned that the elastic

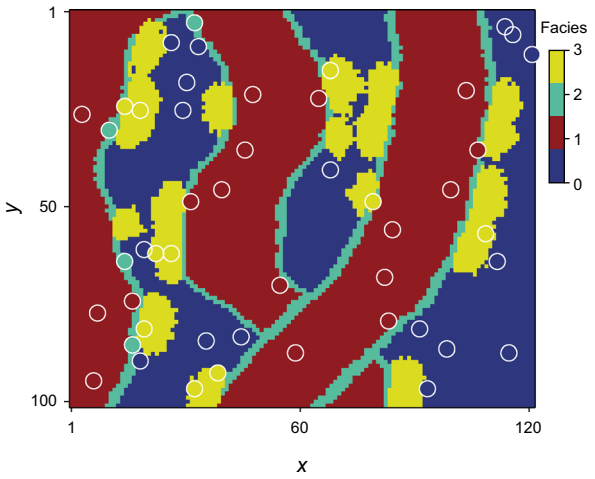

Fig. 2 One slice of the test model, which develops channels, natural dike, point bar and floodplain. Different colors represent different facies. 0, 1, 2 and 3 represent mudstone, sandstone, sandy mudstone and siltstone, respectively. White circles represent training data locations

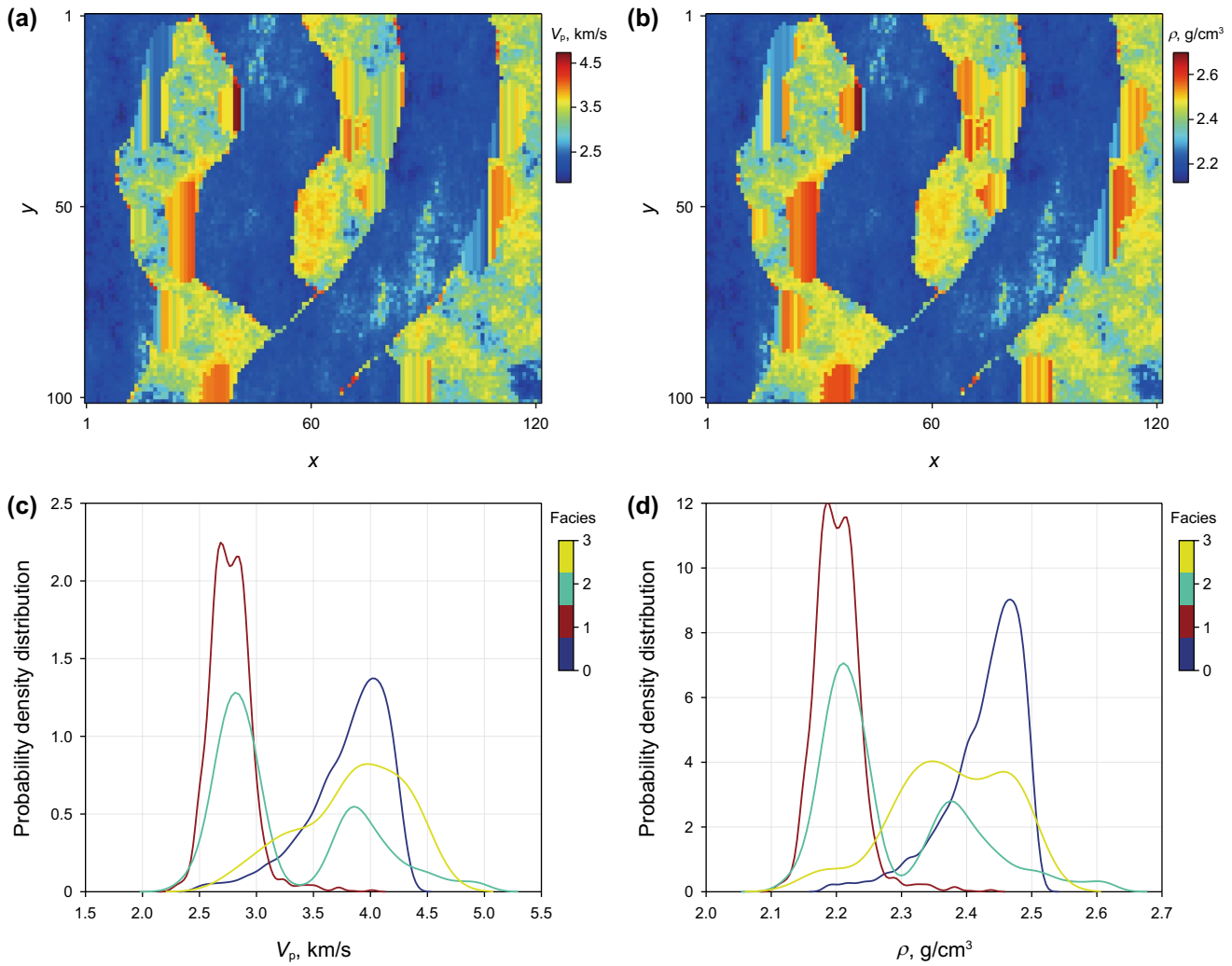

Fig. 3 Elastic attributes of the test model (**a**, **b**) and their probability density distribution (**c**, **d**). (a, c) P-wave velocity; (b, d) density. 0,1,2,3 represent mudstone, sandstone, sandy mudstone and siltstone, respectively. The difference of elastic properties between mudstone and sandstone is remarkable and easy to distinguish, while attribute values of sandy mudstone and siltstone overlap with those of mudstone and sandstone, which brings about challenges to discriminate them

attributes features of siltstone and sandy mudstone are distributed between that of mudstone and sandstone, and their attribute values overlap with sandstone or mudstone in a large range. The relationship between lithofacies and elastic attributes is established by the present method and LMKL-SVM method, respectively. Figure 4a exhibits the discriminated lithofacies using LDMKL-SVM. It can be noted that this method can better identify channel sandstone and floodplain mudstone, while the siltstone and sandy mudstone cannot be fully recognized. Figure 4b shows the discriminated lithofacies by LMKL-SVM. Some locations whose lithofacies are sandstone in model are identified as sandy mudstone. In addition, the lithofacies that should be siltstone and sandy mudstones at some locations are not consistent with the model facies. Although both the two methods closely follow the defined facies and show the promising results, the proposed method produces a more accurate result, which suggests that its ability to classify these facies is stronger than the conventional method. In order to compare and evaluate the performance of the two methods more intuitively and quantitatively, we count the confusion matrices of the two methods (shown

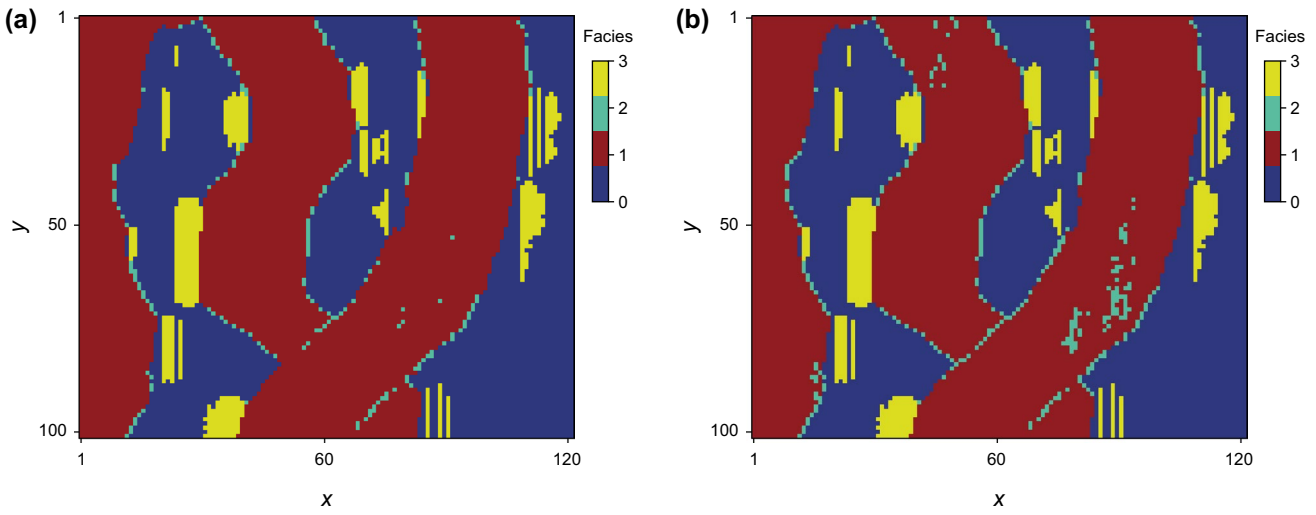

Fig. 4 Classified results of the test model by different methods: **a** identified lithofacies by LDMKL-SVM; **b** identified lithofacies by LMKL-SVM. By contrast, the former is superior to the latter. Mudstone and sandstone are well recognized. But the classification effect of the other two lithofacies is unsatisfactory

Table 1 Confusion matrix of the test model with the LDMKL-SVM method

	Defined mudstone, %	Defined sandstone, %	Defined sandy mudstone, %	Defined siltstone, %
Predicted mudstone	99.97	0.00	0.00	0.03
Predicted sandstone	0.00	99.91	0.09	0.00
Predicted sandy mudstone	0.00	74.46	25.54	0.00
Predicted siltstone	54.36	0.00	0.00	45.64

Table 2 Confusion matrix of the test model with the LMKL-SVM method

	Defined mudstone, %	Defined sandstone, %	Defined sandy mudstone, %	Defined siltstone, %
Predicted mudstone	99.97	0.00	0.00	0.03
Predicted sandstone	0.00	96.40	3.60	0.00
Predicted sandy mudstone	0.00	74.15	25.85	0.00
Predicted siltstone	60.05	0.00	0.00	39.95

in Tables 1 and 2). It is found that the accuracy of the LMKL-SVM is slightly lower than that of the proposed method (about 2.03%). Obviously, the recognition accuracy of the proposed method for different lithofacies is also higher than that of LMKL-SVM. For the whole test model, the execution time of LMKL-SVM is 1674.81 s and is longer than that of the proposed method (354.99 s). This demonstrates the effectiveness and superiority of the lithofacies identification method based on LDMKL-SVM.

4 Application on field data

In order to further verify the validity of lithofacies discriminant method based on LDMKL-SVM, we applied this novel method to actual land logging and 2D seismic data. The filed data are from a China work area. Figure 5 displays the test seismic section that goes through two wells (Well A and B). According to the preliminary study and log interpretation results, the main lithofacies in the study area

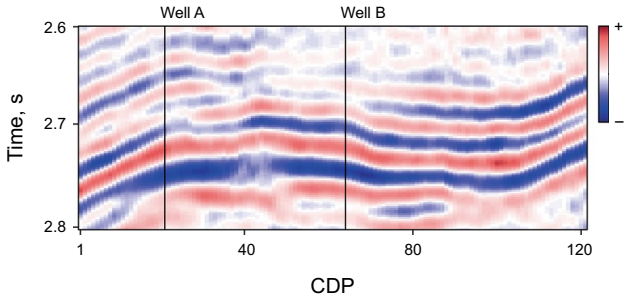

Fig. 5 2D seismic stack profile. Black lines denote locations of the two wells. Black arrows indicate target reservoirs

includes mudstone, water-bearing sandstone and oil-bearing sandstone. Black lines denote locations of the two wells, and the distance between the two wells is 975 m. Black arrows in Fig. 5 indicate the target reservoirs.

4.1 Lithofacies discriminant in wells

Experiments are carried out in wells at first. We borrow density and P- and S-wave velocity as training attributes. Figures 6 and 7 exhibit logging curves of Wells A and B. Figure 8 displays the distribution of elastic attributes in the two wells. All logging data of Well A are used as training data to identify lithofacies of Well A and Well B by the present method and LMKL-SVM, respectively. Figures 9 and 10 show interpretation lithofacies and identified results of Well A and Well B, respectively. The result produced by the present method has an excellent consistency with the interpretation lithofacies and can accurately pick out the oil-bearing sandstone. The performance of the proposed

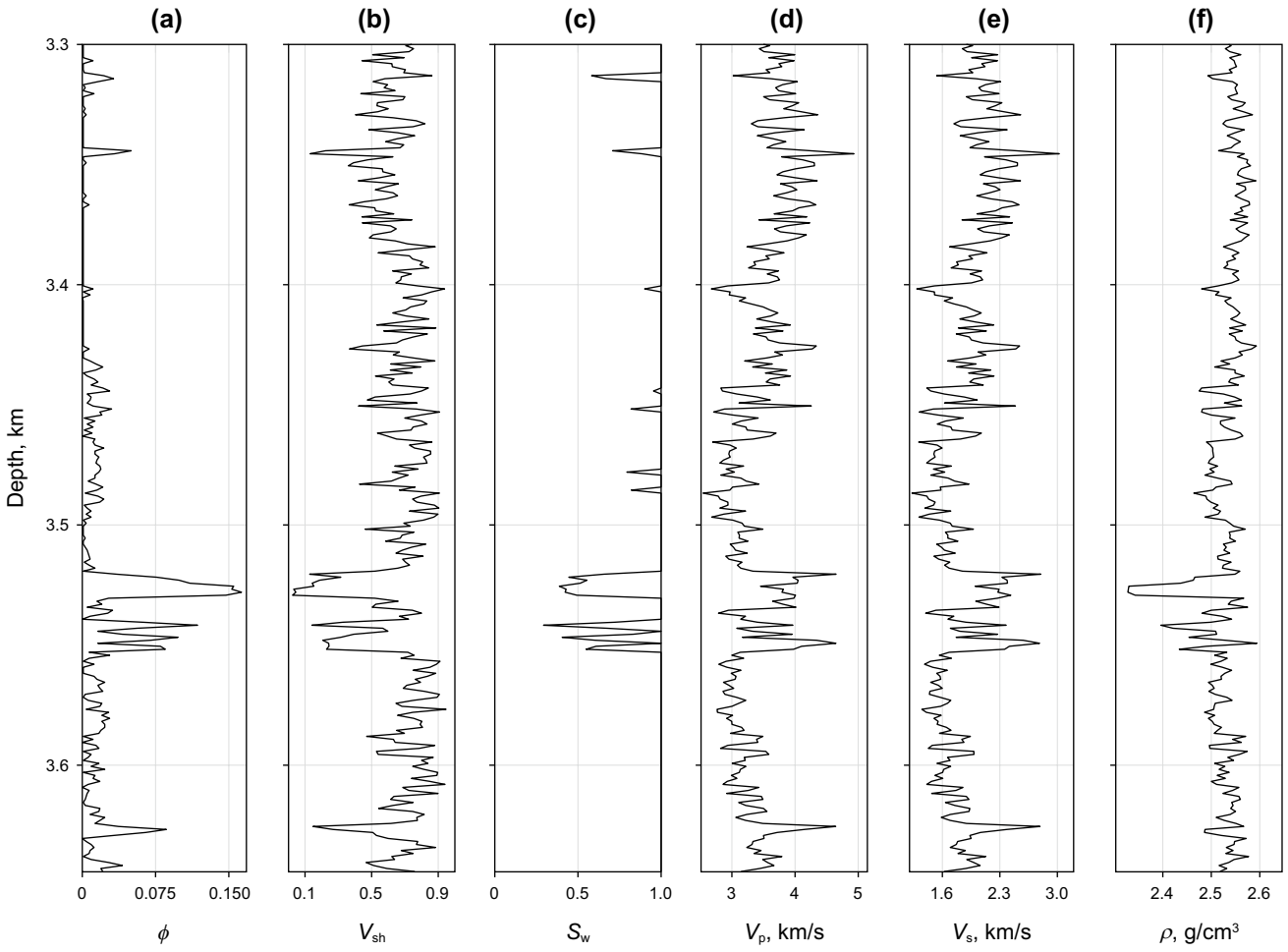

Fig. 6 Logging curves of Well A: **a** porosity; **b** shale content; **c** water saturation; **d** P-wave velocity; **e** S-wave velocity; and **f** density

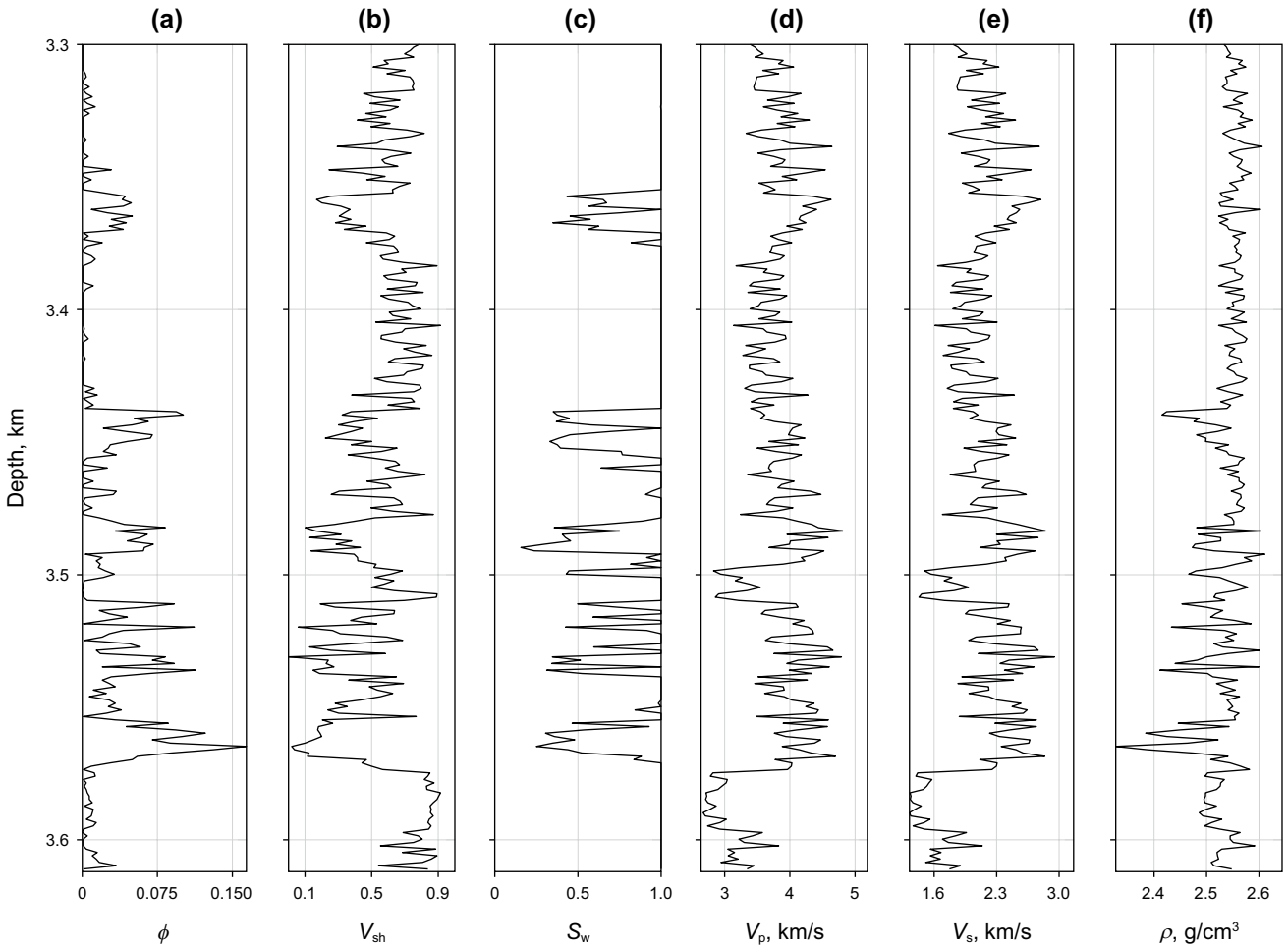

Fig. 7 Logging curves of the Well B: **a** porosity; **b** shale content; **c** water saturation; **d** P-wave velocity; **e** S-wave velocity; and **f** density

method is better than that of the traditional method. Statistical analysis of the confusion matrix exhibited in Table 3 and Table 4 indicates that the misjudgment rate of the proposed method is only 4.87% (Well A) and 6.69% (Well B), while that of the traditional method is 7.89% and 14.68%, correspondingly. The accuracy of each type lithofacies is higher than that of LMKL-SVM. That is, the accuracy is improved by jointly considering the global low- and local high-dimensional features, indicating the dependability of the method quantitatively.

4.2 Seismic facies classification

Ultimately, we applied this novel method to the 2D seismic profile. The characteristics of the geological structure vary with different measurement methods at different scales. Because the observation scales of logging data and seismic data are different, logging data are first coarsened to seismic scale by Backus averaging. The elastic attributes of the area to be predicted are obtained by the prestack seismic inversion method, as shown in Fig. 11. With coarsened well data (well A and well B) as training data, classified lithofacies by

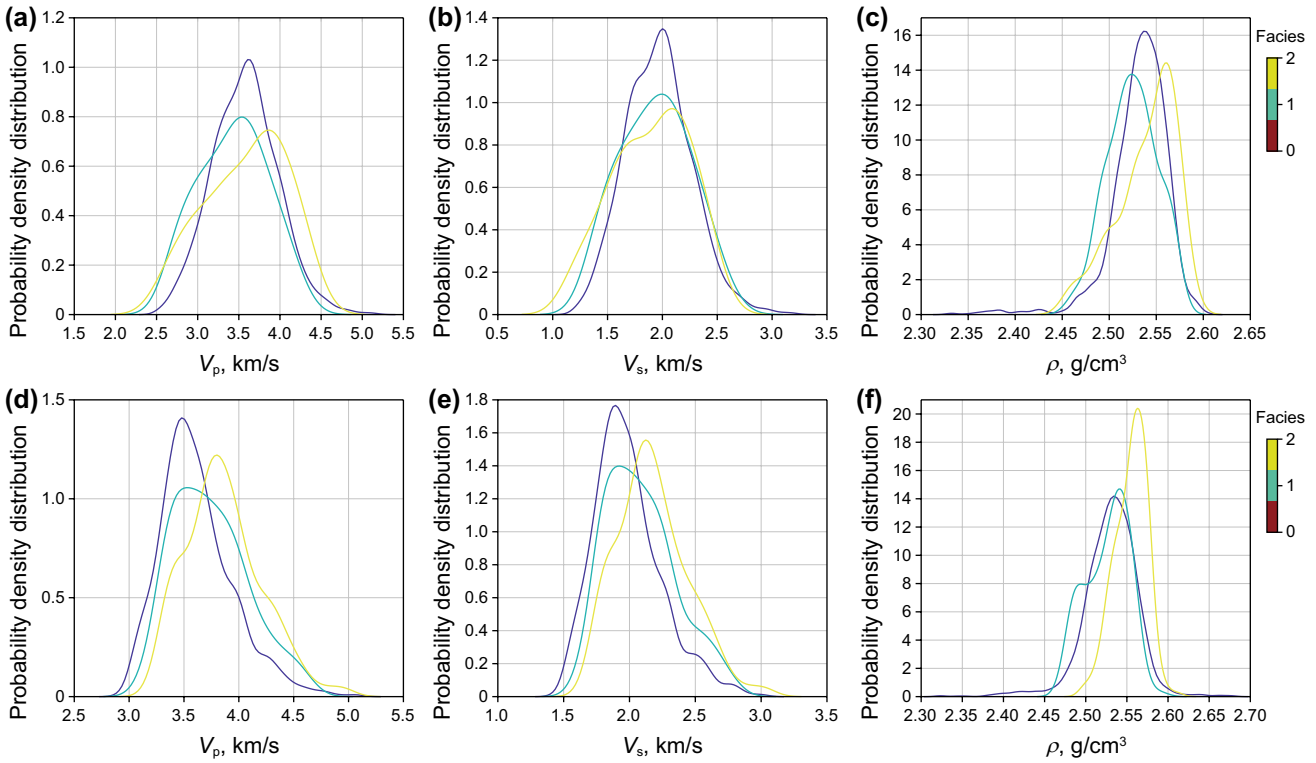

Fig. 8 Probability density distribution of elastic attributes in Wells A and B. **a**, **d** P-wave velocity; **b**, **e** S-wave velocity; **c**, **f** density. The above figures are about Well A, and the below figures are about Well B. 0, 1, 2 represent mudstone, water sandstone and oil sandstone, respectively

different methods is displayed in Fig. 12. It can be noted that both the two methods can effectively identify sandstone and mudstone in the inter-well area and distinguish oil-bearing sandstone from water-bearing sandstone. In this application, we cannot compare the accuracy of the two methods because the true seismic facies is unknown. However, LDMKL-SVM can speed up prediction time over LMKL-SVM. Their running time is 107.5 s and 418.2 s, respectively. Predicted lithofacies of the traces near wells are almost in agreement with the coarsened well facies. The identified target reservoirs follow the actual situation, which also suggests the good performance and practical value of the new method. It can provide reliable information for reservoir prediction and subsequent research. In addition, this method can be applied to any case where lithology and fluid need to be recognized.

5 Conclusion

We describe a new lithofacies identification method based on SVM. The present method draws support from a composite nonlinear kernel consisted of high-dimensional, sparse and computationally deep local features and low-dimensional global features. This method can be used to classify multiple types of lithofacies. One advantage is that it can automatically learn the parameters of kernel functions and SVM at the same time, avoiding the weakness of the traditional lithofacies discriminant method based on SVM. Another potential advantage is that the new method can effectively improve the classification accuracy. At the same time, the computing cost is saved because the local high-dimensional features are

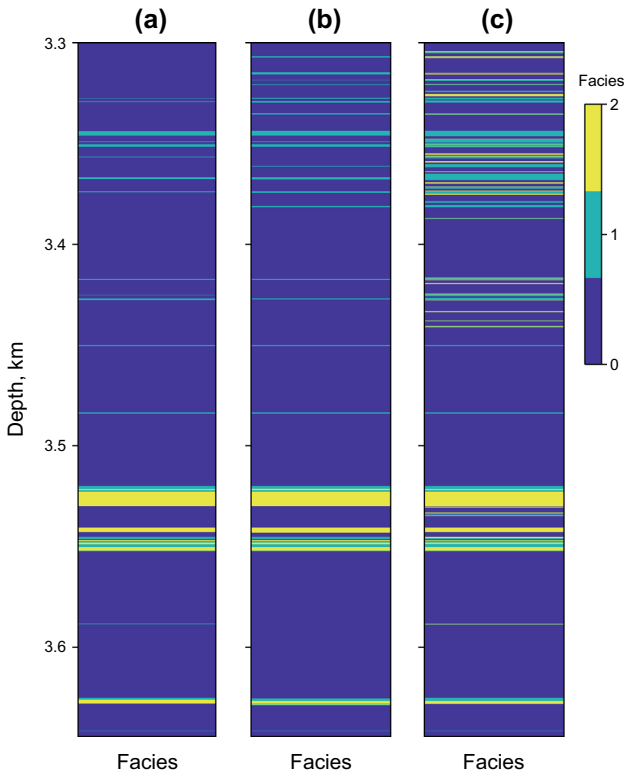

Fig. 9 Defined lithofacies **a** and identified facies by different methods of the test Well A: **b** classified by the proposed method; **c** classified by the traditional method. 0, 1, 2 represent mudstone, water sandstone and oil sandstone, respectively. Because of the higher complexity of the field data than model data, learning of global and high-dimensional local features can better distinguish the attributes of different lithofacies and the accuracy is promoted

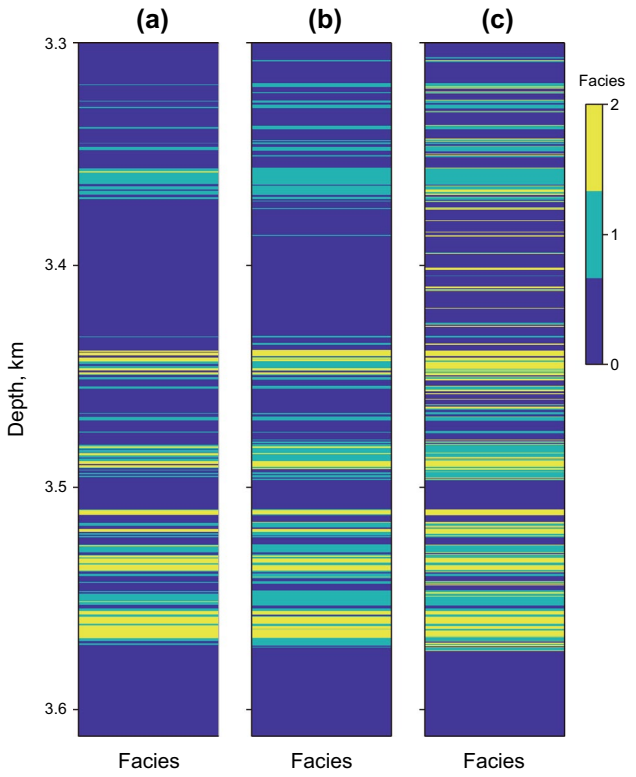

Fig. 10 Defined lithofacies **a** and identified facies by different methods of the test Well B. **b** classified by the proposed method; and **c** classified by the traditional method. 0, 1, 2 represent mudstone, water sandstone and oil sandstone, respectively. Comparing diagrams **b** with **c**, it can be found that the new method improves the accuracy and has a better consistency with the definition of lithofacies

sparse and tree-structured. The numerical example and filed data application illustrate that the proposed method can generate a preferable identified result in a relatively short time. The comparison with traditional methods also calibrates the superiority of the lithofacies discriminant method based on LDMKL-SVM. Profiting from the high accuracy and computation efficiency, a valuable approach is provided for the practical application of seismic lithofacies identification. It is also of great significance to the exploration and development of reservoirs and has good application prospects.

Table 3 Confusion matrix of Well A with different methods

Method		Defined mudstone	Defined water stone, %	Defined oil stone, %
Predicted lithofacies by LDMKL-SVM	Mudstone	97.39%	2.51	0.10
	Water stone	6.38%	92.20	1.42
	Oil stone	0.00%	4.20	95.8
Predicted lithofacies by traditional LMKL-SVM	Mudstone	90.12	5.98	3.90
	Water stone	0.00%	92.91	5.67
	Oil stone	0.00%	6.72	93.28

Table 4 Confusion matrix of Well B with different methods

Method		Defined mudstone, %	Defined water stone, %	Defined oil stone, %
Predicted lithofacies by LDMKL-SVM	Mudstone	91.90	7.16	0.94
	Water stone	1.58	98.15	0.27
	Oil stone	0.00	10.13	89.87
Predicted lithofacies by traditional LMKL-SVM	Mudstone	81.64	9.27	9.09
	Water stone	0.00	90.77	9.23
	Oil stone	0.00	16.46	83.54

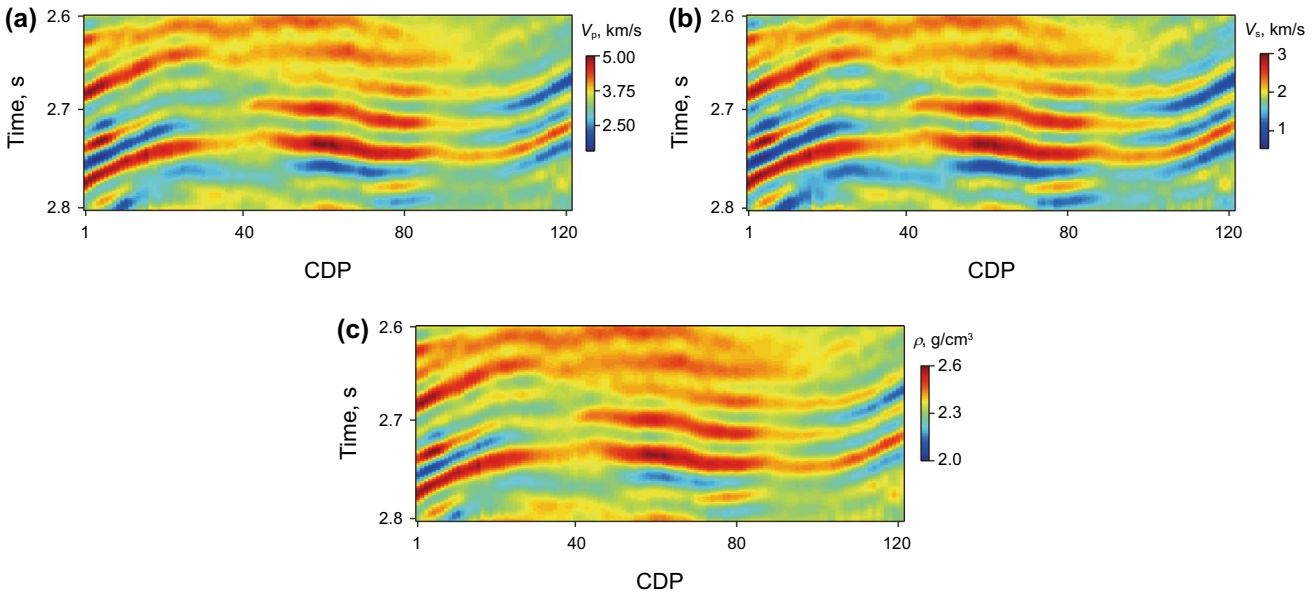

Fig. 11 Profiles of seismic elastic attributes obtained by prestack seismic inversion technique. **a** P-wave velocity; **b** S-wave velocity; and **c** density. The inverted attributes reflect the shapes and characteristics of strata in the study area. A good inversion result of the elastic attribute is helpful to identify the corresponding lithofacies

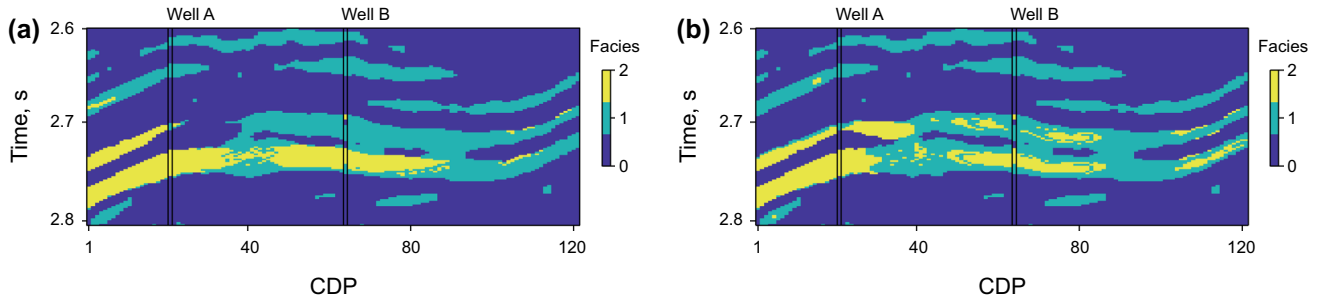

Fig. 12 Classified lithofacies of the field data using different methods. **a** LDMKL-SVM; **b** traditional LMKL-SVM. 0, 1, 2 represent mudstone, water sandstone and oil sandstone, respectively. Different lithofacies are well distinguished, and their shapes are similar to of the strata. Target reservoirs are also well identified

Acknowledgements This work is financially supported by the National Natural Science Foundation of China (41774129, 41904116), the Foundation Research Project of Shaanxi Provincial Key Laboratory of Geological Support for Coal Green Exploitation (MTy2019-20). Comments by the journal's editors and the associated reviewers are very helpful in improving the manuscript. Finally, we would like to thank the reviewers and the editors.

References

Abedi M, Norouzi GH, Bahroudi A. Support vector machine for multi-classification of mineral prospectivity areas. Comput Geosci. 2012;46:272–83. https://doi.org/10.1016/j.cageo.2011.12.014.

Bengio Y, Delalleau O, Simard C. Decision trees do not generalize to new variations. Comput Intell. 2010;26(4):449–67. https://doi.org/10.1111/j.1467-8640.2010.00366.x.

Caruana R, Niculescu-Mizil A. An empirical comparison of supervised learning algorithms. In: Proceedings of the 23rd international conference on machine learning. 2006; 161–68. https://doi.org/10.1145/1143844.1143865.

Chen YK. Automatic microseismic event picking via unsupervised machine learning. Geophys J Int. 2017;212(1):88–102. https://doi.org/10.1093/gji/ggx420.

Chen YK. Fast waveform detection for microseismic imaging using unsupervised machine learning. Geophys J Int. 2018;215(2):1185–99. https://doi.org/10.1093/gji/ggy348.

Chen YK, Zhang M, Bai M, et al. Improving the signal-to-noise ratio of seismological datasets by unsupervised machine learning. Seismol Res Lett. 2019. https://doi.org/10.1785/0220190028.

Cheng JW, Chen XH, Liu XY, et al. Lithofacies discrimination based on adaptive kernel function of support vector machines. In: 80th EAGE conference and exhibition. 2018. https://doi.org/10.3997/2214-4609.201800921.

Crammer K, Singer Y. On the algorithmic implementation of multiclass kernel-based vector machines. J Mach Learn Res. 2001;2(12):265–92. https://doi.org/10.1162/15324430260185628.

Ding Y. A kind of improved localized multiple kernel learning algorithm. Chongqing Technol Business Univ (in Chinese). 2014;31(11):56–61. https://doi.org/10.3969/j.issn.1672-058X.2014.11.012.

Duan KB and Keerthi SS. Which is the best multiclass SVM method? an empirical study. In: International workshop on multiple classifier systems. Springer, Berlin, Heidelberg. 2005;278–85. https://doi.org/10.1007/11494683_28.

Fung G, Mangasarian OL. Proximal support vector machine classifiers. In: 7th ACM Sigkdd international conference on knowledge discovery & data mining. 2001. https://doi.org/10.1145/502512.502527.

Gönen M and Alpaydin E. Localized multiple kernel learning. In: Proceedings of the 25th international conference on machine learning. 2008; 352–59. https://doi.org/10.1145/1390156.1390201.

GöNen M, AlpaydıN E. Localized algorithms for multiple kernel learning. Pattern Recognit. 2013;46(3):795–807. https://doi.org/10.1016/j.patcog.2012.09.002.

Huang WL, Wang R, Yuan Y, et al. Signal extraction using randomized-order multichannel singular spectrum analysis. Geophysics. 2016;82(2):V69–84. https://doi.org/10.1190/geo2015-0708.1.

Huang W, Wang R, Gong X, et al. Iterative deblending of simultaneous-source seismic data with structuring median constraint. IEEE Geosci Remote Sens Lett. 2017a;15(1):58–62. https://doi.org/10.1109/LGRS.2017.2772857.

Huang WL, Wang RQ, Chen XH, et al. Double least-squares projections method for signal estimation. IEEE Trans Geosci Remote Sens. 2017b;55(7):4111–29. https://doi.org/10.1109/TGRS.2017.2688420.

Jalalalhosseini SM, Ali H, Mostafazadeh M. Predicting porosity by using seismic multi-attributes and well data and combining these available data by geostatistical methods in a South Iranian oil field. Pet Sci Technol. 2014;32(1):29–37. https://doi.org/10.1080/10916466.2011.584102.

Jalalalhosseini SM, Eskandari S, Mortezazadeh E. The technique of seismic inversion and use of the relation between inversion results and porosity log for predicting porosity of a carbonate reservoir in a south Iranian oil field. Energy Sour Part A Recovery Utilization Environ Effects. 2015;37(3):265–72. https://doi.org/10.1080/15567036.2011.580326.

Jose C, Goyal P, Aggrwal P, et al. Local deep kernel learning for efficient non-linear svm prediction. In: International conference on

machine learning. 2013;486–494. https://doi.org/10.1145/2623330.2623759.
Kobrunov A, Priezzhev I. Hybrid combination genetic algorithm and controlled gradient method to train a neural network. Geophysics. 2016;81(4):IM35–43. https://doi.org/10.1190/geo2015-0297.1.
Li J, Castagna J, Li DA, Bian X. Reservoir prediction via SVM pattern recognition. In: 74th SEG technical program expanded abstracts. 2004; 425–8. https://doi.org/10.1190/1.1845255.
Li X, Mao W, Jiang W. Multiple-kernel-learning-based extreme learning machine for classification design. Neural Comput Appl. 2016;27(1):175–84. https://doi.org/10.1016/j.neucom.2013.09.072.
Li X, Zhou J, Li H, et al. Computational intelligent methods for predicting complex ithologies and multiphase fluids. Pet Explor Develop. 2012;39(2):261–7. https://doi.org/10.1016/s1876-3804(12)60041-x.
Li Y, Wen D, Wang K, et al. Multiple kernel MtLSSVM and its application in lung nodule recognition. J Jilin Univ (in Chinese). 2014;44(2):508–15. https://doi.org/10.13229/j.cnki.jdxbgxb201402037.
Lin YY, Liu TL, Fuh CS. Local ensemble kernel learning for object category recognition. In: 2007 IEEE Conference on computer vision and pattern recognition. 2007; 1–8. https://doi.org/10.1109/cvpr.2007.383084.
Liu XY, Li JY, Chen XH, et al. Bayesian discriminant analysis of lithofacies integrate the Fisher transformation and the kernel function estimation. Interpretation. 2017;5(2):SE1–10. https://doi.org/10.1109/cvpr.2007.383084.
Liu XY, Li JY, Chen XH, et al. Stochastic inversion of facies and reservoir properties based on multi-point geostatistics. J Geophys Eng. 2018;15(6):2455–68. https://doi.org/10.1190/INT-2016-0025.1.
Liu XY, Chen XH, Li JY, et al. Facies identification based on multi-kernel relevance vector machine. IEEE Trans Geosci Remote Sens. 2020;58(10):1–14. https://doi.org/10.1109/TGRS.2020.2981687.
Mangasarian OL, Wild EW. Multisurface proximal support vector machine classification via generalized eigenvalues. IEEE Trans Pattern Anal Mach Intell. 2005;28(1):69–74. https://doi.org/10.1109/TPAMI.2006.17.
Mao S, Journel AG. Generation of a reference petrophysical/seismic data set. In: The Stanford V reservoir. In Stanford center for reservoir forecasting annual meeting. SCRF Report, Stanford University. 1999.
Mou D, Wang Z, Huang Y, et al. Lithological identification of volcanic rocks from SVM well logging data: case study in the eastern depression of Liaohe Basin. Chin J Geophys (in Chinese). 2015;58(5):1785–93. https://doi.org/10.6038/cjg20150528.
Orabona F, Jie L, Caputo B. Online-batch strongly convex multi kernel learning. In: 2010 IEEE computer society conference on computer vision and pattern recognition. 2010; 787–794. https://doi.org/10.1109/cvpr.2010.5540137.
Qin Y. Application of multi-kernel function method in reservoir lithology identification Northeast Petroleum University (Daqing). 2017.
Qu S, Guan Z, Verschuur E, et al. Automatic high-resolution microseismic event detection via supervised machine learning. Geophys J Int. 2019;218(3):2106–21. https://doi.org/10.1093/gji/ggz273.
Suykens JA, Vandewalle J. Least squares support vector machine classifiers. Neural Process Lett. 1999;9(3):293–300. https://doi.org/10.1023/a:1018628609742.
Torres A, Reveron J. Lithofacies discrimination using support vector machines, rock physics and simultaneous seismic inversion in clastic reservoirs in the Orinoco Oil Belt, Venezuela. In: 83rd SEG technical program expanded abstracts. 2013; 2578–2582. https://doi.org/10.1190/segam2013-0039.1.
Vapnik VN. An overview of statistical learning theory. IEEE Trans Neural Networks. 1999;10(5):988–99. https://doi.org/10.1109/72.788640.
Wang P, Wang Z, Ni N. Identification of the lithology in tight sandstone reservoir in sulige gas field based on SVM optimized by cross validation. Chin Manganese Industry (in Chinese). 2016;34(6):53–6. https://doi.org/10.14101/j.cnki.issn.1002-4336.2016.06.017.
Xiong W, Wan ZH, Chen MS, et al. Semi-automatic determination of the number of seismic facies in waveform classification. In: 72nd EAGE conference and exhibition incorporating SPE EUROPEC 2010. https://doi.org/10.15258/sst.2008.36.3.10.
Zhang G, Wang Z, Chen Y. Deep learning for seismic lithology prediction. Geophys J Int. 2018;215(2):1368–87. https://doi.org/10.1093/gji/ggy344.
Zhang Z, Ye H, Wang G, et al. Leak detection in transport pipelines using enhanced independent component analysis and support vector machines. In: International conference on natural computation. 2005; 95–100. https://doi.org/10.1007/11539117_16.
Zhao T, Jayaram V, Marfurt KJ, et al. Lithofacies classification in Barnett Shale using proximal support vector machines. In: 84th SEG technical program expanded abstracts. 2014; 1491–1495. https://doi.org/10.1190/segam2014-1210.1.
Zhao T, Jayaram V, Roy A, et al. A comparison of classification techniques for seismic facies recognition. Interpretation. 2015;3(4):SAE29–58. https://doi.org/10.1190/int-2015-0044.1.
Zhu YF, Tian LF, Mao ZY, et al. Mixtures of kernels for SVM modeling. In: International conference on natural computation. 2015; 601–607. https://doi.org/10.1007/11539087_76.