

Original Paper

Working condition recognition of sucker rod pumping system based on 4-segment time-frequency signature matrix and deep learning



Yun-Peng He ^{a, b, c, d}, Hai-Bo Cheng ^{a, b, c, *}, Peng Zeng ^{a, b, c, **,} Chuan-Zhi Zang ^e,
Qing-Wei Dong ^{a, b, c, d}, Guang-Xi Wan ^{a, b, c, d}, Xiao-Ting Dong ^{a, b, c, d}

^a State Key Laboratory of Robotics, Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang 110016, Liaoning, China

^b Key Laboratory of Networked Control Systems, Chinese Academy of Sciences, Shenyang 110016, Liaoning, China

^c Institutes for Robotics and Intelligent Manufacturing, Chinese Academy of Sciences, Shenyang 110169, Liaoning, China

^d University of Chinese Academy of Sciences, Beijing, 100049, China

^e School of Artificial Intelligence, Shenyang University of Technology, Shenyang 110870, Liaoning, China

ARTICLE INFO

Article history:

Received 4 April 2023

Received in revised form

12 June 2023

Accepted 28 August 2023

Available online 28 August 2023

Edited by Jia-Jia Fei and Min Li

Keywords:

Sucker-rod pumping system

Dynamometer card

Working condition recognition

Deep learning

Time–frequency signature

Time-frequency signature matrix

ABSTRACT

High-precision and real-time diagnosis of sucker rod pumping system (SRPS) is important for quickly mastering oil well operations. Deep learning-based method for classifying the dynamometer card (DC) of oil wells is an efficient diagnosis method. However, the input of the DC as a two-dimensional image into the deep learning framework suffers from low feature utilization and high computational effort. Additionally, different SRPSs in an oil field have various system parameters, and the same SRPS generates different DCs at different moments. Thus, there is heterogeneity in field data, which can dramatically impair the diagnostic accuracy. To solve the above problems, a working condition recognition method based on 4-segment time-frequency signature matrix (4S-TFSM) and deep learning is presented in this paper. First, the 4-segment time-frequency signature (4S-TFS) method that can reduce the computing power requirements is proposed for feature extraction of DC data. Subsequently, the 4S-TFSM is constructed by relative normalization and matrix calculation to synthesize the features of multiple data and solve the problem of data heterogeneity. Finally, a convolutional neural network (CNN), one of the deep learning frameworks, is used to determine the functioning conditions based on the 4S-TFSM. Experiments on field data verify that the proposed diagnostic method based on 4S-TFSM and CNN (4S-TFSM-CNN) can significantly improve the accuracy of working condition recognition with lower computational cost. To the best of our knowledge, this is the first work to discuss the effect of data heterogeneity on the working condition recognition performance of SRPS.

© 2023 The Authors. Publishing services by Elsevier B.V. on behalf of KeAi Communications Co. Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

Sucker rod pumping (SRP) is one of the most widespread and reliable artificial-lift methods that is dominant in oil-recovery engineering. In production practice, due to harsh working environment, the sucker rod pumping system (SRPS) is prone to failure during operation, which not only affects the normal exploration of oil fields, reduces well production rate but also increases the cost of oil recovery (Li et al., 2018; Zhang et al., 2022). The shape of the

dynamometer card (DC) can reflect the downhole circumstances of the SRP. Therefore, the analysis and diagnosis of SRPS are mostly based on DC (Han et al., 2022; Lv et al., 2021b; Zheng et al., 2020).

Recently, intelligent recognition technologies such as self-organizing competitive networks (Xu et al., 2007), radial basis function (RBF) neural networks (Zhou et al., 2019), spectral clustering (Li et al., 2015), designated component analysis theory (Han et al., 2019; Li et al., 2013b), hidden Markov models (Zheng et al., 2019a), and support vector machines (SVMs) (Cheng et al., 2020; Liu et al., 2021) have been rapidly developed, particularly the deep learning-based well condition recognition method (Abdalla et al., 2020; Sun et al., 2022; Tian et al., 2021). These approaches have significantly improved the accuracy of well condition recognition. Li et al. (2013a) used the moment-curve approach to obtain the DC

* Corresponding author.

** Corresponding author.

E-mail addresses: chenghaibo@sia.cn (H.-B. Cheng), zp@sia.cn (P. Zeng).

characteristics and subsequently performed working condition identification using a modified SVM. Zheng and Gao (2017) extracted seven geometric features using curvature and barycentric decomposition methods and used a continuous hidden Markov model as a classifier for diagnosing well operating conditions. Zhou et al. (2019) proposed an approximate polygon-based Fourier descriptor and adopted an improved RBF neural network to build a diagnostic model. Wang et al. (2021) suggested an intelligent recognition model based on convolutional neural network (CNN), which implicitly extracts features of images through convolution and pooling operations and accomplishes well fault diagnosis. Zhou et al. (2018) extracted 12 DC features and 7 electrical power signal features based on mechanistic analysis, expert knowledge and priori information, used SVM for multi-view co-training and Hessian regularization to achieve pumping well working state identification. Pan et al. (2021) suggested a decision fusion method based on online Bayesian techniques and Bayesian probability formulations for the unsupervised fault detection of oil-pumping units. The above research focuses on improving the characterization capability of features and the classification performance of models to improve diagnostic results.

However, two challenges are encountered in actual sucker rod pumping systems. (i), deep learning-based well condition recognition methods usually use the DC as an image for feature extraction and classification (Song et al., 2023; Yin et al., 2023; Zhao et al., 2017). Different from a normal image containing complex information, the dynamometer card consists of a closed curve with only load and displacement data. The DC is input to the deep neural network as an image for model training, which further increases the computational effort. (ii), different SRPSs in the oil field have different system parameters, and the same SRPS generates different DCs at different moments, so the data have a certain degree of heterogeneity, which may affect the accuracy of diagnosis to some extent. The current working condition recognition of SRPS based on DCs either relies on extracting features from the DCs and then using classifiers to automatically classify the features (Lv et al., 2021c; Ye et al., 2020; Zhou et al., 2019), or inputting the DCs as an image into a deep neural network for working condition diagnosis (Song et al., 2023; Yin et al., 2023; Zhao et al., 2017). Additionally, methods such as dictionary-based transfer learning (Zhang and Gao, 2019), meta learning (He et al., 2023a), and meta-transfer learning (Zhang et al., 2022) have been used to solve the problem that different SRPSs in the oilfield have different system parameters resulting in training data that do not have similar distributions, and have achieved good results in working condition recognition. However, all the above methods use individual DC data for analysis, while the DC shape of oil wells changes slowly with time, and the heterogeneity of well data is usually ignored. At present, little research has been mentioned on the heterogeneity of oil well data.

Therefore, efficient feature extraction to reduce the input dimensionality of the depth model and suppress the heterogeneity of well data are worthy of investigation. Based on the above analysis, a fault diagnostic approach based on 4S-TFSM-CNN is proposed in this paper. 4S-TFSM-CNN includes two steps: feature extraction and working condition detection. In this method, a 4S-TFS feature is proposed to extract the graphical features of the DC, which can significantly reduce the computing power requirements compared with the image method. However, the issue of heterogeneity of well data is still ignored. A new data fusion method, 4S-TFSM, is introduced to solve the data heterogeneity problem. Specifically, the data heterogeneity is suppressed by enhancing the feature differences between different types of samples by subjecting multiple 4S-TFS to relative normalization and matrix computation. Finally, a specifically designed CNN is introduced to extract key features from 4S-TFSM for oil well working condition detection.

In summary, the following contributions are made in this paper.

- 1) To address the issue of high computational complexity due to the input of the DC as an image to the neural network, a 4-segment time-frequency signature (4S-TFS) extraction technique is suggested to perform feature extraction while compressing the data.
- 2) Considering the heterogeneity problem of the well data, a 4S-TFSM method is proposed to enhance the feature differences between different types of samples by relative normalization and matrix computation.
- 3) For the selection of features, the features that are sensitive to the working condition categories are automatically selected by deep learning algorithms.

The remainder of this work is arranged in the following manner. The studies related to our research are reviewed in Section 2. The suggested working condition identification methods are introduced in Section 3. The experimental comparisons are reported in detail in Section 4. Section 5 summarizes the conclusions and future research of the paper.

2. Related works

The working condition recognition of SRPS is a typical classification problem, and there are many methods of working condition recognition of pumping machines. Feature extraction and the classifier selection are the major dependencies of the pumping machine working condition identification method. Liu et al. (2021) used improved Fourier descriptors for feature extraction and SVMs for pumping well fault diagnosis. Chen et al. (2021a) extracted 16 features by combining the working mechanism of the sucker rod pump and identified the operating state of the sucker rod pump using the XGBoost algorithm with high generalization capability. Han et al. (2022) extracted five feature vectors with Freeman chain codes, and employed an optimized density peak clustering (DPC) method and an improved brainstorm optimization (BSO) algorithm to achieve online diagnosis.

The feature-extraction step is critical in the defect diagnosis procedure. High-quality features can ensure efficient and accurate working condition identification. Although Fourier descriptors (Zhou et al., 2019), curvilinear moments (Li et al., 2013a), wavelet transforms (Tagirova et al., 2021), statistical features (Zheng et al., 2019b), and other techniques have been used to extract DC features, these approaches still suffer from computational complexity, sensitivity to noise, and high feature dimensionality. Therefore, powerful feature-extraction methods with high computational efficiency and low feature dimensionality need to be investigated (Han et al., 2022).

In recent years, deep learning-based fault diagnosis methods represented by CNNs have been proposed for the identification of working conditions in SRPSs. Compared with methods such as random forest (RF), k-nearest neighbors (k-NN), spectral clustering (Li et al., 2015), and density peak clustering (DPC) (Wang et al., 2019), deep learning techniques show extremely strong pattern mining ability in environments with a large number of features, low sensitivity and complex data types, and are widely used in object detection (Yang et al., 2022), image recognition (He et al., 2023b), behavior prediction (He et al., 2022a), 3D reconstruction (Lee et al., 2021), visual relocalization (Chen et al., 2021b), and other fields. In addition, deep learning-based methods can also use deep neural networks to automatically extract image features from DCs. Zhao et al. (2017) developed image-based CNN and data-based-CNN approaches, and experiments revealed that recognition accuracy of CNN-based methods is higher than that of machine

learning algorithm represented by RF and k-NN. Song et al. (2023) improved a GoogLeNet framework for oil well condition recognition and compared the performance with nine models such as VGG-16, ResNet-18, DarkNet-19, and DenseNet-201, showing that deeper neural network layers do not always correspond to higher accuracy. Studies such as (He et al., 2022b; Sharaf, 2018; Tian et al., 2021; Yin et al., 2023) also validated the effectiveness of image-based CNN for well fault classification.

However, different SRPSs in the oil field have different system parameters, and the same SRPS generates different DCs at different moments, so the data have a certain degree of heterogeneity, which may affect the accuracy of diagnosis to some extent. For the problem of heterogeneous data among oil wells, Lv et al. (2021a) proposed an adaptive diagnostic method that uses simulated schematic power cards generated near the target operating point as training data and uses additional DC location features to improve the DC features by increasing the distance between different specimens. Zhang and Gao (2019) used dictionary-based transfer subspace learning to build transformation matrices, allowing the training and target data to be moved to a shared low-dimensional subspace. Lv et al. (2021c) proposed the static apparent stiffness features algorithm to preserve the fault information of DCs, and an incremental algorithm and generated DCs to improve the generalization capability of the support vector machine diagnostic model. He et al. (2023a) and Zhang et al. (2022) employed a meta learning framework using existing well data to solve the problem of few-shot samples of new oil well conditions. Although the above methods address the problem that data from different wells do not have similar distributions, they all use individual dynamometer card data for analysis. While the shapes of DCs for a well change slowly with time, the heterogeneity of data from the same well is usually ignored. At present, little research has been mentioned on the heterogeneity of oil well data. Additionally, although CNN has powerful feature extraction and classification capabilities, the DC directly input to the deep neural network as an image for model training, which further increases the computational effort. Therefore, a more accurate and efficient method for automatic well working condition identification would be a major advancement for promoting the construction of smart oil fields.

3. Methodology

3.1. 4S-TFS feature extraction

The theoretical DC is drawn under certain ideal conditions. The goal is to use the theoretical DC and the actual DC for comparison, from the differences in the graph can be determined for the downhole pump at the working condition of the judgment. The variation law of the static load at the suspension point is shown in Fig. 1, and the graph is a parallelogram abcd, which is also known as the theoretical static load DC. In this figure, the abc process is the line of static load change in the upstroke, where ab is the loading line. The traveling valve (TV) is closed during the loading operation, but the standing valve (SV) is not yet open, so the pumping pump does not pump oil. Only when the upward displacement of the suspension point reaches point b does the SV open and begin to suck oil. Therefore, bc is the pump suction process. The cda process is the downstroke static load change line, where cd is the unloading line. Again, in this process, although the suspension point moves downward, at this time, the SV and the TV are in the closed state. Until the end of unloading at point d, there is a relative displacement between the plunger and the pump barrel, the TV is opened, and the pump begins to delivery.

As shown in Fig. 1, the initial DC data of the well is a two-dimensional closed curve chart formed by the load-displacement.

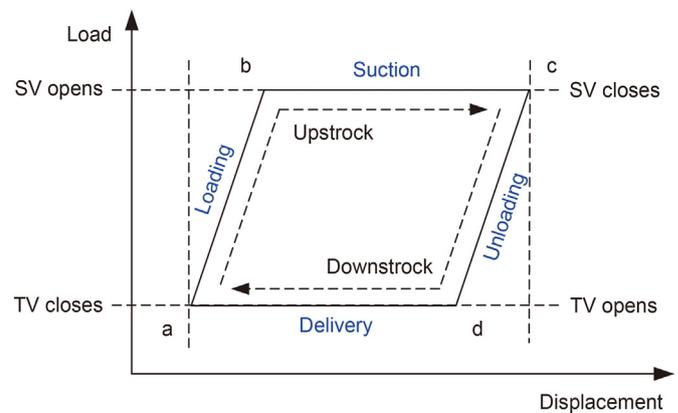


Fig. 1. The theoretical static load dynamometer card.

Usually, two-dimensional images have more information than one-dimensional signals, but in the parameter learning process, the calculation quantity is too large. The Lanczos Algorithm (Cheng et al., 2020; He et al., 2022b) is a common method for compressing images. However, due to the specificity of the dynamometer card data, the data needs to be preparatory converted and extracted to obtain comprehensive DC information. As a result, this work proposes a 4S-TFS feature-extraction approach with time- and frequency-domain features. The major purpose of the 4S-TFS aims to increase the feature-extraction capabilities by gaining a wealth of knowledge about the working circumstances involved in various features retrieved from each domain while minimizing the feature-extraction computational cost.

Fig. 2 depicts the 4S-TFS feature-extraction flowchart. (a) shows four points identified according to the theoretical analysis of the dynamometer card: a, b, c and d; i.e., the load-displacement curve is split into four segments (S_1, S_2, S_3 and S_4). Accurately extracting the opening and closing points of the pump valve is one of the key steps to extract the 4S-TFS features. According to the physical meaning of the opening and closing points, the opening and closing points of the pump valve are identified by analyzing the variation pattern of the load slope and time relationship curve (Zhang et al., 2021). The equation for the load slope K_L is

$$K_L = \frac{\Delta L_t}{\Delta t} = \frac{L_t(t+1) - L_t(t)}{\Delta t} \quad (1)$$

where $L_t(t)$ is the load point of the DC at time t and Δt is the time interval between two adjacent points of the DC. Since the relative time intervals at the time of acquisition are equal, Δt can be set to 1 (Zhang et al., 2021). Specifically, point b is located in the region of $p_1 - u_1$ and point d is located in the region of $p_2 - u_2$. Points p_1 and p_2 are the intersection points of median load and DC curve, and points u_1 and u_2 are the intersection points of median displacement and DC curve. According to Eq. (1), the maximum point of load slope change is obtained in each region, and the points b and d can be determined. At this time, the loads and displacements (L_t^a, D_t^a) , (L_t^b, D_t^b) , (L_t^c, D_t^c) , (L_t^d, D_t^d) are acquired for the four points a, b, c and d, respectively. Here, L_t and D_t denote the original load and displacement obtained by sampling with time.

After that, the upstroke load and downstroke load are sampled using the one-dimensional nearest neighbor interpolation technique to generate a new load-displacement curve, as shown in Fig. 2(b). Here, with displacement D as the independent variable and load L_D as the dependent variable, a new load-displacement curve independent of time is obtained by one-dimensional nearest neighbor interpolation. It can be expressed as follows.

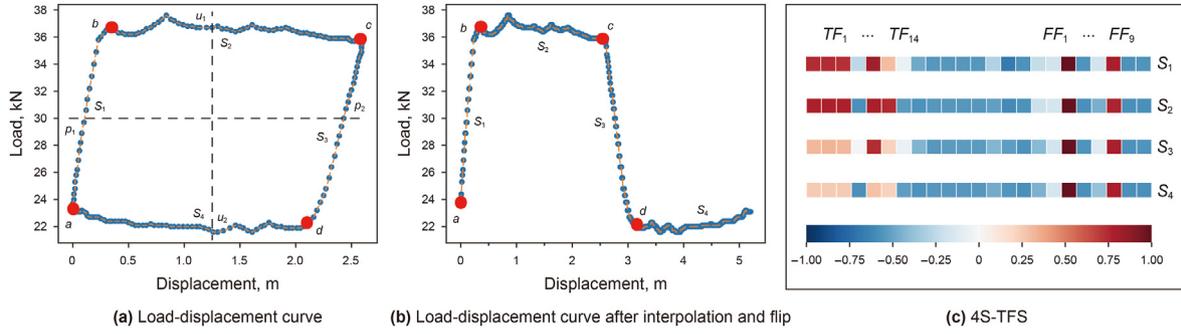


Fig. 2. Flowchart of 4S-TFS feature extraction.

$$L_D^{up} = \text{interp1d}(L_t^{up}, D_t^{up}, D^{up}) \quad (2)$$

$$L_D^{down} = \text{interp1d}(L_t^{down}, D_t^{down}, D^{down}) \quad (3)$$

$$D^{up} = [d_a, d_a + 0.01, d_a + 0.02, \dots, d_c] \quad (4)$$

$$D^{down} = [d_c, d_c - 0.01, d_c - 0.02, \dots, d_a] \quad (5)$$

$$d_a = \text{round}(D_t^a, 2) \quad (6)$$

$$d_c = \text{round}(D_t^c, 2) \quad (7)$$

where L_D^{up} and L_D^{down} denote the load after interpolation during the upstroke and downstroke, respectively, interp1d denotes the one-dimensional nearest neighbor interpolation, $\text{round}(D_t^a, 2)$ denotes the retention of 2 decimal places at the displacement point D_t^a , L_t^{up} and D_t^{up} denote the original load and displacement of the upstroke, respectively, and D^{up} denotes the interpolated displacement point of the upstroke process. According to the characteristics of the actual DC displacement data, the interpolation interval of displacement D^{up} and D^{down} are set to 0.01. At this time, the new load data $L_D = [L_D^{up}, L_D^{down}]$, $D = [d_a, d_a + 0.01, d_a + 0.02, \dots, 2(d_c - d_a + 0.01) + 0.01]$ for the variation in load L_D with displacement D is acquired.

Finally, the frequency-domain features FF and time-domain features TF of the load data that vary with displacement are chosen to constitute the feature set F , where S_i is the i th segment TF and FF_j^i is the j th element of the i th feature, and J is the number of time-frequency features of the feature set (He et al., 2023a), as shown in Eq. (8) and Fig. 2(c).

$$F = \begin{bmatrix} S_1 \\ S_2 \\ S_3 \\ S_4 \end{bmatrix} = \begin{bmatrix} f_1^1 & f_1^2 & \dots & f_1^J \\ f_2^1 & f_2^2 & \dots & f_2^J \\ f_3^1 & f_3^2 & \dots & f_3^J \\ f_4^1 & f_4^2 & \dots & f_4^J \end{bmatrix} \quad (8)$$

J denotes the 23 time- and frequency-domain features selected in this article. To obtain comprehensive information of the DC while taking into account the non-negativity of the load data that vary with displacement, 9 frequency-domain statistical features $FF_1 \sim FF_9$ transformed by fast Fourier transform (FFT) (Cooley and Tukey, 1965) and 14 time-domain statistical features $TF_1 \sim TF_{14}$ are

extracted from 4 segments of each curve. The FFT is based on the radix-2 butterfly block and the fast Fourier transform $y(k)$ of the load $L_D(n)$ is

$$y(k) = \sum_{n=1}^N L_D(n) \cdot W_N^{k \cdot n} \quad (9)$$

$$W_N = e^{-j \frac{2\pi}{N}} \quad (10)$$

where W_N is the N th twiddle factor, N is the number of points in the FFT, and j is the imaginary unit (Ferreira et al., 2021; He et al., 2023a).

$TF_1 \sim TF_8, TF_{15}, TF_{16}$ are referenced as the dimensional statistics, while $TF_9 \sim TF_{14}$ are referenced as the dimensionless statistics, as indicated in Table 1. Table 1 shows the series $L_D(n), n = 1, 2, 3, \dots, N$, where N is the number of points after load interpolation. Considering the nonnegativity of the load-displacement data after interpolation, i.e., $L_D(n) > 0$, the first 14 time-domain features are chosen in this paper.

In Table 2, the frequency-domain expressions of the characteristics are displayed. The feature FF_1 indicates the amplitude of vibration energy in the frequency domain, the features $FF_2 \sim FF_5, FF_9, FF_{11} \sim FF_{13}$ indicate the spectrum dispersion, and the feature $FF_6 \sim FF_8, FF_{10}$ indicates the major frequency band position shift. The FFT spectrum of the load data $L_D(n)$ is provided by $y(k), k = 1, 2, \dots, K$ in Table 2, where K is the number of spectrum lines and f_k is the frequency value of the k th spectrum line (He et al., 2023a; Yu et al., 2021). The first 9 frequency-domain features are considered in this paper, as some $FF \rightarrow 0$ or ∞ .

3.2. 4S-TFSM construction

Once the 4S-TFS features of a single DC have been extracted, the 4S-TFSM of the oil well DC data can be constructed. The sensors collect data at a certain time, and it is a common practice to generate l DCs of these data according to the production cycle of the pumping pump, as shown in Fig. 3(a). The method we use is to form a 4S-TFSM with l DCs. As shown in Fig. 3(d), the signature matrix M_i of the i th segment F is generated with l DCs, in particular, by the inner product between the l feature vector pairs of F . It can be expressed as

$$M_i = S_i^1 T \cdot S_i^2 \cdot S_i^3 T \cdot \dots \cdot S_i^l \quad (11)$$

$$l = 2, 4, 6, 8, \dots \quad (12)$$

The dimension of M_i is $J \times J$, which is 23×23 in this paper. Clearly, the TFSM of each segment M_i is symmetric about the major

Table 1
Time-domain feature expression.

Feature expression			
$TF_1 = \frac{1}{N} \sum_{n=1}^N L_D(n)$	$TF_5 = \max(L_D(n))$	$TF_9 = \frac{TF_2}{TF_{15}}$	$TF_{13} = \frac{\frac{1}{N} \sum_{n=1}^N (L_D(n) - TF_1)^3}{(TF_8)^3}$
$TF_2 = \sqrt{\frac{1}{N} \sum_{n=1}^N (L_D(n))^2}$	$TF_6 = \min(L_D(n))$	$TF_{10} = \frac{TF_{16}}{TF_2}$	$TF_{14} = \frac{\frac{1}{N} \sum_{n=1}^N (L_D(n) - TF_1)^4}{(TF_4)^2}$
$TF_3 = \left(\frac{1}{N} \sum_{n=1}^N \sqrt{ L_D(n) } \right)^2$	$TF_7 = TF_5 - TF_6$	$TF_{11} = \frac{TF_{16}}{TF_{15}}$	$TF_{15} = \frac{1}{N} \sum_{n=1}^N L_D(n) $
$TF_4 = \frac{1}{N} \sum_{n=1}^N (L_D(n) - TF_1)^2$	$TF_8 = \sqrt{TF_4}$	$TF_{12} = \frac{TF_{16}}{TF_3}$	$TF_{16} = \max L_D(n) $

diagonal of the matrix. The number of l is set to an integer multiple of 2 in this work.

Next, just as the three red-green-blue (RGB) channels are used to form an image, we sequentially form the four matrices $M_1, M_2, M_3,$ and M_4 into a 4S-TFSM which can be denoted as $M = [M_1, M_2, M_3, M_4]$ to represent the DC of wells, as shown in Fig. 3(e).

In particular, to prevent large differences in the values of different features in the time-frequency feature vector, we applied standard deviation normalization (Z-score normalization) in the 4S-TFS feature extraction and TFSM construction process, and the

transformation function is

$$x^* = \frac{x - \mu}{\sigma} \tag{13}$$

where μ and σ are the mean and standard deviation of the input data x , respectively, and x can be considered as the feature set F or the feature matrix M . In this way, the processed data conform to a conventional normal distribution with a mean of 0 and a standard deviation of 1, which can speed up the convergence of the training network and improve the accuracy.

Table 2
Frequency-domain feature expression.

Feature expression			
$FF_1 = \frac{\sum_{k=1}^K y(k)}{K}$	$FF_5 = \sqrt{\frac{\sum_{k=1}^K [(f_k - FF_{10})^2 y(k)]}{K}}$	$FF_9 = \frac{\sum_{k=1}^K [\sqrt{ f_k - FF_{10}} y(k)]}{K \sqrt{FF_5}}$	$FF_{13} = \frac{\sum_{k=1}^K [(f_k - FF_{10})^4 y(k)]}{K (FF_5)^4}$
$FF_2 = \frac{\sum_{k=1}^K [y(k) - FF_1]^2}{K - 1}$	$FF_6 = \sqrt{\frac{\sum_{k=1}^K (f_k^2 y(k))}{\sum_{k=1}^K y(k)}}$	$FF_{10} = \frac{\sum_{k=1}^K (f_k y(k))}{\sum_{k=1}^K y(k)}$	-
$FF_3 = \frac{\sum_{k=1}^K [y(k) - FF_1]^3}{K (\sqrt{FF_2})^3}$	$FF_7 = \sqrt{\frac{\sum_{k=1}^K (f_k^4 y(k))}{\sum_{k=1}^K (f_k^2 y(k))}}$	$TF_{11} = \frac{TF_5}{TF_{10}}$	-
$FF_4 = \frac{\sum_{k=1}^K [y(k) - FF_1]^4}{K (FF_2)^2}$	$FF_8 = FF_6 \sqrt{\frac{\sum_{k=1}^K (f_k^2 y(k))}{\sum_{k=1}^K (f_k^4 y(k))}}$	$FF_{12} = \frac{\sum_{k=1}^K [(f_k - FF_{10})^3 y(k)]}{K (FF_5)^3}$	-

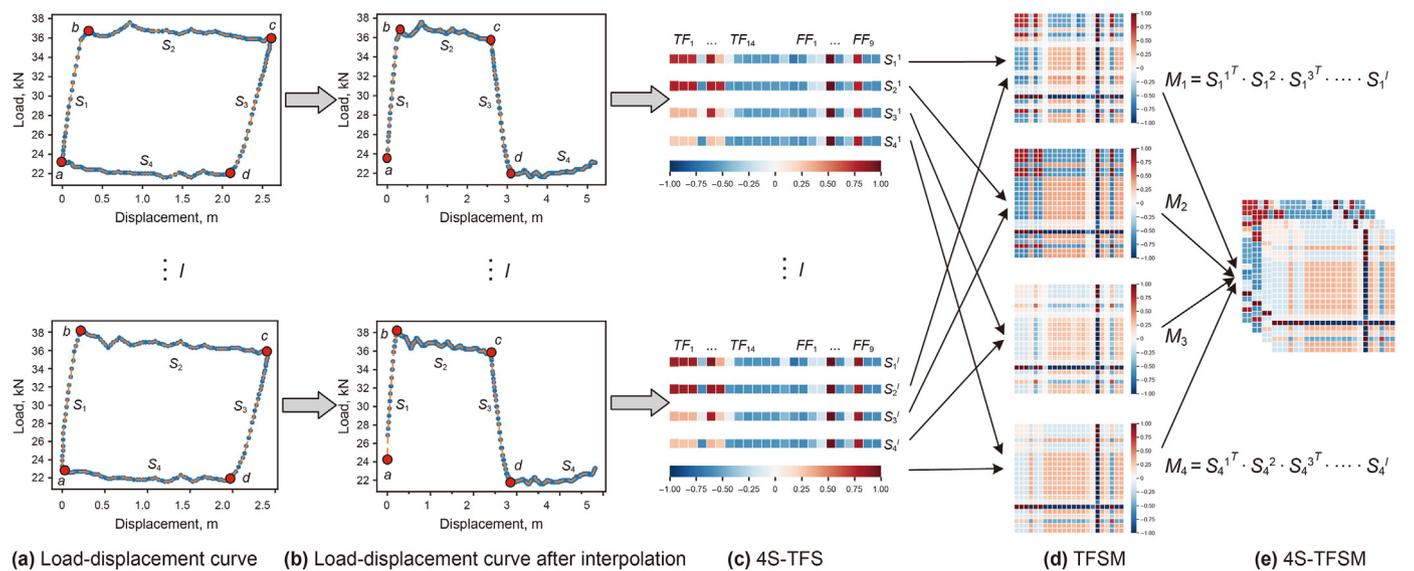


Fig. 3. Flowchart of 4S-TFSM construction.

Similar to the construction of 4S-TFS and 4S-TFSM, the dynamometer card curves in Fig. 3(a) and (b) can also be split into 1, 2 and 8 segments to create the 1S-TFS, 2S-TFS, and 8S-TFS in Fig. 3(c), which in turn form the 1S-TFSM, 2S-TFSM and 8S-TFSM in Fig. 3(e), respectively. Section 4.2 of the experimental section contains the related experimental results.

3.3. Working condition recognition framework of SRPS

With the extraction of 4S-TFSM features, CNNs can be used for working condition identification of SRPS. The framework of the 4S-TFSM-CNN method proposed in this paper is shown in Fig. 4, and the general steps are summarized as follows.

- 1) Data collection. The data are measured by dynamometer sensors attached to the pumping rods and sent to the data center via wireless networks. The displacement data are obtained through the acceleration sensor by quadratic integration. Based on the collected displacement and load data, the DC can be generated. The dynamometer card reflects the running process of the sucker rod.
- 2) 4S-TFS feature extraction. The proposed 4S-TFS feature extraction method is employed to extract the time-frequency-domain features of each dynamometer card, and the data compression is performed at the same time as feature extraction, which effectively reduces the dimensionality of the data and solve the issue of high computational complexity caused by the DC as a two-dimensional image input to the neural network.
- 3) 4S-TFSM construction. The collected 4S-TFS data generated from multiple DCs are used to construct 4S-TFSM by relative normalization and matrix calculation, which can enhance the feature differences between different working conditions and solve the heterogeneity problem of training samples.
- 4) Working condition recognition. The optimal selection of features is accomplished by a specifically designed CNN model to automatically select 4S-TFSM features that are sensitive to the category of working conditions. Specifically, the dataset is

divided into a training set and a test set according to a certain ratio. The CNN model is used to obtain optimal parameters sensitive to task variation by the training set, and the test set is used for working conditions recognition.

4. Experimental evaluation

In this part, Section 4.1 describes the data collection and processing process. Section 4.2 presents the hyperparameter selection and performance evaluation of the 4S-TFSM-CNN method proposed in this paper. Section 4.3 compares the proposed 4S-TFSM-CNN and 4S-TFS-CNN with the traditional method image-CNN in four aspects of classification accuracy, model parametric number, algorithm complexity and training time to verify the effectiveness of the innovation points presented in this paper. Section 4.4 discusses the negative impact of the heterogeneity of training samples on the actual working condition recognition, and explains the rationality of the innovation point by comparing 4S-TFSM-CNN with some traditional machine learning and deep learning methods.

4.1. Data collection and processing

To demonstrate the effectiveness and performance of the proposed diagnostic method, more than 100,000 DCs were collected from an oilfield in northern China. The load and displacement data for each dynamometer card is composed of 200 point pairs. Generally, it is difficult to obtain multiple working conditions in a single well. Therefore, we gathered data from nearly 1300 wells over a period of more than 90 days. These data contained more than ten types of failures. To maximize the generality of the data, we screened 8 common types of working conditions from the sensor data of multiple wells, which are the insufficient liquid supply (ILS), traveling valve leakage (TVL), standing valve leakage (SVL), normal operation condition (NOC), gas interference (GIF), continuous pumping and spraying (CPS), oil pipe leakage (OPL) and upstroke pump bumping (UPB); the corresponding DCs are shown in Fig. 5. For each working condition, 3500 samples were selected from

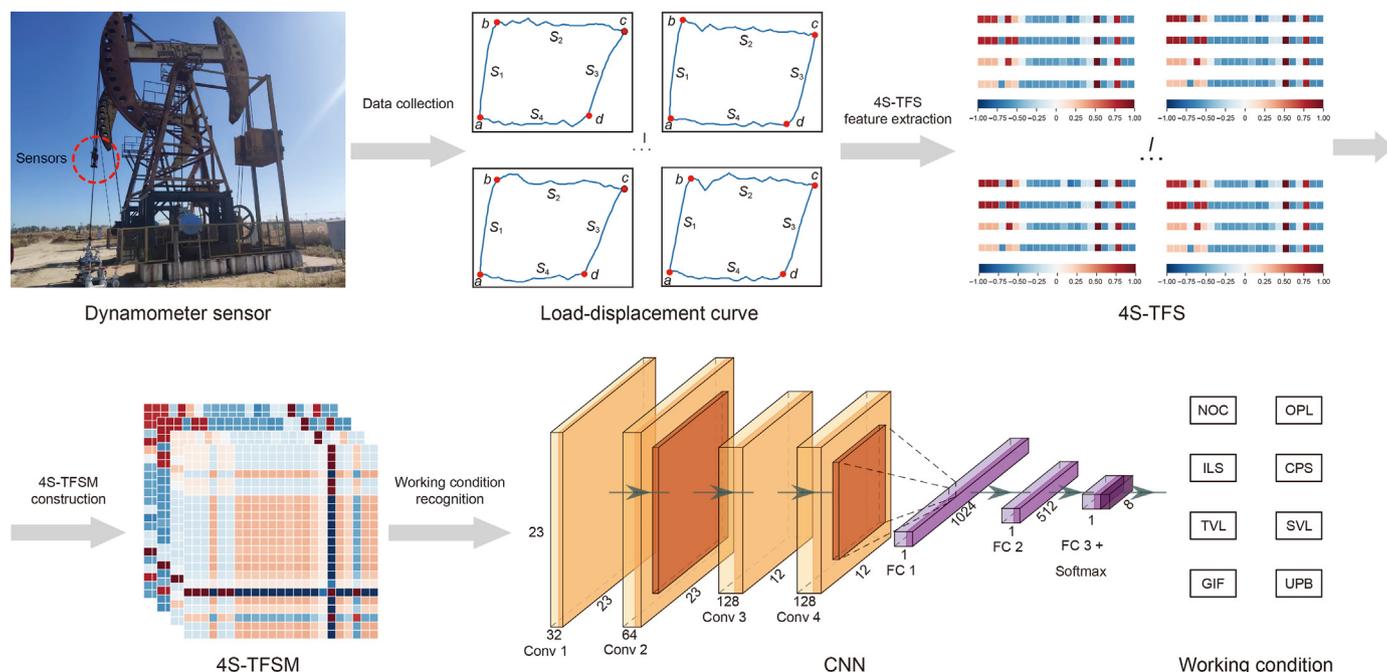


Fig. 4. Framework of 4S-TFSM-CNN method for working condition recognition.

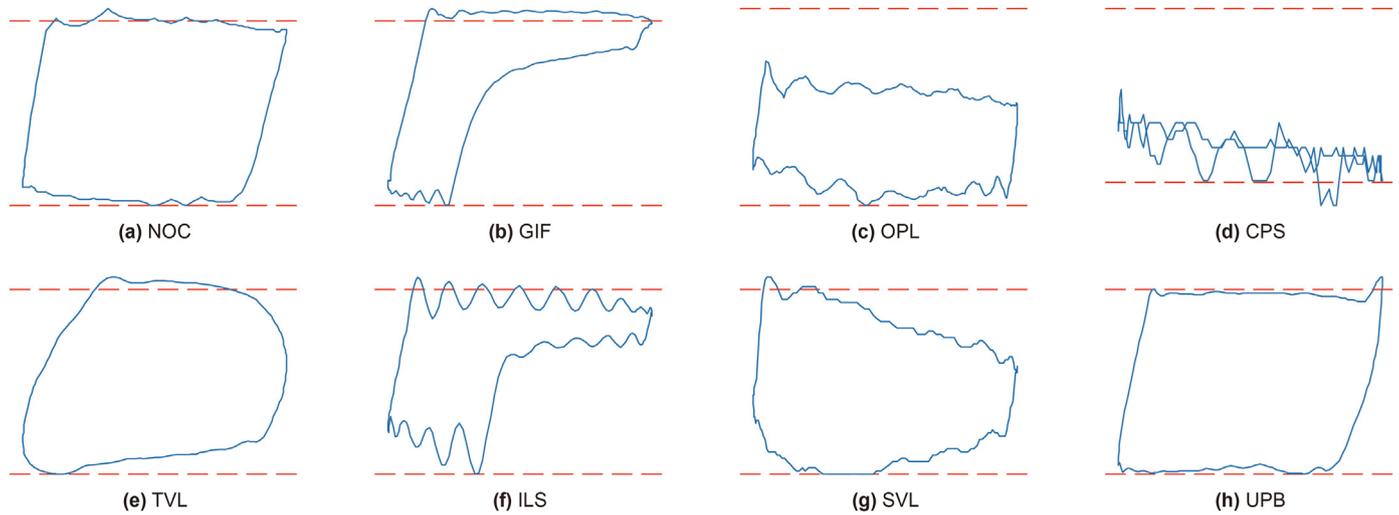


Fig. 5. Shapes of the dynamometer cards for 8 different working conditions. The vertical axis is the load and the horizontal axis is the displacement.

multiple wells for the experiment. Then, the data were divided into training and test datasets according to 6:1, i.e., 3000 training samples and 500 test samples for each type of working condition.

Since the DCs from different wells have different load and displacement ranges, the collected data were processed using the Min-Max normalization method (Zhang and Gao, 2019).

$$D_t^*(i) = \frac{D_t(i) - \min(D_t)}{\max(D_t) - \min(D_t)} \quad (14)$$

$$L_t^*(i) = \frac{L_t(i) - \min(L_t)}{\max(L_t) - \min(L_t)} \quad (15)$$

$$i = 1, 2, 3, \dots, 200 \quad (16)$$

where $D_t^*(i)$ and $L_t^*(i)$ are the normalized displacement and load

data, i is the point sampled with time, $\min(D_t)$, $\max(D_t)$, $\min(L_t)$ and $\max(L_t)$ are the minimum displacement, maximum displacement, minimum load and maximum load, respectively (Cheng et al., 2020).

Fig. 6 shows the shape of the dynamometer card for a well in the same coordinate interval at different times. It can be seen that the shapes of the dynamometer cards are different for the same well at different times under the same working condition. The heterogeneity of the data will undoubtedly reduce the accuracy of working condition recognition.

All experiments were obtained on a desktop workstation with Intel Core i7-10700K 3.8 GHz CPU, 64-GB RAM, GeForce RTX 3090 GPU and Ubuntu 18.04 operating system.

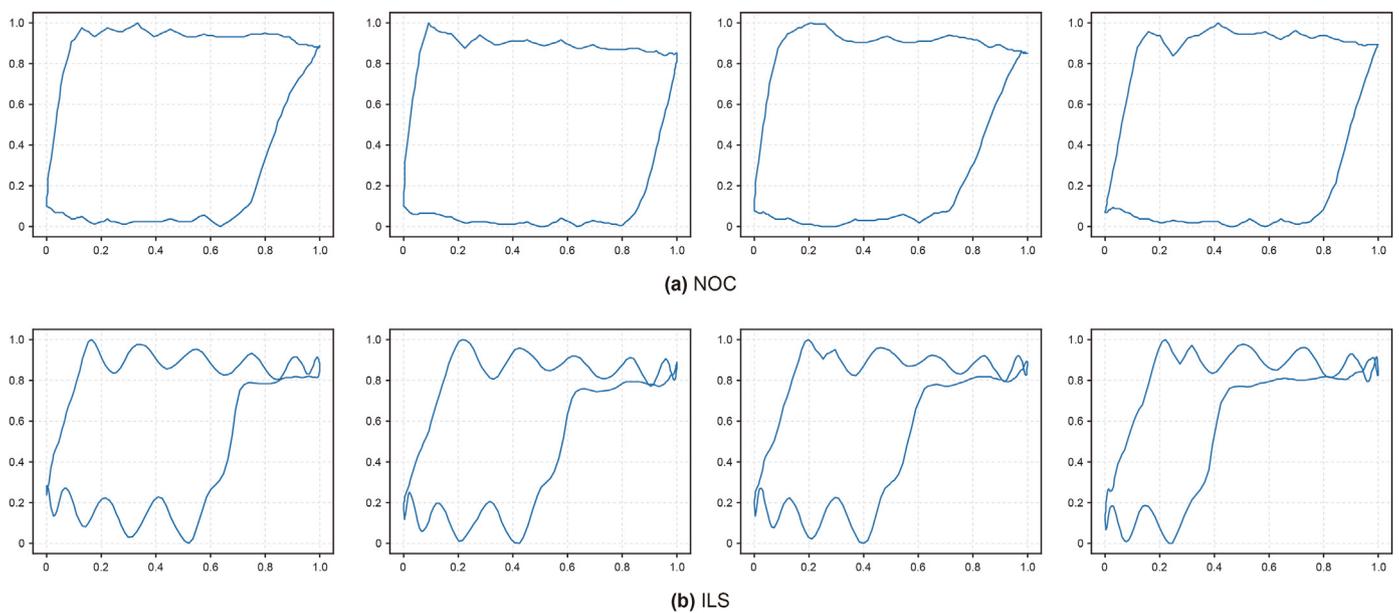


Fig. 6. Shapes of the dynamometer cards for a well at different times. The figure shows two common working conditions, with the horizontal axis indicating the normalized displacement and the vertical axis indicating the normalized load. It is clearly seen that the DCs collected from the same well at different moments are different, and the heterogeneity of the data may degrade the performance of the working condition identification.

4.2. 4S-TFSM-based CNN

Neural networks solve multiclass classification problems in machine learning. In particular, CNNs allow weight sharing and downsampling at several stages of the network, reducing the number of parameters and resulting in an effective architecture for large-scale classification.

The structure and parameters of the 4S-TFSM-CNN network are given in Table 3. During the training process, stochastic gradient descent is used to minimize the loss according to the cross-entropy loss function. The rectified linear unit (ReLU) activation function is employed in the implicit layer, and softmax is applied in the output layer to obtain a probability sum of 1 for each category and to use the predicted category as the category with the highest probability.

The training set was used to generate 320,000 training subtasks in the model training process, with the batch size set to 32, the training step size set to 10,000, and the learning rate set to 0.001. To prevent chance and specificity, 300 task data were selected in the test set and the average test precision and standard deviation of all tasks were obtained.

The accuracy of the oil well working condition recognition using different numbers of DCs constituting the 4S-TFSM is shown in Table 4. It can be seen that after feature extraction of the DC data by 4S-TFS, the highest recognition accuracy of 98.48% and the lowest standard deviation of 0.72% were achieved with the composition of 4 DCs. The training and test processes of 10,000 times for this scenario are shown in Fig. 7.

As shown in Fig. 7, only approximately 3000 training epochs are needed to achieve an accuracy of nearly 98% for working condition recognition on the test dataset, and the loss value drops to a low level and then stabilizes around that value, which indicates that the trained CNN has good stability and also shows that it is feasible to train the CNN model with 3000 samples.

As described in Section 3.2, based on the 4S-TFSM with 4 DCs, we also used 4 DCs to form the 1S-TFSM, 2S-TFSM and 8S-TFSM. Table 5 shows the accuracy of the working condition identification using 4 DCs under different TFSMs.

It is evident that the 1S-TFSM constructed with all interpolated load-displacement data cannot extract enough feature information, thus reducing the recognition accuracy. Similarly, the recognition accuracy of splitting the dynamometer curve into 8 segments to form the 8S-TFSM is relatively low due to the dynamometer curve consists of only 200 points with simple features, and the sample points are too small to extract effective time-frequency features, which reduces the recognition accuracy of the 8S-TFSM. It is worth mentioning that the 2S-TFSM and 4S-TFSM achieve nearly the same result, which is due to the simple curve composition of S_1 and S_3 in Fig. 3, and the two parts also belong to the upstroke load and downstroke load, respectively, which causes the difference in the extracted time-frequency features to not be obvious.

Table 3
The architecture of 4S-TFSM-CNN.

Layer	Type	Feature maps	Filter size	Dropout
0	Input	$23 \times 23 \times 4$	—	—
1	Conv 1	$23 \times 23 \times 32$	3×3	—
2	Conv 2	$23 \times 23 \times 64$	3×3	—
3	Pool 1	$12 \times 12 \times 64$	2×2	—
4	Conv 3	$12 \times 12 \times 128$	3×3	—
5	Conv 4	$12 \times 12 \times 128$	3×3	—
6	Pool 2	$6 \times 6 \times 128$	2×2	—
7	FC 1	$1 \times 1 \times 1024$	—	0.1
8	FC 2	$1 \times 1 \times 512$	—	0.1
9	FC 3	$1 \times 1 \times 8$	—	0.1

4.3. Performance comparison

The method of classifying the dynamometer cards as images has been successfully applied to the working condition recognition of SRPS. Inspired by this result, we conducted a comparative study of image-CNN and the 4S-TFSM-CNN proposed in this paper. We first draw the DC based on the sensor data and generate a 640×480 image. Subsequently, the binarization process is performed. Finally, the binarized image is compressed to 64×64 and used as the input of the image-CNN.

Different from 4S-TFSM-CNN, the input size of image-CNN is $64 \times 64 \times 1$, so we added 1 maximum pooling layer between the first and second convolutional layers. In addition, we also tried to input load-displacement data into the neural network to generate the data-based-CNN, and to input 4S-TFS features extracted from a single DC into the neural network to generate the 4S-TFS-CNN. The structure and parameters of the three neural networks are depicted in Table 6.

The accuracies of the four methods for working condition recognition are shown in Table 7. It is evident that the recognition accuracy of 4S-TFS-CNN is 97.53%, which are significantly higher than the accuracies of 96.83% and 97.39% of image-CNN and data-based-CNN methods. Meanwhile, the standard deviation of 4S-TFS-CNN is 0.91%, which is also lower than 1.08% and 0.99% of image-CNN and data-based-CNN, which indicates that the 4S-TFS feature-extraction method for a single DC can extract effective features for CNN classification. In addition, among the four methods, 4S-TFSM-CNN has the highest recognition accuracy of 98.48% and the lowest standard deviation of 0.72%, which demonstrates that the construction of 4S-TFSM by combining multiple DCs can solve the heterogeneity problem and improve the work condition recognition accuracy.

Fig. 8 shows the working condition recognition accuracies of image-CNN, data-based-CNN, and 4S-TFSM-CNN with different training steps. Fig. 8(a) demonstrates that throughout the training phase, the 4S-TFSM-CNN achieves similar accuracy and loss as the image-CNN and data-based-CNN. On the test set in Fig. 8(b), the accuracy of 4S-TFS-CNN is higher than that of the image-CNN and data-based-CNN overall, which is consistent with the findings in Table 7. This also illustrates that sufficient feature information can be extracted from multiple DCs for CNN working condition recognition by a 4S-TFSM, which confirms the effectiveness of the 4S-TFSM-CNN.

In addition, Table 8 demonstrates the effects of the four methods on CNN performance. It is evident that the proposed 4S-TFS-CNN method has the least algorithmic complexity, time complexity and training time, and the recognition accuracy is higher than that of Image-CNN and Data-based-CNN. Compared with data-based-CNN, the proposed 4S-TFSM-CNN method increases in algorithm complexity, time complexity and training time, but the recognition accuracy is the highest. The analysis of Tables 7 and 8 verifies that the proposed 4S-TFSM can reduce the model complexity and computational cost, and improve recognition accuracy.

4.4. Discussion

We used classical machine learning methods such as RF, k-NN and SVM for prediction on the same dataset, and the results are given in Table 9. It can be seen that image-RF (image as input feature and RF algorithm as classifier) has the lowest expected accuracy of 92.59% and 4S-TFSM-CNN obtains the highest accuracy of 98.48%. Compared with other deep learning-based methods, such as RNN, LSTM and GRU, image-RNN has the lowest accuracy of 87.40%, while LSTM and GRU-based methods have better performance than RNN, but the accuracy is also lower than the proposed

Table 4
Working condition recognition accuracies of the 4S-TFSM-CNN with different numbers of DCs.

Number of DCs	$l = 2$	$l = 4$	$l = 6$	$l = 8$	$l = 10$
Accuracy	98.12% ± 0.81%	98.48% ± 0.72%	97.71% ± 0.89%	98.07% ± 0.78%	97.83% ± 0.91%

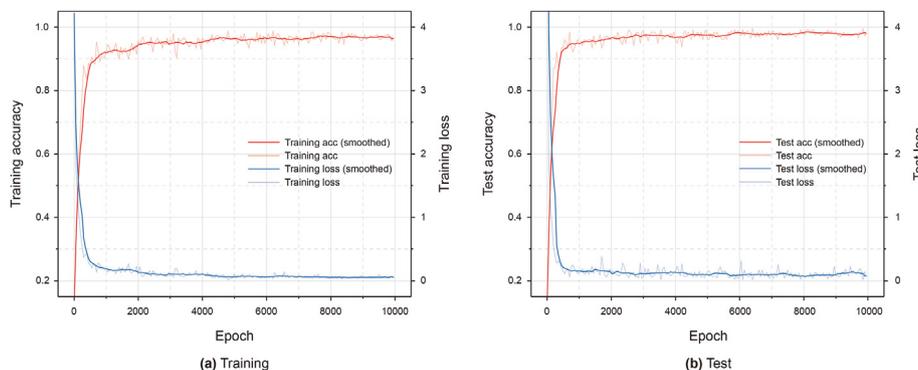


Fig. 7. Training and test accuracies of the 4S-TFSM-CNN with 4 DCs.

Table 5
The accuracy of working condition recognition for different TFSMs with 4 DCs.

Method	1S-TFSM-CNN	2S-TFSM-CNN	4S-TFSM-CNN	8S-TFSM-CNN
Accuracy	95.65% ± 1.29%	98.39% ± 0.76%	98.48% ± 0.72%	97.01% ± 1.04%

4S-TFSM-CNN method. Moreover, the classification accuracies of RF, SVM, RNN, LSTM, GRU and CNN using 4S-TFS as input data

outperformed those of image- and data-based classification results, which illustrates that the 4S-TFS feature extraction for a single DC contributes to the performance improvement of working condition recognition. In addition, the classification results of 4S-TFSM as input data when using SVM, RNN, LSTM and CNN as classifiers also outperformed those of the other three data as input, indicating that feature fusion of multiple DCs can reduce data heterogeneity and improve the working condition recognition accuracy of oil wells.

Table 6
The architectures for image-CNN, data-based-CNN and 4S-TFS-CNN.

Layer	Type	Dropout	Image-CNN		Data-based-CNN		4S-TFS-CNN	
			Feature maps	Filter size	Feature maps	Filter size	Feature maps	Filter size
0	Input	—	64 × 64 × 1	—	2 × 200 × 1	—	4 × 23 × 1	—
1	Conv 1	—	64 × 64 × 32	3 × 3	2 × 200 × 32	1 × 3	4 × 23 × 32	1 × 3
2	Pool 1	—	32 × 32 × 32	2 × 2	2 × 100 × 32	1 × 2	—	—
3	Conv 2	—	32 × 32 × 64	3 × 3	2 × 100 × 64	1 × 3	4 × 23 × 64	1 × 3
4	Pool 2	—	16 × 16 × 64	2 × 2	2 × 50 × 64	1 × 2	4 × 12 × 64	1 × 2
5	Conv 3	—	16 × 16 × 128	3 × 3	2 × 50 × 128	1 × 3	4 × 12 × 128	1 × 3
6	Pool 3	—	—	—	2 × 25 × 128	1 × 2	—	—
7	Conv 4	—	16 × 16 × 128	3 × 3	2 × 25 × 128	1 × 3	4 × 12 × 128	1 × 3
8	Pool 4	—	8 × 8 × 128	2 × 2	2 × 13 × 128	1 × 2	4 × 6 × 128	1 × 2
9	FC 1	0.1	1 × 1 × 1024	—	1 × 1 × 1024	—	1 × 1 × 1024	—
10	FC 2	0.1	1 × 1 × 512	—	1 × 1 × 512	—	1 × 1 × 512	—
11	FC 3	0.1	1 × 1 × 8	—	1 × 1 × 8	—	1 × 1 × 8	—

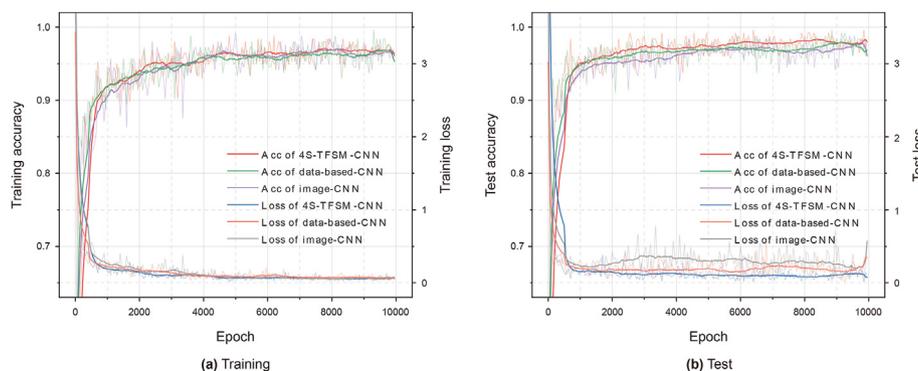


Fig. 8. Training and test accuracies of image-CNN, data-based-CNN and 4S-TFSM-CNN with different training steps.

Table 7
Working condition recognition accuracies based on different methods.

Method	Image-CNN	Data-based-CNN	4S-TFS-CNN	4S-TFSM-CNN
Accuracy	96.83% ± 1.08%	97.39% ± 0.99%	97.53% ± 0.91%	98.48% ± 0.72%

Table 8
CNN algorithm indicators.

Method	Image-CNN	Data-based-CNN	4S-TFS-CNN	4S-TFSM-CNN
Feature map size	64 × 64 × 1	2 × 200 × 1	4 × 23 × 1	23 × 23 × 4
Param, MB	10.21	4.12	3.78	6.03
FLOPs, GB	43.97	5.21	4.00	24.35
Training time, min	162.15	21.83	7.32	99.02

Table 9
Overall prediction accuracy of the methods.

		Classifier						
		RF	k-NN	SVM	RNN	LSTM	GRU	CNN
Feature	Image	92.59%	96.89%	93.79%	87.40%	95.17%	94.59%	96.83%
	Data-based	96.02%	96.99%	97.70%	93.31%	96.67%	96.69%	97.39%
	4S-TFS	97.94%	96.99%	97.74%	96.89%	97.87%	98.07%	97.53%
	4S-TFSM	96.57%	96.02%	97.90%	97.27%	98.00%	97.57%	98.48%

What exactly is the reason for the difference in recognition accuracy? Which types of working conditions are more difficult to distinguish? We visualize the dataset to better understand the data structure. The t-distributed stochastic neighbor embedding (t-SNE) methodology is a recently developed dimensionality reduction method that is frequently used to visualize data. t-SNE can integrate high-dimensional data in two- or three-dimensional space, allowing related objects in the high-dimensional space to approach each other in the low-dimensional space (Zhao et al., 2017). Fig. 9 shows the scatter plot of the sample data after it has been embedded in two-dimensional space.

The data in categories 1 (SVL), 2 (NOC), and 6 (TVL) are depicted as reasonably independent clusters in Fig. 9, which means that these categories may be relatively easy to identify and classify. Other types of samples, become entangled and mixed together, making it impossible to distinguish between them. For example, category 0 (ILS) and category 3 (GIF) are almost indistinguishable in

the plot because their corresponding DCs are highly similar, while category 5 (UPB) is scattered throughout the plot.

Fig. 10 shows the confusion matrix derived on the test dataset using the four data processing approaches, image-CNN, data-based-CNN, 4S-TFS-CNN, and 4S-TFSM-CNN. We can clearly observe that the vast majority of samples lie on the diagonal. It is straightforward to observe from the diagonal of the four confusion matrices that category 0 (ILS), category 3 (GIF), category 4 (OPL), and category 5 (UPB) are the four most difficult categories to distinguish from each other, as these four working conditions have the lowest classification accuracies. This is consistent with the results of the t-SNE visualization discussed in Fig. 9. The recognition accuracy of 4S-TFSM-CNN is also higher than the remaining three schemes in the four most difficult to distinguish categories, indicating that the 4S-TFSM feature-extraction approach can weaken the heterogeneity of the data and significantly affect the diagnosis performance.

The heterogeneity induced by the varied distributions of data between distinct classes, as illustrated in Fig. 10, is an important factor limiting the recognition accuracy of our model. Here, instead of using individual datasets as direct inputs, we use a simple synthetic feature-extraction method, i.e., a feature-extraction technique that synthesizes multiple datasets.

The idea is to “synthesize” training samples from real samples that are more suitable for deep learning to complete the classification task by extracting time-frequency features from the feature space of original dataset for normalization and matrix operations. In addition, despite the high accuracy of 98.48%, we evaluate the incorrectly classified samples from a probabilistic perspective to assess the performance of the 4S-TFSM-CNN. We infer that some misclassifications, such as category 5 (UPB) in Fig. 10, are attributable to label ambiguity. For these samples, it may be difficult to identify clear fault types, even for human experts, or the samples might contain multiple fault types at the same time, thus introducing some label noise throughout the dataset.

In terms of training efficiency, CNNs may require more training time than typical machine learning methods (e.g., RF, k-NN and SVM) due to their processes of learning features. However, traditional machine learning methods are not friendly in the case of

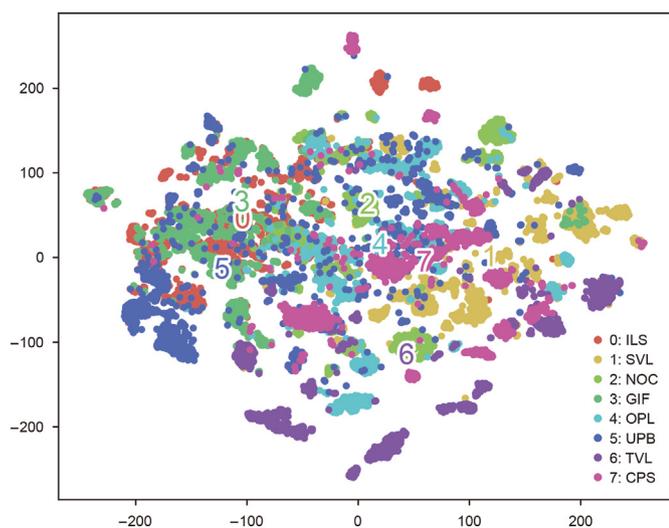


Fig. 9. t-SNE embedding in 2D space.

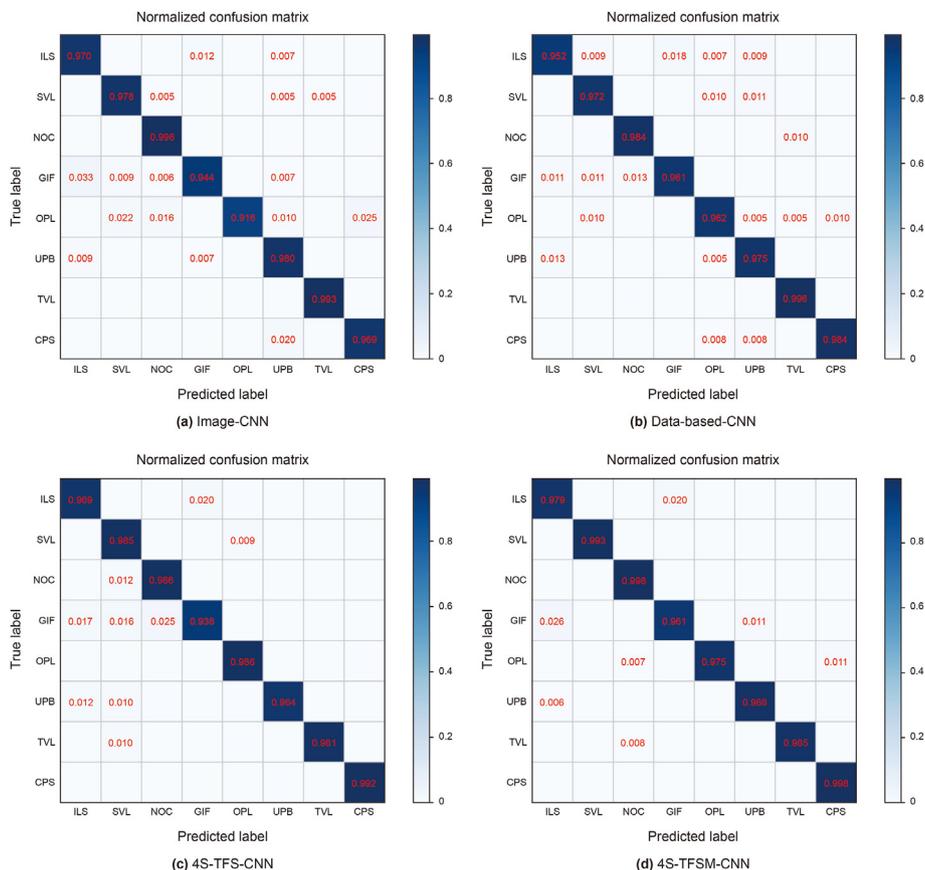


Fig. 10. The confusion matrix for the prediction under different methods. The numbers on the diagonal line represent the probability of correctly categorized samples, while the numbers on the nondiagonal line represent the probability of incorrectly classified samples. For clarity, we retain three decimal places, and the misclassification probabilities below 0.005 are not shown on the confusion matrix.

large quantities of data. Fortunately, once trained, the network can be recovered and utilized for model prediction multiple times, and the training time of our proposed 4S-TFSM-CNN technique is substantially faster than that of image-CNN. Therefore, the 4S-TFSM-CNN-based approach is efficient and feasible in real-world oil well working condition recognition.

5. Conclusions

In this paper, a 4S-TFSM-based feature-extraction technique for oil well DCs is proposed, and a specifically designed CNN is employed to achieve oil well working condition recognition. The technique uses 4S-TFS to extract DC features, which can effectively improve the recognition accuracy of oil well conditions while reducing the computational cost. Subsequently, 4S-TFSM is constructed by relative normalization and matrix calculation to enhance the feature differences between various types of working conditions and suppress the data heterogeneity problem. Finally, a special CNN is designed to automatically select 4S-TFSM features that are sensitive to working condition categories to achieve high-accuracy identification of oil well working conditions. Experimental results with a large amount of real oil well data in a field demonstrate the effectiveness of the proposed method.

Considering the computational complexity of neural networks, a 4-layer CNN is used in this paper, although the accuracy of working condition recognition may be higher as the number of neural network layers increases. The DC data are generated with time series, and in future work, we will also consider temporal convolutional networks (TCNs), which can use convolutional features

while considering temporal information.

Declaration of 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.

Acknowledgements

We would like to thank the associate editor and the reviewers for their constructive comments. This work was supported in part by the National Natural Science Foundation of China under Grant 62203234; in part by the State Key Laboratory of Robotics of China under Grant 2023-Z03; in part by the Natural Science Foundation of Liaoning Province under Grant 2023-BS-025; and in part by the Research Program of Liaoning Liaohe Laboratory under Grant LLL23ZZ-02-02.

A. Appendix

1.1. Parameter setting of RF

We mainly consider the effects of `n_estimators`, `min_samples_split`, `max_depth`, `max_features`, and `min_samples_leaf` on the performance of RF classification, and use `GridSearchCV` to automate the search for the above parameters. The parameters are set as shown in [Table A1](#).

Table A1
Parameter setting of RF.

Input	n_estimators	min_samples_split	max_depth	max_features	min_samples_leaf
Image	190	2	90	41	1
Data-based	190	2	50	21	1
4S-TFS	190	2	70	21	1
4S-TFSM	180	2	40	31	1

1.2. Parameter setting of k-NN

We mainly focus on the effects of n_neighbors, weights, and p on k-NN classification performance and use GridSearchCV to automate the search for the above parameters. Interestingly, the four inputs have the best classification accuracy when all three parameters are equal to 1.

1.3. Parameter setting of SVM

For SVM, we mainly consider the effects of kernel, C, gamma and degree on the classification performance and use GridSearchCV to automate the search for the above parameters. The parameters are set as given in Table A2.

Table A2
Parameter setting of SVM.

Input	Kernel	C	Gamma	Degree
Image	Linear	0.001	0.0005	3
Data-based	RBF	50	0.5	3
4S-TFS	RBF	100	0.5	3
4S-TFSM	RBF	50	0.005	3

References

Abdalla, R., Ela, M.A., El-Banbi, A., 2020. Identification of downhole conditions in sucker rod pumped wells using deep neural networks and genetic algorithms. *SPE Prod. Oper.* 35 (2), 435–447. <https://doi.org/10.2118/200494-pa>.

Chen, L., Gao, X., Li, X., 2021a. Using the motor power and XGBoost to diagnose working states of a sucker rod pump. *J. Petrol. Sci. Eng.* 199, 108329. <https://doi.org/10.1016/j.petrol.2020.108329>.

Chen, Z., Pei, H., Wang, J., et al., 2021b. Survey of monocular camera-based visual relocalization. *Robot* 43 (3), 373–384. <https://doi.org/10.13973/j.cnki.robot.200350> (in Chinese).

Cheng, H.B., Yu, H.B., Zeng, P., et al., 2020. Automatic recognition of sucker-rod pumping system working conditions using dynamometer cards with transfer learning and svm. *Sensors* 20 (19), 1–15. <https://doi.org/10.3390/s20195659>.

Coolley, J.W., Tukey, J.W., 1965. An algorithm for the machine calculation of complex fourier series. *Math. Comput.* 19 (90), 249–259. <https://doi.org/10.1090/S0025-5718-1965-0178586-1>.

Ferreira, G., Paim, G., Rocha, L.M.G., et al., 2021. Low-power fast Fourier transform hardware architecture combining a split-radix butterfly and efficient adder compressors. *IET Comput. Digital Tech.* 15 (3), 230–240. <https://doi.org/10.1049/cdt2.12015>.

Han, Y., Li, K., Ge, F., 2019. Multiple fault diagnosis for sucker rod pumping systems based on matter element analysis with F-statistics. In: 2019 IEEE 8th Data Driven Control and Learning Systems Conference (DDCLS), pp. 66–70. <https://doi.org/10.1109/DDCLS.2019.8908910>.

Han, Y., Li, K., Ge, F., et al., 2022. Online fault diagnosis for sucker rod pumping well by optimized density peak clustering. *ISA Trans.* 120, 222–234. <https://doi.org/10.1016/j.isatra.2021.03.022>.

He, P., Pan, J., Li, Y., 2022a. Long-term dam behavior prediction with deep learning on graphs. *Journal of Computational Design and Engineering* 9 (4), 1230–1245. <https://doi.org/10.1093/jcde/qwac051>.

He, Y.P., Zang, C.Z., Zeng, P., et al., 2023a. Few-shot working condition recognition of a sucker-rod pumping system based on a 4-dimensional time-frequency signature and meta-learning convolutional shrinkage neural network. *Petrol. Sci.* 20 (2), 1142–1154. <https://doi.org/10.1016/j.petsci.2023.02.017>.

He, Y.P., Zang, C.Z., Zeng, P., et al., 2023b. Convolutional shrinkage neural networks based model-agnostic meta-learning for few-shot learning. *Neural Process. Lett.* 55 (1), 505–518. <https://doi.org/10.1007/s11063-022-10894-7>.

He, Y.P., Zang, C.Z., Zeng, P., et al., 2022b. Automatic recognition of sucker-rod

pumping system working conditions using few-shot indicator diagram based on meta-learning. *Lecture Notes on Data Engineering and Communications Technologies* 80, 436–444. https://doi.org/10.1007/978-3-030-81007-8_49.

Lee, H., Lee, J., Kim, H., et al., 2021. Dataset and method for deep learning-based reconstruction of 3D CAD models containing machining features for mechanical parts. *Journal of Computational Design and Engineering* 9 (1), 114–127. <https://doi.org/10.1093/jcde/qwab072>.

Li, K., Gao, X.W., Zhou, H.B., et al., 2015. Fault diagnosis for down-hole conditions of sucker rod pumping systems based on the FBH–SC method. *Petrol. Sci.* 12 (1), 135–147. <https://doi.org/10.1007/s12182-014-0006-5>.

Li, K., Gao, X.W., Tian, Z., et al., 2013a. Using the curve moment and the PSO-SVM method to diagnose downhole conditions of a sucker rod pumping unit. *Petrol. Sci.* 10 (1), 73–80. <https://doi.org/10.1007/s12182-013-0252-y>.

Li, K., Gao, X.W., Yang, W.B., et al., 2013b. Multiple fault diagnosis of down-hole conditions of sucker-rod pumping wells based on Freeman chain code and DCA. *Petrol. Sci.* 10 (3), 347–360. <https://doi.org/10.1007/s12182-013-0283-4>.

Li, K., Han, Y., Wang, T., 2018. A novel prediction method for down-hole working conditions of the beam pumping unit based on 8-directions chain codes and online sequential extreme learning machine. *J. Petrol. Sci. Eng.* 160, 285–301. <https://doi.org/10.1016/j.petrol.2017.10.052>.

Liu, J., Feng, J., Xiao, Q., et al., 2021. Fault diagnosis of rod pump oil well based on support vector machine using preprocessed indicator diagram. In: 2021 IEEE 10th Data Driven Control and Learning Systems Conference (DDCLS), pp. 120–126. <https://doi.org/10.1109/DDCLS52934.2021.9455702>.

Lv, X.X., Wang, H.X., Xin, Z., et al., 2021a. Adaptive fault diagnosis of sucker rod pump systems based on optimal perceptron and simulation data. *Petrol. Sci.* 19 (2), 743–760. <https://doi.org/10.1016/j.petsci.2021.09.012>.

Lv, X.X., Feng, L., Wang, H.X., et al., 2021b. Quantitative diagnosis method of the sucker rod pump system based on the fault mechanism and inversion algorithm. *J. Process Control* 104, 40–53. <https://doi.org/10.1016/j.jprocont.2021.06.001>.

Lv, X.X., Wang, H.X., Zhang, X., et al., 2021c. An evolutionary SVM method based on incremental algorithm and simulated indicator diagrams for fault diagnosis in sucker rod pumping systems. *J. Petrol. Sci. Eng.* 203. <https://doi.org/10.1016/j.petrol.2021.108806>.

Pan, Y., An, R., Fu, D., et al., 2021. Unsupervised fault detection with a decision fusion method based on Bayesian in the pumping unit. *IEEE Sensor. J.* 21 (19), 21829–21838. <https://doi.org/10.1109/JSEN.2021.3103520>.

Sharaf, S.A., 2018. Beam pump dynamometer card prediction using artificial neural networks. *KnE Engineering* 3 (7), 198. <https://doi.org/10.18502/keg.v3i7.3083>.

Song, C.H., Liu, S., Han, G., et al., 2023. Edge intelligence-based condition monitoring of beam pumping units under heavy noise in industrial internet of things for industry 4.0. *IEEE Internet Things J.* 10 (4), 3037–3046. <https://doi.org/10.1109/JIOT.2022.3141382>.

Sun, L., Shi, H., Bai, M., 2022. Intelligent oil well identification modelling based on deep learning and neural network. *Enterprise Inf. Syst.* 16 (2), 249–263. <https://doi.org/10.1080/17517575.2020.1722252>.

Tagirova, C., Vulfin, A., Arslanova, A., 2021. Data Mining of the Dynamometry of Oil Production Sucker Rod Pumping Unit. 2021 International Russian Automation Conference. *RusAutoCon*, pp. 835–840. <https://doi.org/10.1109/RusAutoCon52004.2021.9537412>.

Tian, H., Deng, S., Wang, C., et al., 2021. A novel method for prediction of paraffin deposit in sucker rod pumping system based on CNN indicator diagram feature deep learning. *J. Petrol. Sci. Eng.* 206, 108986. <https://doi.org/10.1016/j.petrol.2021.108986>.

Wang, X., He, Y., Li, F., et al., 2021. A working condition diagnosis model of sucker rod pumping wells based on deep learning. *SPE Prod. Oper.* 36 (2), 317–326. <https://doi.org/10.2118/205015-pa>.

Wang, Y., Wei, Z., Yang, J., 2019. Feature trend extraction and adaptive density peaks search for intelligent fault diagnosis of machines. *IEEE Trans. Ind. Inf.* 15 (1), 105–115. <https://doi.org/10.1109/TII.2018.2810226>.

Xu, P., Xu, S., Yin, H., 2007. Application of self-organizing competitive neural network in fault diagnosis of suck rod pumping system. *J. Petrol. Sci. Eng.* 58 (1–2), 43–48. <https://doi.org/10.1016/j.petrol.2006.11.008>.

Yang, X., Gao, F., Lu, R., et al., 2022. Lightweight aerial object detection method based on improved YOLOv5. *Inf. Control* 51 (3), 361–368. <https://doi.org/10.13976/j.cnki.kk.2021.1240> (in Chinese).

Ye, Z., Liu, Z., Cheng, C., et al., 2020. Efficient evaluation model of beam pumping unit based on principal component regression analysis. *Sci. Prog.* 103 (1), 1–22. <https://doi.org/10.1177/0036850419895769>.

Yin, C., Zhang, K., Zhang, L., et al., 2023. Imbalanced working states recognition of sucker rod well dynamometer cards based on data generation and diversity augmentation. *SPE J.* 1–20. <https://doi.org/10.2118/214661-pa>.

- Yu, C., Ning, Y., Qin, Y., et al., 2021. Multi-label fault diagnosis of rolling bearing based on meta-learning. *Neural Comput. Appl.* 33 (10), 5393–5407. <https://doi.org/10.1007/s00521-020-05345-0>.
- Zhang, A., Gao, X., 2019. Supervised dictionary-based transfer subspace learning and applications for fault diagnosis of sucker rod pumping systems. *Neuro-computing* 338, 293–306. <https://doi.org/10.1016/j.neucom.2019.02.013>.
- Zhang, K., Wang, Q., Wang, L., et al., 2022. Fault diagnosis method for sucker rod well with few shots based on meta-transfer learning. *J. Petrol. Sci. Eng.* 212. <https://doi.org/10.1016/j.petrol.2022.110295>.
- Zhang, R., Yin, Y., Xiao, L., et al., 2021. A real-time diagnosis method of reservoir-wellbore-surface conditions in sucker-rod pump wells based on multidata combination analysis. *J. Petrol. Sci. Eng.* 198, 108254. <https://doi.org/10.1016/j.petrol.2020.108254>.
- Zhao, H., Jian, W., Peng, G., 2017. A deep learning approach for condition based monitoring and fault diagnosis of rod pump system. *Services Transactions on Internet of Things* 1 (1), 32–42. <https://doi.org/10.29268/stiot.2017.1.1.3>.
- Zheng, B., Gao, X., 2017. Sucker rod pumping diagnosis using valve working position and parameter optimal continuous hidden Markov model. *J. Process Control* 59, 1–12. <https://doi.org/10.1016/j.jprocont.2017.09.007>.
- Zheng, B., Gao, X., Li, X., 2019a. Diagnosis of sucker rod pump based on generating dynamometer cards. *J. Process Control* 77, 76–88. <https://doi.org/10.1016/j.jprocont.2019.02.008>.
- Zheng, B., Gao, X., Li, X., 2019b. Fault detection for sucker rod pump based on motor power. *Control Eng. Pract.* 86, 37–47. <https://doi.org/10.1016/j.conengprac.2019.02.001>.
- Zheng, B., Gao, X., Pan, R., 2020. Sucker rod pump working state diagnosis using motor data and hidden conditional random fields. *IEEE Trans. Ind. Electron.* 67 (9), 7919–7928. <https://doi.org/10.1109/TIE.2019.2944081>.
- Zhou, B., Wang, Y., Liu, W., et al., 2018. Identification of working condition from sucker-rod pumping wells based on multi-view co-training and hessian regularization of SVM. 2018 14th. In: *IEEE International Conference on Signal Processing (ICSP)*, pp. 969–973. <https://doi.org/10.1109/ICSP.2018.8652454>.
- Zhou, W., Li, X., Yi, J., et al., 2019. A novel UKF-RBF method based on adaptive noise factor for fault diagnosis in pumping unit. *IEEE Trans. Ind. Inf.* 15 (3), 1415–1424. <https://doi.org/10.1109/TII.2018.2839062>.