



## Original Paper

# The real-time dynamic liquid level calculation method of the sucker rod well based on multi-view features fusion



Cheng-Zhe Yin<sup>a</sup>, Kai Zhang<sup>a, b, \*</sup>, Jia-Yuan Liu<sup>c</sup>, Xin-Yan Wang<sup>a</sup>, Min Li<sup>a</sup>, Li-Ming Zhang<sup>a</sup>, Wen-Sheng Zhou<sup>d, e</sup>

<sup>a</sup> School of Petroleum Engineering, China University of Petroleum, Qingdao, 266580, Shandong, China

<sup>b</sup> School of Civil Engineering, Qingdao University of Technology, Qingdao, 266580, Shandong, China

<sup>c</sup> Tarim Oilfield Company, China National Petroleum Corporation, Kuerle, 841000, Xinjiang, China

<sup>d</sup> State Key Laboratory of Offshore Oil Exploitation, Beijing, 100028, China

<sup>e</sup> CNOOC Research Institute Ltd., Beijing, 100028, China

## ARTICLE INFO

## Article history:

Received 30 July 2023

Received in revised form

12 March 2024

Accepted 8 May 2024

Available online 9 May 2024

Edited by Jie Hao and Meng-Jiao Zhou

## Keywords:

Dynamic liquid level

Multi view

Features fusion

Sucker rod well

Dynamometer cards

## ABSTRACT

In the production of the sucker rod well, the dynamic liquid level is important for the production efficiency and safety in the lifting process. It is influenced by multi-source data which need to be combined for the dynamic liquid level real-time calculation. In this paper, the multi-source data are regarded as the different views including the load of the sucker rod and liquid in the wellbore, the image of the dynamometer card and production dynamics parameters. These views can be fused by the multi-branch neural network with special fusion layer. With this method, the features of different views can be extracted by considering the difference of the modality and physical meaning between them. Then, the extraction results which are selected by multinomial sampling can be the input of the fusion layer. During the fusion process, the availability under different views determines whether the views are fused in the fusion layer or not. In this way, not only the correlation between the views can be considered, but also the missing data can be processed automatically. The results have shown that the load and production features fusion (the method proposed in this paper) performs best with the lowest mean absolute error (MAE) 39.63 m, followed by the features concatenation with MAE 42.47 m. They both performed better than only a single view and the lower MAE of the features fusion indicates that its generalization ability is stronger. In contrast, the image feature as a single view contributes little to the accuracy improvement after fused with other views with the highest MAE. When there is data missing in some view, compared with the features concatenation, the multi-view features fusion will not result in the unavailability of a large number of samples. When the missing rate is 10%, 30%, 50% and 80%, the method proposed in this paper can reduce MAE by 5.8, 7, 9.3 and 20.3 m respectively. In general, the multi-view features fusion method proposed in this paper can improve the accuracy obviously and process the missing data effectively, which helps provide technical support for real-time monitoring of the dynamic liquid level in oil fields.

© 2024 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

The dynamic liquid level of a sucker rod well is the liquid level of the annular space between the tubing and the casing during the oil production. It reflects the relationship between the fluid supply and

discharge from the formation to the wellbore, and is also an important basis for determining a reasonable working operation for the oil well. The dynamic liquid level measurement methods by devices mainly consist of float, pressure, and acoustic measurement method (Chen et al., 2008; Zhang, 2003). The device measurement can provide more reliable results, but it has a higher cost and cannot be used for obtaining the dynamic liquid level in real-time. A large number of dynamic data (Mohammadpoor and Torabi, 2018; Nguyen et al., 2020) have been accumulated in the process

\* Corresponding author. School of Petroleum Engineering, China University of Petroleum, Qingdao, 266580, Shandong, China.

E-mail address: [zhangkai@upc.edu.cn](mailto:zhangkai@upc.edu.cn) (K. Zhang).

of oil well production, so these data can be fully used for real-time calculation of the liquid level by data driven methods.

Traditionally, the soft measurement based on dynamometer cards is always used as a data-driven method to calculate the dynamic liquid level (Zhang et al., 2007). In order to eliminate the influence of the dynamic load, the ground dynamometer cards were converted into the down-hole dynamometer cards to calculate the dynamic liquid level (Chen et al., 2015). Based on the down-hole dynamometer cards, the results reliability was improved by correcting the pressure gradient of the annular space (Zhang et al., 2011). The pressure in the oil tubing and the annular space is the key to determining the dynamic liquid level, which can be calculated more accurately by the multiphase pipe flow method (Yang, 2010; Lu, 2017). Although researchers have improved the calculation method based on dynamometer cards, there are still some assumptions hardly considered in the calculation formula due to the influence of some factors difficult to describe, such as mixed liquid density, equipment friction and so on. When these factors have obvious influence on the artificial lifting, the accuracy of the calculation results will be affected obviously as well. Therefore, there are limitations when calculating the dynamic liquid level with the existing methods. More advanced data-driven methods are needed to set up a more universal model.

Machine learning, as an advanced data-driven analysis method, can describe the complicated factors which are difficult to be considered in mechanical formula due to its special model structure and hyper-parameters. It is not limited to expression forms and has strong nonlinear fitting ability. Least squares support vector machine (LSSVM) was used with the Grey Wolf optimization algorithm (Tian, 2021). An incremental learning method was set up based on Gaussian regression to increase the calculation efficiency (Li et al., 2015). The Adaboost was used for dynamic liquid level prediction and whether the model need updated based on the prediction error index was determined (Wang et al., 2017). In order to solve the problem of insufficient data, generative adversarial network (GAN) was used for samples generation, which improved the calculation accuracy (Hou et al., 2019). Some researchers hold the view that the change of the dynamic liquid level have chaotic characteristics. The chaotic time series prediction method was used to study the prediction of dynamic liquid level time series with a maximum Lyapunov index prediction model set up (Yu et al., 2018). The chaotic time series was taken as the features and extreme learning machine (ELM) was used with PSO optimization to calculate the dynamic liquid level (Yu, 2020). There are some researchers interested in the dynamic liquid level prediction based on signal analysis. The resonant frequency difference (RFD) of the annular resonant acoustic signal was extracted to calculate the dynamic liquid level (Zhou et al., 2018). Adaptive filtering was used

to extract features and the dynamic liquid level prediction model was set up based on long short-term memory (LSTM) (Liang and Zhang, 2021). The time-frequency features were extracted by interactive self-organizing data analysis technique algorithm (ISODATA) and Gaussian process regression (GPR) was set up to calculate dynamic liquid level (Li et al., 2016). By the fusion of the physical process and machine learning methods, an analytical model was set up with stochastic configuration networks (SCN) as the compensation to obtain more accurate calculation results (Han et al., 2022).

Various machine and deep learning algorithms have been widely applied in the dynamic liquid level calculation. In terms of the feature engineering, researchers not only used dynamometer cards, pressure and so on as features, but also extracted the chaotic time series and time-frequency features. For describing the dynamic liquid level more accurately, these features were concatenated as well. Therefore, the features usually come from different data sources (multi-view data) in the data-driven process. The key to the research is how to choose the views and combine them. In terms of the dynamic liquid level calculation, there are few researches of multi-view data fusion and features interaction. In fact, there are differences between the views in data form and physical meaning. Based on the accurate features extraction results, the differences should be eliminated by the fusion and then different views should complement each other.

The multi-view data fusion methods can be divided into early fusion, late fusion and hybrid fusion (Baltrusaitis et al., 2018). Early fusion is always the concatenation of the original features. Late fusion is usually the fusion of the results (labels). Hybrid fusion, which falls between the two, is the most popular method for features fusion. Stack auto encoder was used for the multi-models feature extraction and the extracted results and their residuals were concatenated into new features for fault diagnosis (Li et al., 2020a). An FAC-CNN network, which does not need complicated data preprocessing and is not limited by the data source channel, was proposed when performing data fusion (Li et al., 2020b). A parallel convolutional neural network (PCNN) was proposed for multi-sensor features fusion which has been demonstrated effective by tool monitoring and bearing fault diagnosis experiments (Xu et al., 2020). Based on the above research results, it can be seen that multi-view data fusion has been used in the industry widely.

Similarly, in petroleum industry, multi-view data need be integrated when calculating the dynamic liquid level. In the artificial lifting of the sucker rod well, the pump is always under the wellbore liquid, which indicates that the dynamic liquid level is shallower than the depth of the pump. The submergence pressure changes with the liquid column load in the upstroke. In this process, the submergence pressure is also influenced by the dynamic

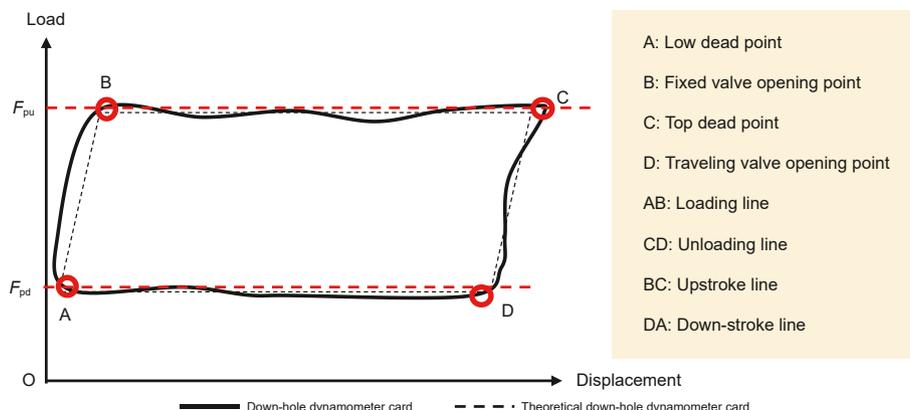


Fig. 1. Down-hole dynamometer card.

**Table 1**  
Detailed features used for dynamic liquid level calculation.

Name of different views	Name of detailed features
Load	Load of the low dead point Load of the fixed valve opening point Load of the travelling valve opening point Load of the top dead point Average load of the upstroke line Average load of the down-stroke line Maximum load Minimum load
Production	Wellhead pressure Casing pressure Return pressure Water cut Pump depth Pump diameter Stroke Stroke frequency
Image features of down-hole dynamometer cards	Image grayscale matrix

liquid level and casing pressure. With the submergence pressure as a bond, the dynamic liquid level is closely related with the load, pressure, fluid properties and pumping parameters. Therefore, it is suitable to use the multi-view features fusion method for dynamic liquid level calculation. However, the data quantity and quality of different sources vary greatly with different data acquisition facilities or test difficulty. When the acquisition frequency is different, there may be even several-times difference in quantity between data sources. If there are missing or abnormal data, the difference of data quality will be more obvious. Therefore, during the features fusion, it is necessary to consider not only the correlation but also the difference between different views. In order to solve the above problems, the EmbraceNet network structure (Choi and Lee, 2019) is used and its classification structure has been transformed into a regression structure in this paper. For adapting the data type and quantity of this paper, the network structure and parameters have been modified as well.

In summary, the main content of this paper is as follows. Firstly, according to the mechanical equilibrium relationship between the

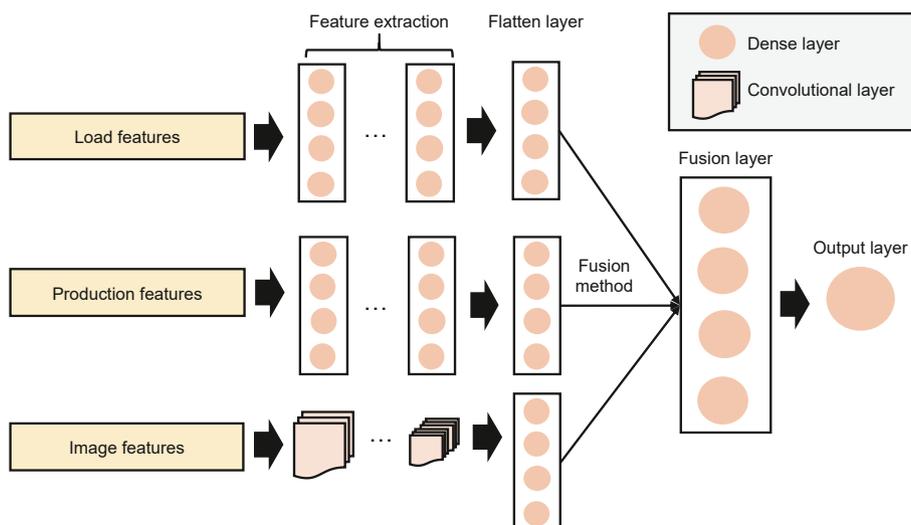
sucker rod string and the fluid in the wellbore, the features that affect the dynamic liquid level were analyzed (Section 2). Secondly, a multi branch neural network was set up for features extraction, and then a special fusion layer was added to the network for features fusion (Section 3). Thirdly, the comparative experiments of different views combinations were conducted to analyze the interaction between different features and then get the combinations with good performance (Section 4.1). At last, based on the above results, the comparative experiments with different missing ratios were conducted to demonstrate the advantage of the proposed method in this paper when processing missing data automatically (Section 4.2).

The highlights of the paper consist of the following two parts. Firstly, the real-time dynamic liquid level calculation model can be more accurate and generalized by multi-source data fusion and the interaction between features can be analyzed. Secondly, this method can guarantee that the model still has higher accuracy and the data can be utilized more fully even if there are missing data.

## 2. Features analysis

The soft measurement method is mainly based on dynamometer cards and production dynamic data. Due to the effect of dynamic loads, there are deviations when calculating dynamic liquid level by surface dynamometer cards. The surface dynamometer cards can be converted into down-hole dynamometer cards by solving the one-dimensional damped wave equation with Fourier series and taking surface load and displacement as the boundary conditions. In this way, the calculation deviations caused by dynamic loads such as vibration, inertia, and friction can be eliminated to some degree (Gibbs, 1963; Zhang and Wu, 1984).

In the upstroke process, the travelling valve is closed. When the submergence pressure is greater than the pressure in the pump, the fixed valve opens. In the downstroke process, the fixed valve is closed. When the pressure in the pump is greater than the liquid pressure, the travelling valve opens. After the loading in the upstroke and the unloading in the downstroke, the load difference from the bottom of the sucker rod string can be expressed as Eq. (1) (Chen et al., 2015)



**Fig. 2.** Features fusion route.

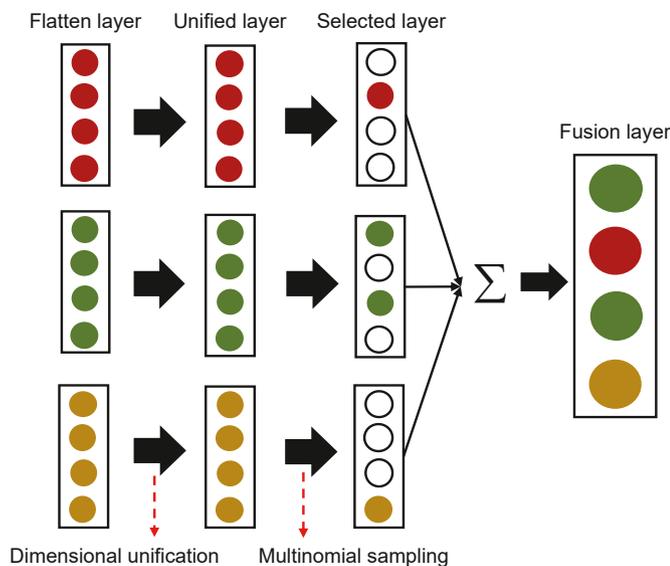


Fig. 3. The structure of EmbraceNet model.

$$F_{pu} - F_{pd} = (p_p - p_n + \Delta p)f_p + 2f \tag{1}$$

where  $p_p$  is the pump discharge pressure,  $p_n$  is the submergence pressure,  $\Delta p$  is the pressure drop when the fluid passes through the travelling and fixed valve,  $f_p$  is the cross-sectional area of the plunger, and  $f$  is the friction between plunger and pump barrel. Therefore, the submergence pressure can be expressed as Eq. (2)

$$p_n = (p_p + \Delta p) - \frac{F_{pu} - F_{pd}}{f_p} + \frac{2f}{f_p} \tag{2}$$

The pump discharge pressure can be calculated by the wellbore multiphase flow rule. The load difference  $F_{pu} - F_{pd}$  can be obtained by down-hole dynamometer cards (Fig. 1).

At last, the dynamic liquid level can be calculated based on the casing pressure, pump depth, submergence pressure and the pressure distribution of the wellbore and annular space.

Since the actual production is more complicated, complex phenomenon cannot be reflected by the above ideal physical formula, such as friction between wellbore equipment, load deviation caused by paraffin deposition and so on. However, they can be learned automatically by machine or deep learning algorithms, which can help set up a more universal calculation method. Based on the above analysis, it can be seen that real-time calculation of the dynamic liquid level needs the load of dynamometer cards, dynamic production data, and pumping parameters. Due to the modal differences and information overlap among various data, in this paper, the above features were fused by the multi-view learning. Data sources include the following three views:

- 1) The dynamometer cards measured in real time by the dynamometer,
- 2) the pressure and water cut measured in real time by wellhead sensors,
- 3) the pumping parameters such as pump depth, pump diameter, stroke, and stroke rate set by technical staff according to the actual condition.

The dynamometer cards and production dynamic data can be collected in real-time based on the sensors. However, the pumping parameters usually remain constant in a period of time, which is not fit for a single view due to the lack of richness. Therefore, the pumping parameters and production dynamic data are combined together as a single view (Table 1).

In the load view, real-time load of sucker rod in upstroke and down-stroke can be provided by the load of fixed and travelling valve opening point, load of the top and low dead point, average load and maximum load. In the production view, the fluid dynamic information in the wellbore can be provided by the pressure and water cut. The liquid productivity of the mechanical facilities can be reflected by the pump depth, pump diameter, stroke and stroke frequency. The image shape of the down-hole dynamometer cards (hereinafter referred to as dynamometer cards) is able to show different working conditions of the oil well, so it can be supplemented as the third view.

### 3. Feature fusion method

#### 3.1. Background of theoretical methods

Based on Section 2, there are differences in terms of the physical meaning between production and load features, and they are different from dynamometer cards images in data form as well. Therefore, these three kinds of data show multi-view characteristics. The data dimension from different views need unified before features fusion by features extraction. Due to the flexibility of neural network, data from different views can be input independently at the same time and then features are further extracted by feedforward propagation (Fig. 2).

It can be seen that the production and load features are extracted by the dense layers, while the dynamometer card image features are extracted by the convolutional and pooling layers. Then, the extracted results were unified in dimension by the flatten layers. At last, the outputs of the flatten layers can be fused with proper fusion method to calculate dynamic liquid level.

A multi-view fusion structure called EmbraceNet has been proposed (Fig. 3) (Choi and Lee, 2019). Its good performance has been verified on the MNIST datasets (LeCun et al., 1998), gas sensor arrays dataset (Vergara et al., 2013), and OPPORTUNITY dataset (Chavarriga et al., 2013). With this method, the outputs of the flatten layers from different branch networks are selected by multinomial sampling and fused by adding them together. In this way, each neuron of the fusion layer is contributed by only one view. So, all the neurons of the fusion layer contain all the views information without mutual interference.

After features extraction, the results are flattened and unified to the same dimension. Multinomial sampling is used to select

Table 2 Comparison of data imputation methods.

Research filed	Characteristics	Limitations	References
Interpolation	It is simple and easy to be implemented.	It cannot reproduce the complicated rules of real data.	Ordóñez and Roggen (2016).
Data augmentation	It can fill miss data by generating new data directly.	The generation results have randomness and its distribution may be different from real data.	Hou et al. (2022); Liguori et al. (2023); Eitel et al. (2015).
Model structure improvement	It can handle missing data during model training.	It is limited to specific model architectures or numbers of modalities.	Jaques et al. (2017); Srivastava and Salakhutdinov (2012); Gu et al. (2017).

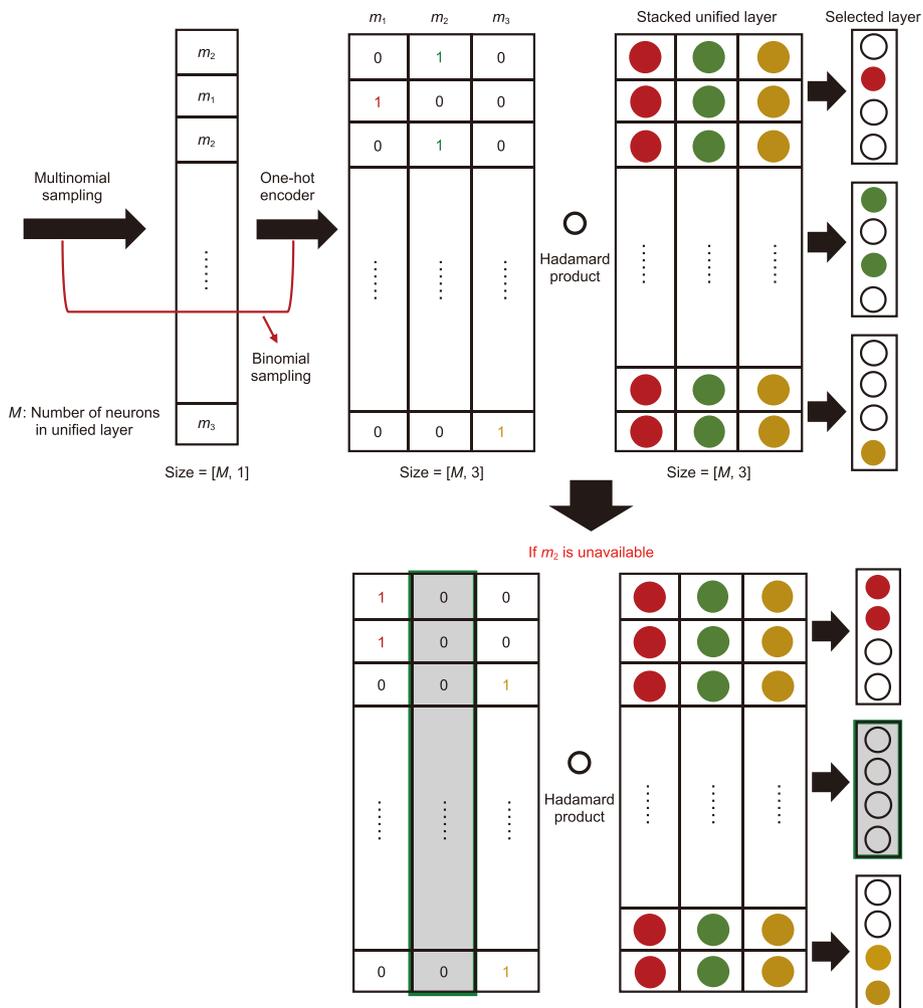


Fig. 4. Calculation process of neurons selection with multinomial sampling.

Table 3  
Branch network structure and parameters of load and production features.

Layer name	Activation function	Dimension	
		Load features	Production features
Input layer	/	8	8
Dense layer1	Relu	100	100
Dense layer2	Relu	100	100
Dense layer3	Relu	50	50

neurons of the unified layers, and the selected neurons are fused by addition to obtain the fusion layer. If there are missing data in some view, the multinomial sampling probability in the corresponding neurons is 0. This network can ignore the missing contents without

Table 4  
Network structure and parameters of load and production features concatenation.

Layer name	Activation function	Dimension
		Load and production features concatenation
Input layer	/	16
Dense layer1	Relu	100
Dense layer2	Relu	100
Dense layer3	Relu	50

influence on other views. In this way, the training and validation process can be conducted normally without completing miss contents, which is different from imputing the missing data directly (Table 2).

It can be seen that although these methods have been applied in different fields successfully, there some limitations to some degree. EmbraceNet can provide a more flexible way to handle missing data without data and model structure limitations and then can be used for multi-view features fusion in various scenarios. The detailed calculation steps are as follows:

Assuming a single sample is input, the detailed calculation process by multinomial sampling is conducted (Fig. 4). The view number is denoted as  $m_j$ . In this paper,  $m_1, m_2, m_3$  represent load features, production features and image features respectively.

Assuming that there are  $N$  views when the data of all views are available, the sampling probability for each view by multinomial sampling can be expressed as:

$$p = \left[ \frac{1}{N}, \frac{1}{N}, \dots, \frac{1}{N} \right] \in \mathbb{R}^N \tag{3}$$

Assuming that there are  $M$  neurons in the unified layer for each view, the sampling results can be expressed as a one-dimension vector  $r \in \mathbb{R}^M$  consisting of the view number  $m_j$ , where

**Table 5**  
Branch network structure and parameters of dynamometer cards image features.

Layer name	Activation function	Dimension	Stride value
Input layer	/	32 × 32 × 1	/
Convolutional layer1	Relu	30 × 30 × 16	(3,3)
Pooling layer1	/	15 × 15 × 16	(2,2)
Convolutional layer2	Relu	13 × 13 × 32	(3,3)
Pooling layer2	/	6 × 6 × 32	(2,2)
Convolutional layer3	Relu	4 × 4 × 64	(3,3)
Pooling layer3	/	2 × 2 × 64	(2,2)
Convolutional layer4	Relu	1 × 1 × 128	(2,2)
Flatten layer	Relu	128 × 1	/

$$r_k \in \{m_j | j = 1, 2, \dots, N\}, k = 1, 2, \dots, M \quad (4)$$

By encoding  $r$  with one hot encoding, the matrix  $\hat{r} \in \mathbb{R}^{M \times N}$  can be obtained, where

$$\hat{r}_{ij} \in \{0, 1\}, i = 1, 2, \dots, M, j = 1, 2, \dots, N \quad (5)$$

Stack the corresponding unified layer of each view, the matrix  $\hat{u} = [u_{m_1}, u_{m_2}, \dots, u_{m_N}] \in \mathbb{R}^{M \times N}$  can be obtained. Then Hadamard product  $\hat{s} = \hat{r} \circ \hat{u}$  can be expressed as follows:

$$\hat{s} = [s_{m_1}, s_{m_2}, \dots, s_{m_N}] \in \mathbb{R}^{M \times N} \quad (6)$$

where  $s_{m_i}$  is the multinomial sampling result of unified layer for  $m_i$ . By adding them, the result of features fusion  $f$  can be obtained and expressed as

$$f = \sum_{i=1}^N s_{m_i} \in \mathbb{R}^M \quad (7)$$

One-hot encoding is equivalent to turning multinomial sampling into binomial sampling. In this way, only one view can be selected for each row in the stacked unified layer, and therefore, each neuron in the fusion layer is contributed by only one view. This sampling method not only considers the correlation between views, but also avoids information redundancy caused by excessive fusion between views. In addition, the binomial sampling method can remove some neurons in the unified layer during training, which can further prevent the model from overfitting.

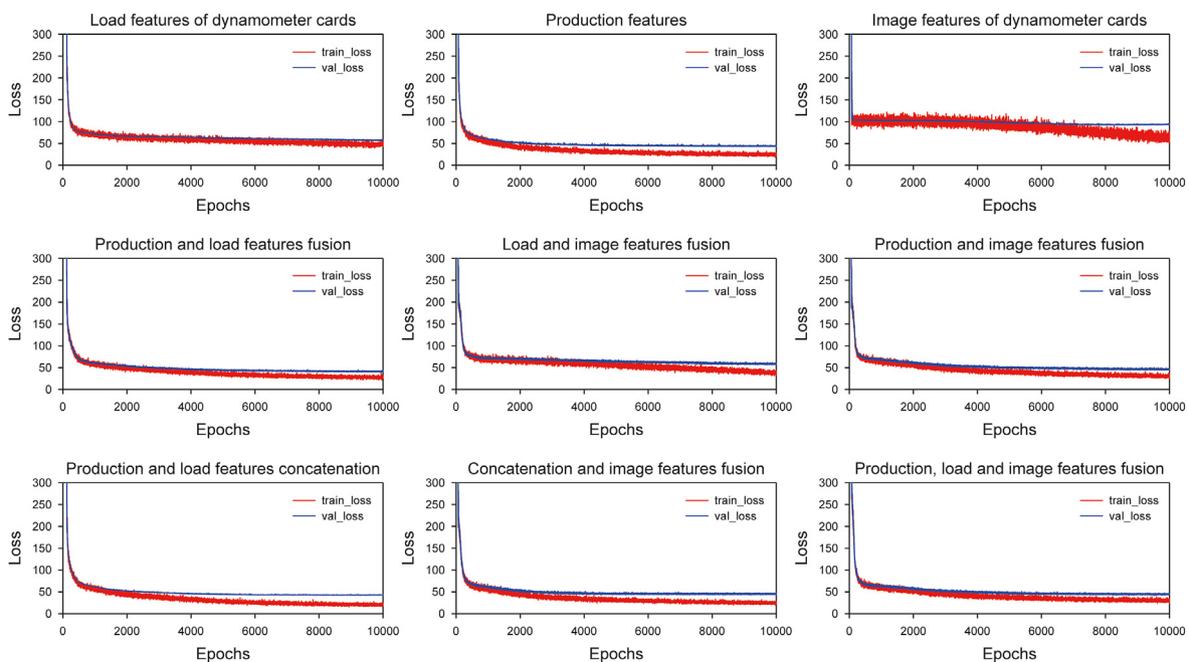
When the data of some view is unavailable, the sampling probability of multinomial sampling in this view should be zero. Therefore, assuming that there are  $T$  views unavailable, the sampling probability of each view can be expressed as

$$p = \left[ \dots, \frac{1}{N-T}, \dots, 0, \dots, \frac{1}{N-T}, \dots \right] \in \mathbb{R}^N \quad (8)$$

The following calculation process is the same as that without unavailable data. Assuming that there are missing data in  $m_2$ , the detailed calculation process is shown in Fig. 4. It can be seen that the view with missing data will be ignored during the multinomial sampling, so the information of this view no longer exists in the fusion layer and is also no longer involved in the backpropagation. Other views without missing data can still be involved in feature fusion and model training. Therefore, if there are missing data for

**Table 6**  
Network structure and parameters from unified layer to output layer.

Layer name	Activation function	Dimension		
		Load features	Production features	Image features
Input layer	/	50	50	128
Unified layer	Relu	50	50	50
Fusion layer	/	50		
Output layer	Relu	1		



**Fig. 5.** Loss curves of different features combination.

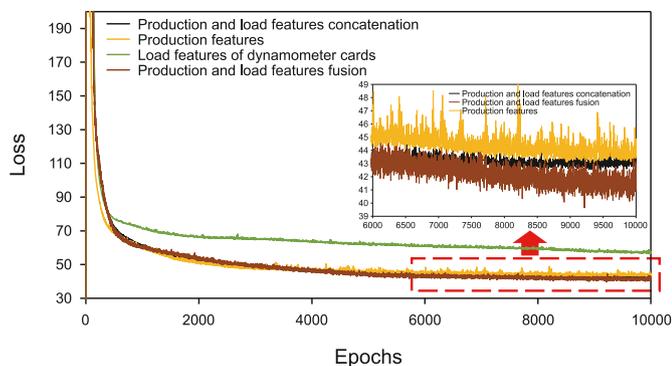


Fig. 6. Validation set loss curves of production and load features.

some view, it can be avoided that a large number of samples are unavailable.

### 3.2. Model structure and parameters

The network consists of the input layer, branch networks of different views (load, production and image features), fusion layer and output layer. The branch networks of the structured features (load and production features) consist of the dense layers with the same network structure (Table 3).

When directly concatenating the load and production features, it is only needed to merge the two branches (Table 3) and then the new network structure can be gotten (Table 4).

The branch network of unstructured features (dynamometer cards image features) consists of convolutional and pooling layers (Table 5).

The dimensions from the unified layer to the output layer are the same in all views (Table 6).

## 4. Experiment results analysis

### 4.1. Model performance of multi-view features under different combination forms

The experiment data in this paper comes from the Fifth Oil Production Plant of XX Oilfield, covering 196 sucker rod wells. The

Table 7  
MAE of the best models under different features combination of the validation set.

Names of features combination	MAE of the best model, m
Load features of dynamometer cards	56.47
Production features	43.07
Image features of dynamometer cards	92.61
Production and load features fusion	39.63
Load and image features fusion	56.35
Production and image features fusion	44.34
Production and load features concatenation	42.47
Concatenation and image features fusion	42.54
Production, load and image features fusion	42.16

dynamic liquid level data were selected from these wells for the past two years. After data processing, a total of 2901 samples were obtained, of which 70% were used as the training set and 30% were used as the validation set. Both in training and validation stage, the dynamic liquid level data are all from real monthly measurements.

Based on Sections 2 and 3, dynamic liquid level calculation models of single view and different combinations of views were set up, including 1) load features of dynamometer cards, 2) production features, 3) image features of dynamometer cards, 4) production and load features fusion, 5) load and image features fusion, 6) production and image features fusion, 7) production and load features concatenation, 8) concatenation and image features fusion and 9) production, load and image features fusion.

To get the complete training process, the iteration number was set to 10,000 after a number of attempts. Then the loss change of training and validation can be obtained (Fig. 5).

It can be seen that the training loss fluctuates greatly when the model is trained by the image grayscale matrix of dynamometer cards. After 6000 epochs, it began to decline. The loss fluctuation of the validation set is more stable, but the loss is still high and there is no downward trend in the later stage. In contrast, the loss fluctuations of the other experiments are much smaller in the training and validation sets, which indicates that the training process is more stable. The loss gap of the production and load features concatenation between the training and validation set is most obvious, which indicates that its overfitting is more significant compared with the others.

The loss curves of validation set before and after fusion or concatenation of production and load features can be obtained (Fig. 6).

It can be seen that the validation loss of the single feature

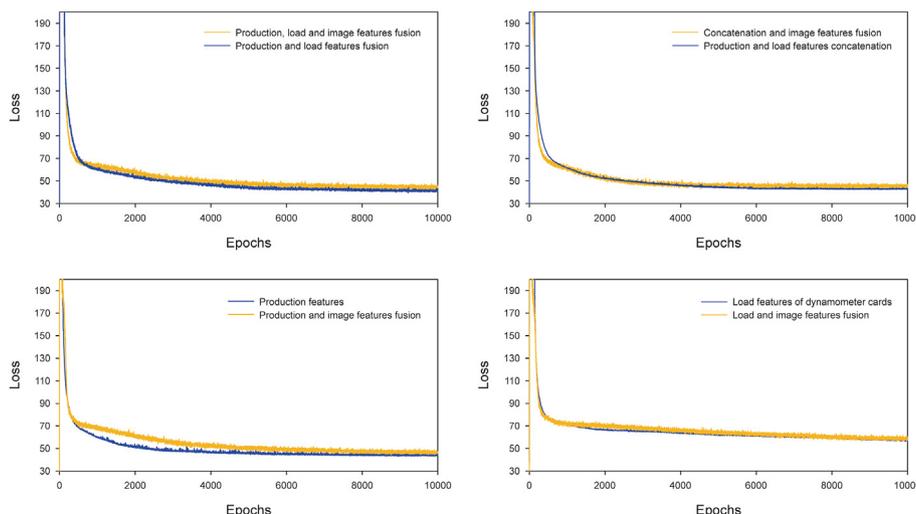


Fig. 7. Validation set loss curves of image, production and load features.

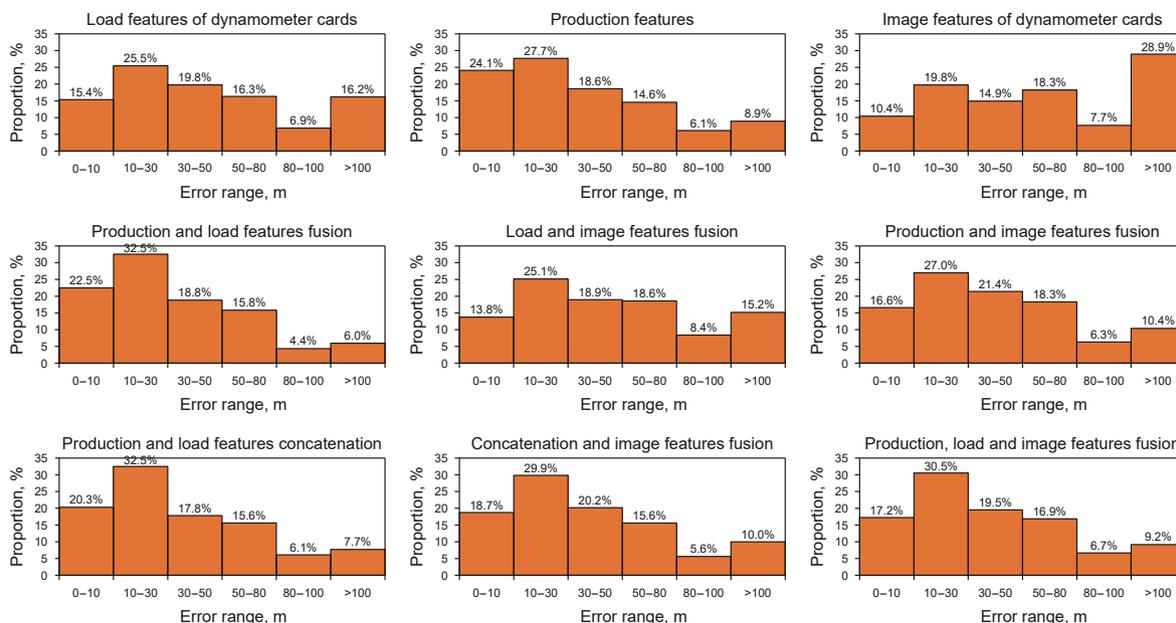


Fig. 8. Error proportion distribution histogram.

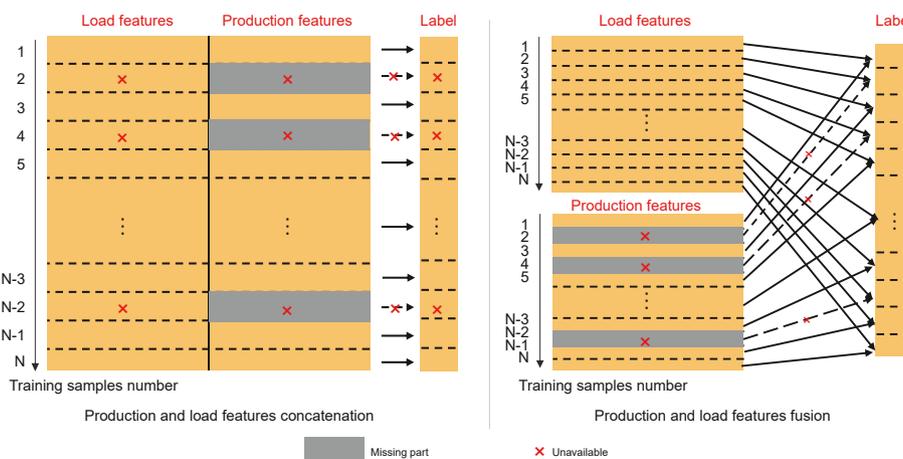


Fig. 9. Handling ways of missing data of features concatenation and fusion.

especially the load feature is generally greater than that of features fusion or concatenation. In the early stage of training, due to sufficient information of features concatenation in the input layer, its loss is similar to that of features fusion. Unlike that, the features fusion layer of multi-view neural network is close to the output layer, so its performance should be improved by gradual training to get accurate network parameters. Therefore, the loss of features fusion is lower than that of features concatenation after 6000 epochs, and the ability of information fusion from different views gradually highlights.

The loss curves of the validation set after fusing production and load features with the image feature of the dynamometer cards can be obtained (Fig. 7).

It can be seen that the image feature of the dynamometer card has not improved the performance of models trained with production and load features. Although the shape of the dynamometer card is closely related to the working conditions of the sucker rod well, there is a significant gap between the dynamic liquid level calculation and working conditions classification. The one-

dimensional form of the dynamometer card can be obtained from the original two-dimensional space of the grayscale matrix by convolutional and pooling processing. In this process, the shape information of the image can be more refined. However, the special points which can reflect the changes of the dynamic liquid level on the dynamometer card are far less obvious after this process. Therefore, the load features inside the images may be weakened, resulting in poor performance of the model.

Based on the validation loss of different features combinations, the models with the best performance can be obtained. Their performance can be measured by mean absolute error (MAE) (Table 7).

It can be seen that the model only trained with the image feature has the worst performance. The model trained by the production and load features fusion has the lowest MAE, followed by the production and load features concatenation. Because the features fusion can take differences and interactions between features into account, its performance is better than the production and load features concatenation. According to the best models of different feature combinations, the error distribution of the

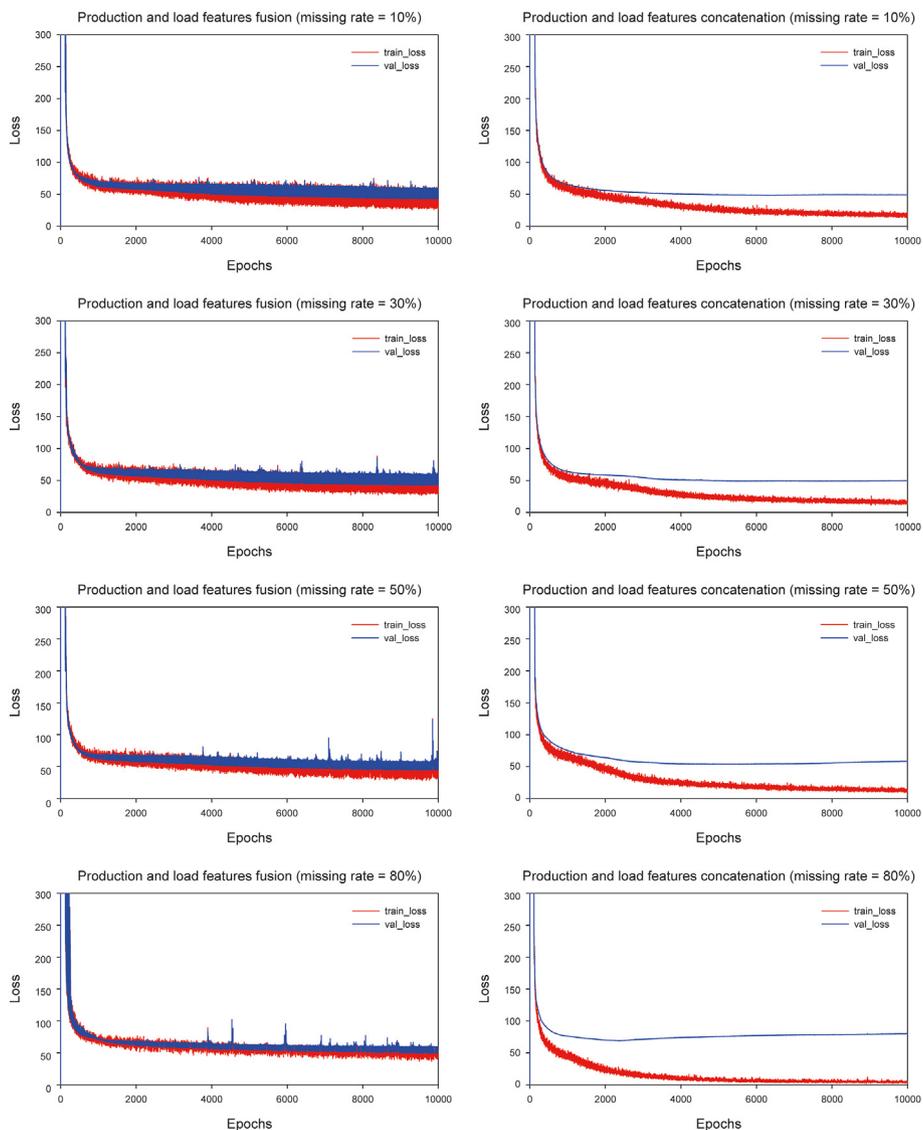


Fig. 10. Model performance under different data missing rates.

validation set can be calculated (Fig. 8).

It can be seen that the samples proportion of production and load feature fusion is lowest in the high error range (>100 m), at 6.0%, followed by production and load feature concatenation, at 7.7%. And the samples proportion of the image features in this range is highest, at 28.9%. In the low error range (<50 m), the samples proportion of production and load features fusion is highest, at 73.9%, followed by production and feature concatenation, at 70.6%. Therefore, the result of the production and load features fusion is more accurate and credible, which provides higher practical value compared with the other features combinations.

#### 4.2. Results comparison under different data missing rates

Based on the results in Section 4.1, only the production and load features fusion and concatenation were worth studying further due to their better performance than the image feature. We assume that there are missing data in the production features view, and evaluate the model performance under different missing rates between features fusion (multi-view calculation method) and concatenation (traditional calculation method). There is difference of dealing with

missing data between features fusion and concatenation (Fig. 9). Compared with the features concatenation, only the view which contains missing samples needs to be removed when using the features fusion instead of removing the samples entirely.

The comparison experiments were conducted when production features were lack of 10%, 30%, 50%, and 80% (Fig. 10). It can be seen that by features concatenation, the loss values of the validation set rise with the increase of missing rates, but that of the training set decreases, indicating that the overfitting is greatly serious. By features fusion implemented by multi-view network, the loss values of the training and validation sets both rise with the increase of missing rates.

It is worth noting that when there are missing samples, the loss values fluctuate greater during the training of the multi-view network. Because random batch samples are used for training, at different epochs, the samples with or without missing views appear alternatively, which results in significant differences among loss values. In fact, although the fluctuation causes instability of the training, it can provide more gradient descent directions for the training and prevent it falling into local optima. For example, at a certain epoch, although all the training batch samples contain

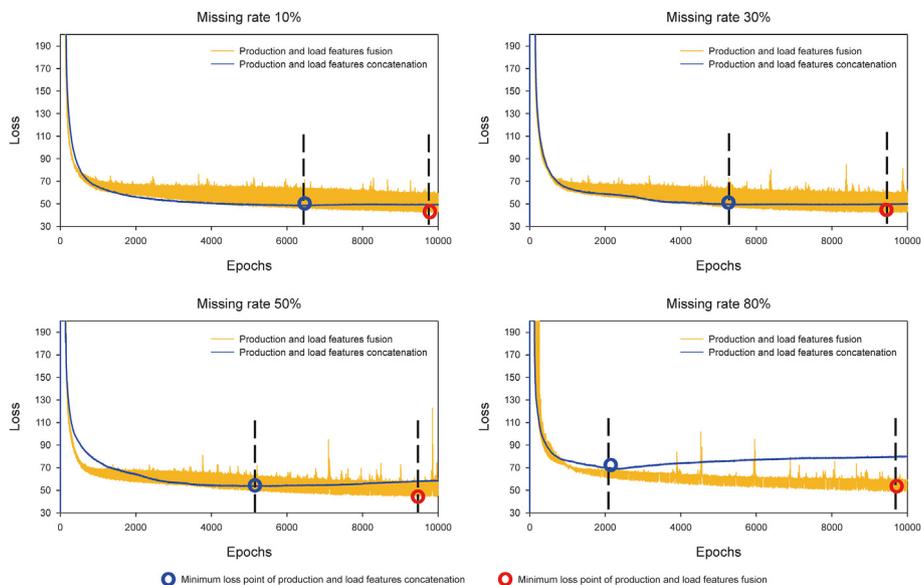


Fig. 11. Loss comparison of the validation set under different data missing rates.

**Table 8**  
MAE of the best model under different data missing rates of the validation set.

Missing rate	MAE of the best model, m	
	Production and load features fusion	Production and load features concatenation
10%	42.5	48.3
30%	42.0	49.0
50%	44.3	53.6
80%	48.4	68.7

missing data of a certain view, the model can be still trained by the other views without missing data. In this situation, the gradient descent direction is not completely wrong and the model just focuses more on the views without missing data. As the different training batch samples are alternately selected, the loss decreases in fluctuations. At last, the model that performs best on the validation set is selected for the dynamic liquid level calculation.

In addition, as the missing rate of production features increases, the loss difference of the training and validation set decreases in the features fusion (left side of Fig. 10). This phenomenon of overfitting attenuation results from the two following aspects: Firstly, the missing of production features causes the loss increase of the training and validation set. Secondly, as the missing rate increases, the model is trained less toward the lower loss direction, which results in the weakening of loss fluctuations. Therefore, as the missing rate increases, although the degree of overfitting decreases, it is more difficult for the model to get parameters with higher accuracy, and meanwhile the degree of under-fitting increases.

Under different missing rates, their validation loss can be obtained (Fig. 11). By features concatenation, the number of epochs which the model needs to reach the best performance is generally smaller than that of the features fusion. Due to the missing data, its loss appears to rise as the epochs increase further (especially when the missing rate is greater than 50%), indicating the overfitting is very obvious.

Based on the MAE of the best model under different missing rates of the validation set (Table 8), it can be seen that in the features fusion, the larger the missing rate is, the worse the validation set performs. Therefore, the influence of under-fitting due to the data missing is more significant than that of overfitting. More

noteworthy is that the MAE of the model trained by features fusion is much lower than by features concatenation. Even if the missing rate reaches 80%, it is still less than 50 m, indicating that it is less sensitive to data missing and has significant advantages when dealing with missing data.

In actual application, if the sensors fail to work during the sucker rod well production, there will be a large amount of missing data from some view in a period of time. By the multi-view features fusion, the model can still be updated in real-time, which helps improve data utilization. When testing the dynamic liquid level by the model in real-time, the mark indicating whether the view is available can be used for adapting the data missing. In this way, the risk of the real-time monitoring function failure due to the sensors faults can be reduced.

### 5. Conclusion

Based on the analysis of features, they can be divided into three views: load features of dynamometer cards, production features, and image features of dynamometer cards which can be used for calculating the dynamic liquid level.

In order to study the interaction between different features, three views were fused by multi-view neural network. The results have shown that the production features and load features are complementary for improving the calculation accuracy. The fusion performance of the model is better than a single view and features concatenation. In contrast, when they are fused with image features, the accuracy cannot be improved.

The features fusion method proposed in this paper can deal with them automatically when there are missing data. With multi-view

network, the view without missing data can be preserved completely. Therefore, there is no need to remove all the unavailable samples with missing data, and then the model performance is significantly improved in the presence of data missing.

The method proposed in this paper is fit for sucker rod well of the conventional sandstone oilfield with water driving development. In addition, the available data collected in this study is limited. In the future, it may be possible to enhance the performance of the model by incorporating acoustic wave data as features and mining temporal features based on the dynamic fluid level data with higher testing frequencies. As new data accumulates, it is necessary to update the model. Therefore, determining when to update, which data to use for updating, and how to update quickly and accurately are also important problems to be solved in future research.

### CRedit authorship contribution statement

**Cheng-Zhe Yin:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Methodology, Investigation, Conceptualization, Data curation, Formal analysis. **Kai Zhang:** Formal analysis, Conceptualization, Project administration, Funding acquisition. **Jia-Yuan Liu:** Resources, Methodology, Formal analysis. **Xin-Yan Wang:** Methodology, Formal analysis. **Min Li:** Validation. **Li-Ming Zhang:** Writing – original draft, Formal analysis. **Wen-Sheng Zhou:** Formal analysis.

### Declaration of competing interest

No potential conflict of interest was reported by the authors.

### Acknowledgements

This work is supported by the National Natural Science Foundation of China under Grant 52325402, 52274057, 52074340 and 51874335, the National Key R&D Program of China under Grant 2023YFB4104200, the Major Scientific and Technological Projects of CNOOC under Grant CCL2022RCPS0397RSN, 111 Project under Grant B08028.

### References

- Baltrusaitis, T., Ahuja, C., Morency, L.P., 2018. Multimodal machine learning: a survey and taxonomy. *IEEE T. Pattern. Anal.* 41 (2), 423–443. <https://doi.org/10.1109/TPAMI.2018.2798607>.
- Chavarriga, R., Sagha, H., Calatroni, A., et al., 2013. The opportunity challenge: a benchmark database for on-body sensor-based activity recognition. *Pattern Recogn. Lett.* 34 (15), 2033–2042.
- Chen, D.C., Zhang, R.C., Meng, H.X., et al., 2015. The study and application of dynamic liquid level calculation model based on dynamometer card of oil wells. *Sci. Technol. Eng.* 15 (32), 32–35. <https://doi.org/10.3969/j.issn.1671-1815.2015.32.006> (in Chinese).
- Chen, D.F., Han, X.L., Yang, J., 2008. Discuss on survey method for liquid level of oil well. *Well Test.* 17 (2), 60–61. <https://doi.org/10.3969/j.issn.1004-4388.2008.02.024> (in Chinese).
- Choi, J.H., Lee, J.S., 2019. EmbraceNet: a robust deep learning architecture for multimodal classification. *Inf. Fusion* 51, 259–270. <https://doi.org/10.1016/j.inffus.2019.02.010>.
- Eitel, A., Springenberg, J.T., Spinello, L., et al., 2015. Multimodal deep learning for robust RGB-D object recognition. In: Proceedings of the IEEE/RISJ International Conference on Intelligent Robots and Systems (IROS 2015), September 28–October 2, Hamburg, Germany, pp. 681–687. <https://doi.org/10.1109/IROS.2015.7353446>.
- Gibbs, S.G., 1963. Predicting the behavior of sucker-rod pumping systems. *JPT* 15 (7), 769–778. <https://doi.org/10.2118/588-PA>.
- Gu, Z.P., Lang, Bo, Yue, T.Y., et al., 2017. Learning joint multimodal representation based on multi-fusion deep neural networks. In: Proceedings of the International Conference on Neural Information Processing (ICONIP 2017), November 14–18, Guangzhou, China, pp. 276–285.
- Han, Y., Song, X.P., Li, K., et al., 2022. Hybrid modeling for submergence depth of the pumping well using stochastic configuration networks with random sampling. *J. Pet. Sci. Eng.* 208, 109423. <https://doi.org/10.1016/j.petrol.2021.109423>.
- Hou, J.L., Jiang, H.C., Wan, C.F., et al., 2022. Deep learning and data augmentation based data imputation for structural health monitoring system in multi-sensor damaged state. *Measurement* 196, 111206. <https://doi.org/10.1016/j.measurement.2022.111206>.
- Hou, Y.B., Gao, X.W., Li, X.Y., 2019. Prediction for dynamic liquid level of sucker rod pumping using generation of multi-scale state characteristics in oil field production. *CIESC J.* 70 (S2), 311–321. <https://doi.org/10.11949/0438-1157.20190352> (in Chinese).
- Jaques, N., Taylor, S., Sano, A., et al., 2017. Multimodal autoencoder: a deep learning approach to filling in missing sensor data and enabling better mood prediction. In: Proceedings of the International Conference on Affective Computing and Intelligent Interaction (ACII 2017), October 23–26, San Antonio, Texas, USA, pp. 202–208. <https://doi.org/10.1109/ACII.2017.8273601>.
- LeCun, Y., Bottou, L., Bengio, Y., et al., 1998. Gradient-based learning applied to document recognition. *Proc. IEEE* 86 (11), 2278–2324.
- Li, X.Y., Gao, X.W., Hou, Y.B., 2015. Online dynamic Gaussian process regression for dynamic liquid level soft sensing of sucker-rod pumping well. *CIESC J.* 66 (6), 2150–2158. <https://doi.org/10.11949/j.issn.0438-1157.20141791> (in Chinese).
- Li, X.Y., Gao, X.W., Li, K., et al., 2016. Ensemble soft sensor modeling for dynamic liquid level of oil well based on multi-source information feature fusion. *CIESC J.* 67 (6), 2469–2479. <https://doi.org/10.11949/j.issn.0438-1157.20151673> (in Chinese).
- Li, Z.C., Tian, L., Jiang, Q.C., et al., 2020b. Fault diagnostic method based on deep learning and multimodal feature fusion for complex industrial processes. *Ind. Eng. Chem. Res.* 59 (40), 18061–18069. <https://doi.org/10.1021/acs.iecr.0c03082>.
- Li, S., Wan, H.Q., Song, L.Y., 2020a. An adaptive data fusion strategy for fault diagnosis based on the convolutional neural network. *Measurement* 165, 108122. <https://doi.org/10.1016/j.measurement.2020.108122>.
- Liang, X., Zhang, Z.H., 2021. Research on depth of oil well moving liquid surface based on short-term energy and LSTM. *Comput. Mod.* 308, 15–19+26. <https://doi.org/10.3969/j.issn.1006-2475.2021.04.003> (in Chinese).
- Liguori, A., Markovic, R., Ferrando, M., et al., 2023. Augmenting energy time-series for data-efficient imputation of missing values. *Appl. Energy* 334, 120701. <https://doi.org/10.1016/j.apenergy.2023.120701>.
- Lu, C., 2017. A Real-time Forecasting Method of Dynamic Liquid Level Calculation Based on the Indicator Diagram of Pumping Unit Well. Master Thesis. China University of Petroleum (East China), Qingdao, China (in Chinese).
- Mohammadpoor, M., Torabi, F., 2018. Big data analytics in oil and gas industry: An emerging trend. *Petroleum* 7 (2), 241–242. <https://doi.org/10.1016/j.petlm.2018.11.001>.
- Nguyen, T., Gosine, R.G., Warriani, P., 2020. A systematic review of big data analytics for oil and gas industry 4.0. *IEEE Access* 8, 61183–61201. <https://doi.org/10.1109/ACCESS.2020.2979678>.
- Ordóñez, F.J., Roggen, D., 2016. Deep convolutional and LSTM recurrent neural networks for multimodal wearable activity recognition. *Sensors* 16 (1), 1–25. <https://doi.org/10.3390/s16010115>.
- Srivastava, N., Salakhutdinov, R., 2012. Multimodal learning with deep Boltzmann machines. In: Proceedings of the Advances in Neural Information Processing Systems (NIPS 2012), December 3–6, Lake Tahoe, Nevada, United States, pp. 2222–2230.
- Tian, Z.D., 2021. Kernel principal component analysis-based least squares support vector machine optimized by improved grey wolf optimization algorithm and application in dynamic liquid level forecasting of beam pump. *Trans. Inst. Meas. Control* 42 (6), 1135–1150. <https://doi.org/10.1177/0142331219885273>.
- Vergara, A., Fonollosa, J., Mahiques, J., et al., 2013. On the performance of gas sensor arrays in open sampling systems using Inhibitory Support Vector Machines. *Sensor. Actuator. B Chem.* 185, 462–477.
- Wang, T., Duan, Z.W., Li, K., 2017. Adaptive ensemble modeling for dynamic liquid level of oil well based on improved AdaBoost method. *J. Electron. Meas. Instrum.* 31 (8), 1342–1348. <https://doi.org/10.13382/j.jemi.2017.08.025> (in Chinese).
- Xu, X.W., Tao, Z.R., Ming, W.W., 2020. Intelligent monitoring and diagnostics using a novel integrated model based on deep learning and multi-sensor feature fusion. *Measurement* 165, 108086. <https://doi.org/10.1016/j.measurement.2020.108086>.
- Yang, L.P., 2010. The dynamic liquid level calculation of the sucker rod well by dynamometer cards. *Petrol. Geol. Eng.* 24 (5), 101–103. <https://doi.org/10.3969/j.issn.1000-7393.2011.06.030> (in Chinese).
- Yu, D.L., Qi, W.G., Ding, B., et al., 2018. Study of forecasting producing fluid level of submersible reciprocating pump on the basis of chaotic time series. *Comput. Sci. Appl.* 8 (6), 1034–1044. <https://doi.org/10.12677/CSA.2018.86115> (in Chinese).
- Yu, L., 2020. Research on the Liquid Level Prediction and the Optimization Frequency of Stroke of Submersible Plunger Pump. Master Thesis. Harbin Engineering University, Harbin, China (in Chinese).
- Zhang, H., 2003. Discussion of detecting fluid level by pressure gauge. *Well Test* 12

- (5), 49–50. <https://doi.org/10.3969/j.issn.1004-4388.2003.05.018> (in Chinese).
- Zhang, H.L., Li, P., Xie, Q., et al., 2007. Preliminary study and application of the method for calculating dynamic liquid level by dynamometer cards. *Qinghai Shiyou* 25 (2), 31–35 (in Chinese).
- Zhang, Q., Wu, X.D., 1984. Pumping well diagnostic technique and its application. *Journal of east China Petroleum Institute* 2, 145–159 (in Chinese).
- Zhang, S.L., Luo, Y., Wu, Z.M., et al., 2011. Corrected algorithm for calculating dynamic fluid level with indicator diagram for rob-pumped well. *Oil Drill. Prod. Technol.* 33 (6), 122–124. <https://doi.org/10.3969/j.issn.1000-7393.2011.06.030> (in Chinese).
- Zhou, W., Liu, J., Gan, L.Q., 2018. Dynamic liquid level detection method based on resonant frequency difference for oil wells. *Turk. J. Electr. Eng. Co.* 26 (6), 2968–2976. <https://doi.org/10.3906/elk-1805-68>.