Fault Diagnosis of Sucker Rod Pump Based on Deep-Broad Learning Using Motor Data

JINGLIANG WEI AND XIANWEN GAO
Department of Information Science and Engineering, Northeastern University, Shenyang 110819, China
Corresponding author: Xianwen Gao (gaoxianwen@mail.neu.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61573088 and Grant 61573087, and in part by the China Scholarship Council under Grant [2018] 3101.

ABSTRACT Conventional fault diagnosis methods of sucker rod pump (SRP) mainly focus the operating status of oil well by identifying the dynamometer cards (DCs), which are limited by the sensor maintenance and calibration, battery replacement and safety hazards for staff. Motor power, as the most basic parameter providing the energy source for the oil well, is directly related to the real-time operation state of oil well. Therefore, a novel deep and broad learning system (DBLS) based on motor power data for fault diagnosis of sucker rod pump is proposed in this paper. Considering the key parameters such as mechanical wear and balance weight, the motor power data are labeled by the DCs with typical working conditions. Furthermore, CNN-based feature extractor is designed to make up for the lack of expert experience in motor power, which is obtained by merging the output of the CNNs with the manual features extracted based on mechanical analysis. And then the broad learning system is employed as the classifier to solve the problem of real-time update of system structure. Finally, a dataset containing six different working states collected from the oilfield by a self-developed device is employed to verify the proposed method experimentally and compared with other methods.

INDEX TERMS Broad learning, convolutional neural network, fault diagnosis, motor power, sucker rod pump.

I. INTRODUCTION
Sucker rod pump (SRP) is the major artificial lift device employed by substantially more than 80 percent oil wells worldwide. During the oil extraction processing, many fault states such as mechanical structure damage and unstable reservoir supply may occur because of the downhole portion of the equipment often operates in poor conditions thousands of meters underground. These fault states will seriously affect the production efficiency of the SRPs and even cause production safety risks when they are not diagnosed in time. Therefore, timely and accurate diagnosis of these fault states is essential in oilfield production. Many fault diagnosis methods have been proposed in order to mitigate the aforementioned problem. Conventional fault diagnosis methods are based on dynamometer card (DC) data measured by load sensor installed on the horse-head. The DC reflects the complex working state underground [1], [2]. Relying on the field experience of diagnostic engineer and staff, Derek et al. regularized the manual rules and established the expert system to diagnosis the working state of SRP system based on DCs [3]. In order to get rid of the subjective experience of the diagnostician, many artificial intelligence based methods have been proposed such as rough set theory [4], support vector machine [5], self-organizing artificial neural network [6], fuzzy theory [7], designed component analysis [8], and hidden Markov model [9]. Zhang and Gao [10] designed a transform matrix to transfer DC data from different wells to the same subspace. Zheng and Gao [11] explored the characteristic parameters of the typical fault DCs, and employed the hidden Markov model for sucker rod pump system diagnosis. Li et al. [12] adopted an online sequential extreme learning machine (OS-ELM) to update parameters in real time, and realized continuous monitoring of downhole conditions. As shown in Fig.1, the major dynamometer card test method is the long-term fixed placement. However, the existing long-term fixed dynamometer card test method has several defects as follows.
The horse-head load needs to be removed when the equipment is installed, which is cumbersome to operate and has certain safety risks;

Long-term fixed placement test method with battery power supply, the daily maintenance will be large, and the oil production will be affected when the oil wells are shut down because of the battery replacement. Therefore, the test interval is generally long, and the downhole working condition monitoring is not timely;

If wired power supply is used, the cable is easy to be damaged along with the movement of the horse-head, and the wire walking distance is relatively long, which is inconvenient for site operation;

In addition, the dynamometer card mainly reflects downhole working conditions, which is difficult to reflect the operating state of surface equipment. The purpose of sucker rod pump is to convert electrical energy of the motor into potential energy of well fluid, so as to lift the liquid to the ground. As the energy resource of the surface mechanical system, the motor contains rich information about the lifting process. Moreover, compared with the traditional DCs, the motor power data has the advantages of real-time, convenient installation and maintenance. Therefore, in order to comprehensively and timely grasp the overall operation of the sucker rod pump, the detection and analysis of motor power is essential.

At present, the SRP fault diagnosis methods based on motor power data can be divided into direct and indirect approaches. The indirect method models the mechanism of the surface transmission part of oil well [13], and obtains the dynamometer card data converted from the motor power data. Essentially, this approach is based on dynamometer cards, the accuracy of the mechanism models directly affects the classification of the working status. Therefore, the limitation of the method occurs because of the approximation of some key parameters (such as torque factor) in the mechanism model. For the direct motor power diagnosis approach, how to extract effective and complete features is the very first step of fault diagnosis. Zheng et al. [14] defined eight features by analyzing the mechanism of motor work and data distribution of the curve, and the experimental results show the satisfaction of the diagnosis method based on the proposed features.

However, the effect of feature learning depends on whether the extracted features can suppress irrelevant changes and keep the discriminant information. Although the expert prior experience based on dynamometer card has been very sufficient in oil well diagnosis, the domain knowledge of motor power data has not yet been systematically summarized. The motor power features selected by hand-crafted analysis may not optimally characterize the motor power data.

Deep learning (DL) has been paid widely attention in fault diagnosis field due to its powerful feature extraction function endowed by its multi-layer structure. In recent years, many works achieved excellent performance merging the expert prior into deep learning process [15]–[17]. Several fault diagnosis methods based on deep belief network (DBN) [18] and convolutional neural network (CNN) [19] have shown more advanced performance compared to the conventional diagnosis methods.

However, it is worth mentioning that these deep structural neural networks are mostly training time-consuming because of the hierarchical architecture and large number of hyper parameters. Chen and Liu [20] proposed a broad learning system (BLS) with random vector functional link neural network (RVFLNN) to mitigate the difficulty of deep structure and hyper parameters optimization. Moreover, BLS employed the incremental learning algorithms to reconstruct the network efficiently when new data is added, which provides an effective solution when real-time changes occur during operating cycle.

To circumvent the difficulties aforementioned, this paper proposes a connection network based on deep and broad learning technique. The overall diagnosis framework is shown as Fig. 2. The main insights and contributions of this paper are summarized as follows.

In order to obtain the training dataset, the corresponding motor power curves under typical work conditions are obtained based on the mechanism analysis of the surface transmission system.

According to mechanism analyzing the sucker rod pump system, a fault diagnosis method is developed based on CNN and BLS, six manual features are proposed by analyzing the working position points and states of valve, and the task-specific feature extractor provides the appropriate features for the BLS diagnosis model.

The proposed DBLS is evaluated using the actual motor power data collected by a self-developed device.

The rest of this paper is organized as follows. Section II collects labeled motor power data by modeling the surface transmission system. Section III describes the proposed method. In Section IV, the proposed DBLS is evaluated using the
up and down, and then the pump is driven to extract oil. The rotational motion is turned into a reciprocating movement in the influence of the crankshaft and four-bar linkage, the motor’s surface transmission system. As shown in Fig.3, under the four-bar linkage is the main mechanical structure of the sucker rod pumping system, through which the polished rod load is transformed into the motor power; and the suspended point displacement is transformed into the crank shaft angle by modeling the four-bar linkage.

The specific angle and parameter calculations related to modeling in the four-bar linkage are defined as follows.

\[
\begin{align*}
\beta &= \arccos \frac{B^2 + L^2 - K^2 - R^2 + 2KR\cos\theta}{2BL} \\
\alpha_1 &= \arccos \frac{H}{K} \\
\varphi_{\text{bottom}} &= \arccos \frac{B^2 + K^2 - (L + R)^2}{2BL} \\
\varphi_{\text{top}} &= \arccos \frac{B^2 + K^2 - (L - R)^2}{2BL} \\
\varphi &= \varphi_{\text{bottom}} - \frac{S_0}{S_C} (\varphi_{\text{bottom}} - \varphi_{\text{top}}) \\
J &= \sqrt{B^2 + K^2 - 2BK\cos(\varphi)} \\
\theta &= \arccos \frac{J^2 + K^2 - B^2}{2BJ} + \arccos \frac{J^2 + R^2 - L^2}{2JR} \\
\theta_o &= \theta - \alpha_1
\end{align*}
\]

where \( R \) is the radius of crank, \( L \) is the length of the pitman, \( B \) is the length of backward beam, \( A \) is the length of forward beam, \( K \) is the length of the fixed rod, \( J \) is the auxiliary line connecting the beam and the crank gravity center, \( O \) and \( O_1 \) are reference points, \( H \) is the vertical height of the support point to the center of the gearbox, \( S_C \) denotes the stroke, \( S_0 \) denotes the suspension point displacement data. Set the clockwise direction positive, the motor parameters are collected at 12 o’clock of crank, \( \theta_o \) denotes the angle of crank spinning clockwise. Thus, the suspension point displacement is transformed into the crank shaft angle by modeling the four-bar linkage.

B. MOTOR POWER GENERATION

By modeling the four-bar linkage, the polished rod load is converted into motor power. The moment balance analysis of the walking beam fulcrum \( O \) and the crankshaft fulcrum \( O_1 \) are obtained as follows.

\[
F_L \sin\beta \cdot B = L \cdot A
\]

\[
T + P_L \sin\alpha \cdot R + (W_{cb} \cdot R_{cb} + W_C \cdot R) \sin\theta_o = 0
\]

where \( L \) is the polished rod load, \( F_L \) is the force of the pitman on the walking beam fulcrum \( O \), \( W_{cb} \) is the gravity of the balance block of the crank, \( W_C \) is the gravity of the crank, and \( R_{cb} \) is the radius of the balance block.

Considering the transmission efficiency and unbalanced weight of the four-bar linkage, and the motor torque \( T \) is obtained as follows.

\[
T = \frac{\overline{T_F}(P - B) \cdot \eta}{\sin\beta} - (W_{cb} \cdot R_{cb} + W_C \cdot R) \sin\theta_o
\]

\[
\overline{T_F} = \frac{A}{B} \cdot \frac{\sin\alpha}{\sin\beta}
\]

\[
\sigma = \begin{cases} 
-1, & \theta_o \in [0, \pi] \\
1, & \theta_o \in [\pi, 2\pi]
\end{cases}
\]

where \( \overline{T_F} \) is the torque factor, \( \eta \) is the transmission efficiency, and \( t \) is the times of stroke in a minute. Based on the above actual motor power data collected by a self-developed device. Finally, conclusions are displayed in Section V.
The equation, the motor power $P$ is obtained as follows:

$$P = T \cdot \frac{t}{9545} \cdot \eta^\tau$$

(7)

$$\tau = \begin{cases} -1, & T_n \leq 0 \\ 1, & T_n \geq 0 \end{cases}$$

(8)

Through the establishment of the surface transmission system model, the DC data with polished rod load $L$ and displacement $S_0$ is successfully transformed into the motor power data with motor power $P$ and the crankshaft angle $\theta_o$. Thus, the generated data set is adopted to train the fault diagnosis model. Notably, different from [13], the torque factor, as one of the important parameters for surface transmission model, does not need to be compensated near zero in our method.

### III. PROPOSED METHOD

The overview of the proposed fault diagnosis system is illustrated in Fig. 4, which consists of task-specific feature extractor based on CNN and fault diagnosis model based on BLS.

#### A. CNN-BASED FEATURE EXTRACTOR MERGING PRIOR KNOWLEDGE

Inspired from the work of LeCun et al., convolutional neural network (CNN), as one of the most effective DLs, has been vastly employed to extract the features of one-dimension (1-D) [21] signal in electrocardiogram (ECG) classification [22], motor fault detection [23] and real-time structural health monitoring [24] fault diagnosis recently. In this paper, we design a novel task-specific feature extractor based on a 1-D CNN merging mechanism analysis prior knowledge, which consists of several convolution layers, pooling layers, two fully connected layers, and one feature concatenating layer.

**Data Pre-processing layer**: Due to the different mechanical parameters and work interval of motors in different oil well, the raw motor power data cannot be used for feature extraction directly. To reduce the interference from noise and to extract the features accurately, the normalization and filtering of the motor power data $(\hat{\theta}, \hat{P})$ is operated as follows.

$$\begin{align*}
\theta &= \frac{\hat{\theta} - \hat{\theta}_{min}}{\hat{\theta}_{max} - \hat{\theta}_{min}} \\
\hat{P} &= \frac{\hat{P} - \hat{P}_{min}}{\hat{P}_{max} - \hat{P}_{min}}
\end{align*}$$

(9)

$$P(t) = \frac{1}{5} \sum_{t-2}^{t+2} \hat{P}(t)$$

(10)

where $\hat{\theta}$ and $\hat{P}$ are the crank shaft angle and motor power, $\hat{\theta}_{min}$, $\hat{\theta}_{max}$, $\hat{P}_{min}$, $\hat{P}_{max}$ are the minimal and maximal crank shaft angle and motor power respectively. The size of filter windows is set to be 5.

**Convolutional layer**: In the convolutional layer, the 1-D convolution kernels are slide with a window size of $m$ over the whole input data with length of $L$ to extract the feature map. Concretely, the $j$th node output $z_j$ of the convolutional layer is defined as follows.

$$z_j = \text{ReLU}(\sum_{i=1}^{n} (w_i)^T P_{j+i+m-1} + b_j)$$

(11)

where $w_i \in \mathbb{R}^m$ represents the $i$th convolution kernel, and $b_j$ denotes the corresponding bias, $P_{j+i+m-1}$ is the $j$th segmented signal of input data with length of $m$. ReLU(·) is an activation function.

**Pooling layer**: Each convolutional layer is connected with a pooling layer, which plays a role of dimensionality reduction. The max pooling $k_j$ is used in this paper to extract the most important and shift invariant feature as follows.

$$k_j = \max_{(j-1)p+1 \leq i \leq jp} \{z_i\}$$

(12)

where $p$ is the pooling length, and $k_j$ is the pooling output of the $j$th point.
Fully connected layer: After several convolutional layers and pooling layers, the mapped features are flattened and then enter the fully connected layer. The output of the fully connected layer is calculated as follows.

\[ f = \sigma(w^T z + b) \]  

where \( \sigma(\cdot) \) is the activation function, \( z \) is the input flattened features, \( w \) is the weight of the fully connected layer, and \( b \) is the corresponding bias.

Feature fusion layer: The feature fusion layer is a novel layer proposed in this paper for solving the issue of incomplete feature of the motor power data. The function of the feature fusion layer is to learn the merging feature of the CNN-based and hand-crafted. In this layer, several meaningful mechanism features denoted by \( z' \) based on expert advice are selected to label the data, the output the feature fusion layer is represented as follows.

\[ Z = \text{ReLU}(w^T z + w'^T z' + b) \]  

where \( \text{ReLU}(\cdot) \) is an activation function, \( z \) is the flattened features outputted from the CNN, \( w \) is the weight of the fully connected layer, \( w' \) is the weight of the mechanism features, and \( b \) is the corresponding bias. Hence, the combination of the selected mechanism features is used as input to learn the novel complete features along with the features extracted from the CNN.

In order to extract the motor power curve features with mechanism meanings, we need to discover the changing rule of the motor power curve by corresponding to the dynamometer card. Similar to the dynamometer card, the power curve is divided into four stages with the change of crank angle. As shown in Fig.5, two crest values of the motor power curve appear at the 90° and 270° of crank, corresponding to loading point B and unloading point D. Therefore, the four valve working points on the theoretical dynamometer card are matched one by one in the power curve.

In this paper, we employ the following six variables as the distinguished (relatively independent) features of a motor power curve.

**Skewness variation:** The skewness measures the deviation degree of data distribution. Compared to the theoretical motor power curve, the skewness variation between the up-stroke and down-stroke is defined as follows.

\[ \text{Skew} = \frac{1}{N_u} \sum_{i=1}^{N_u} \left( \frac{P_u(i) - \mu_u}{\sigma_u} \right)^3 - \frac{1}{N_d} \sum_{i=1}^{N_d} \left( \frac{P_d(i) - \mu_d}{\sigma_d} \right)^3 \]  

where \( N_u \) is the number of the up-stroke samples, \( N_d \) is the number of the down-stroke samples, \( P_u(i) \) and \( P_d(i) \) are the ith sample in up-stroke and down-stroke, \( \mu_u \) and \( \mu_d \) are the mean value of the up-stroke and down-stroke samples, \( \sigma_u \) and \( \sigma_d \) are the standard deviation of the up-stroke and down-stroke samples respectively. The early (delayed) loading (unloading) of the polished rod load in one stroke can be presented in the form of skewness change in the motor power curve. Thus, it will be a powerful feature when the fault conditions such as Insufficient liquid supply and valve leakage occur.

**Kurtosis difference:** The fullness degree of the two crests in the up and down stroke of the motor power curve directly describes the operation of the oil well, so we adopt the kurtosis difference as another major feature, which is defined as follows.

\[ \text{kurtosis} = \frac{1}{N_u} \sum_{i=1}^{N_u} \left( \frac{P_u(i) - \mu_u}{\sigma_u} \right)^4 - \frac{1}{N_d} \sum_{i=1}^{N_d} \left( \frac{P_d(i) - \mu_d}{\sigma_d} \right)^4 \]  

Kurtosis difference describes the operation state of the SRPs in the loading and unloading stages.

The ratio of work done in the up and down stroke describes the global characteristic oil well operation state.

**The ratio of the up-stroke work:**

\[ R_u = \frac{\int_0^{N_u} P_u}{\int_0^{N_u+N_d} P} \]  

**The ratio of the down-stroke work:**

\[ R_d = \frac{\int_0^{N_u+N_d} P_d}{\int_0^{N_u+N_d} P} \]  

Work efficiency in one stroke: As an indispensable criterion for judging the motor and SRP system working statue, the work efficiency is defined as the ratio between average motor power and rater motor power.

\[ R = \frac{\sum_{i=1}^{N_u+N_d} P_i}{(N_u+N_d) \times P_r} \]  

where \( P_r \) is the rated power of the motor.
Unbalance coefficient: The balance weight is the general premise to judge the working condition of SRP system. In this paper, the peak power relation is adopted to describe the unbalanced coefficient.

\[
e = \frac{[\text{max}(P_i) - \text{max}(P_j)]}{\text{max}(P_u, P_d)}
\]

(20)

All the above six features need to be extracted on the premise of clear valve working point. In view of the principle of dynamometer card, we divided the motor power curve into 5 regions with the power mean as the limit, and the point with the biggest curvature change in the area including crests is the valve working point.

Since a single feature cannot fully express one working condition, they are divided into four groups according to their physical meanings, among which every two groups are integrated with the features extracted from CNN. The skewness and kurtosis, the ratio of up-stroke and down-stroke work are grouped respectively, while work efficiency and unbalance coefficient are in each own group apiece. Therefore, 6 CNNs with the same structure need to be trained, and the output of the feature extractor is defined as follows.

\[
Z^n = [Z_1, Z_2, \ldots, Z_n] \in R^{N \times S}
\]

(21)

where \(Z_i (i = 1, 2, \ldots, n)\) denotes the mapped feature output from the \(i\)th CNN-based feature extractor, \(N\) is the number of the samples and \(S\) is the total number of neurons in feature concatenating layer from \(n\) CNNs.

B. SUCKER ROD PUMP FAULT DIAGNOSIS BASED ON BLS

The BLS proposed by C.L.P. Chen is built in flat network based on the random vector functional link neural network (RVFLNN), of which the mapped features and enhancement nodes are directly connected to the output. The flatten structure enables the output coefficients can be obtained by the pseudo-inverse matrix. Furthermore, the BLS extend the network structure using fast incremental learning without the retraining the full network. With the advantages of fast training and low rank approximation, the BLS has been widely used in pipeline leak detection [25], event camera [26] and fast PolSAR image classification [27] in last several years.

BLS architecture: The outline of the entire BLS structure is illustrated in Fig. 4. The mapped features \(Z^n\) are sent to the input layer of the BLS system. The weights \(W^j\) and the bias \(\beta^j\) between the feature nodes and enhancement nodes are randomly generated and the corresponding enhancement nodes \(H_j\) is obtained by \(\xi_j(Z^nW^j + \beta^j)\), where \(\xi_j()\) is a nonlinear function. \(j = 1, 2, \ldots, m\), where \(m\) is the number of groups of enhancement nodes. Consequently, feature nodes and enhancement nodes are consolidated to the output layer, the BLS could be modeled as follows.

\[
Y = [Z_1, Z_2, \ldots, Z_nH_1, H_2, \ldots, H_m]W^m
\]

\[
= [Z^nH^m]W^m
\]

(22)

where \(H^m = [H_1, H_2, \ldots, H_m]\), \(W^m\) is the output weight of BLS. Define \(Q^n_m = [Z^nH^m]\) as the pattern matrix, then the output weight could be derived by least square method. In the broad learning system, we use the ridge regression to solve it.

\[
W^m_n = \left(\lambda I + Q^n_m(Q^m_m)^T\right)^{-1}(Q^m_m)^TY
\]

(23)

Incremental learning: In actual oil production, the diagnosis capability of origin BLS system may not match the real-time update of sucker rod pumping system. The major reasons for this are the insufficiency of features or enhancement nodes. The general practice in most deep structure network is either to increase the number of the windows or to increase the number of layers, which requires resetting parameters for each new adding windows or layers. This often involves a more tedious and time-consuming adaptive learning process. Instead, in the BLS model, the whole architecture can be constructed to operate two kinds of update (feature and enhancement nodes) without the need to reset and retrain the entire network.

First, in order to improve the diagnosis function of the BLS system, the addition enhancement nodes \(Q^n_{add}\) are defined as follows.

\[
Q^n_{add} = \xi_{add}(Z^nW^h_{add} + \beta^h_{add})
\]

(24)

where \(W^h_{add}\) and \(\beta^h_{add}\) are the weights and bias for addition nodes.

Second, when unfamiliar faults occur, system has insufficient features to detect them. To address this problem, the new features \(Z_{add}\) are added to update the enhancement nodes as follows.

\[
H_{ex_m} = [\xi_1(Z^n_{add}W^h_{ex1} + \beta^h_{ex1}), \xi_2(Z^n_{add}W^h_{ex2} + \beta^h_{ex2}), \ldots, \xi_m(Z^n_{add}W^h_{exm} + \beta^h_{exm})]
\]

(25)

where \(W^h_{ex}\) and \(\beta^h_{ex}\) are randomly generated, and \(Q^m_{add} = [Z_{add}, H_{ex_m}]\). And the addition enhancement nodes are obtained as follows.

\[
Q_A = \begin{cases} 
Q^n_{add} & \text{add enhancement nodes} \\
Q^m_{add} & \text{add features}
\end{cases}
\]

(26)

Thus, we obtained the new pattern matrix \(\hat{Q} = [Q^n, Q_A]\), and the output weight \(W\) are obtained as follows

\[
\hat{W} = \left(\lambda I + \hat{Q}(\hat{Q})^T\right)^{-1}(\hat{Q})^TY
\]

\[
= (\hat{q})^{-1}(Q^n_m, Q_A)^TY
\]

(27)

where \(q = Q_A(Q_A)^T\). \(q = \lambda I + Q^m_n(Q^m_n)^T\). And then,

\[
[Q^n_m, Q_A]^T = \left[(qW), Q_A\right]^T
\]

\[
= \left[(\hat{q} - Q_A(Q_A)^T)W\right]
\]

(28)
Therefore, the output weight is updated as follows.

\[
\hat{W} = (\hat{q})^{-1}[Q_A Q_A^T]^T Y \\
\quad = \left[ (1 - (\hat{q})^{-1} Q_A Q_A^T) W \right]^T Y
\]

where \((\hat{q})^{-1} = q^{-1} - q^{-1} Q_A Q_A^T q^{-1} Q_A + I q^{-1} Q_A^T q^{-1}.\)

**IV. EXPERIMENTAL RESULTS AND DISCUSSION**

**A. DATA COLLECTION**

the real-time motor power data of oil production site is collected by a device called NEU Multi-function Monitor, which was developed by the research team in Northeastern University, China. In order to meet the requisition of the real-time parameters’ measurement, the monitor selects ATT7022B as the core chip and obtains the active power of the motor by continuous calculation of power factor. Outside of the device, the proximity switch is used to measure the crank trajectory and set the working period. Besides, this device can carry out online monitoring of important parameters such as dynamic fluid level, pump efficiency, degree of balance and cumulative power consumption, and make real-time fault diagnosis of the working state for the SRPs. The installation picture is shown as Fig. 6, this device is embedded in the cabinet (as shown in Fig. 1) to get the working currents and voltages from the motor, and then a liquid crystal display and membrane keyboards are facilitated to data display and data query.

**FIGURE 6.** The installation of the monitor device.

To create a labeled database, the motor power data should be labeled by the dynamometer cards (DCs) using the mechanism model in Section II. As shown in Fig. 7, the blue line indicates the motor power obtained by the mechanism model in Section II and the red line is the actual motor power data. The overall motor power curves of the six working conditions are shown as Fig.8. Due to the influence of noise generated by the reducer and pulley, the actual motor power data fluctuates more violently, which have little impact on our labeling fault data set.

**FIGURE 7.** The collected motor power and the inferred motor power.

**FIGURE 8.** Six working conditions.

Based on the long-term field industrial experiment of the device, the motor power data in six typical working conditions were collected and sorted into a database to verify the proposed algorithm, including six working conditions naming Normal, Insufficient liquid supply, Gas affected, Gas locked, Standing valve leakage, and Parting rod. The database contains 24000 sets of labeled data, with an average of 4000 sets for each working condition. Among them, 3200 sets of data are taken as the training sets and remaining 800 sets as the test sets.

**B. PARAMETER SETTINGS**

All of the CNNs are designed with the same structure, which consists of two alternating Conv layers, pooling layers and two fully connecting layers. The first Conv has \(5 \times 1\) convolution kernels with number of 12, which is denoted as \(5 \times 1 \times 12\). The second Conv is \(5 \times 1 \times 8\). The pooling layers use maxpooling with \(2 \times 1\) filters. Dropout rate for the second fully connecting layer is set to 0.3. The Adam algorithm with batch size of 100 is employed as the optimizer. The number of enhancement nodes is set to 11000.

**C. FEATURES COMPLETENESS ANALYSIS**

Since the proposed method is characterized by incremental learning, the completeness of our feature selection is tested by increasing the number of features. In this subsection, features
TABLE 1. Feature reconstruction.

| Combination | Number | Features     |
|-------------|--------|--------------|
| $G_1$       | $\rho^3$ | Skewness     |
|             | $\rho^4$ | Kurtosis     |
| $G_2$       | $\rho^5$ | Ratio of up-stroke |
|             | $\rho^6$ | Ratio of down-stroke |
| $G_3$       | $\rho^7$ | Work efficiency |
| $G_4$       | $\rho^8$ | Unbalance coefficient |
| $G_5$       | $\rho^9$ | Distance $L_a$ |
|             | $\rho^9$ | Distance $L_b$ |
|             | $\rho^9$ | Distance $L_c$ |

are extracted and reconstructed in certain combination with manual experience, which is listed in Table 1. The last three features presented in [28] are adopted for incremental learning in this paper. Similar with aforementioned in Section III, features $\rho^7, \rho^8$ and $\rho^9$ are packaged as one group. Then the five groups are randomly recombined, and the number of groups included in the new combination increases from 3 to 5. Therefore, 16 experiments are conducted totally.

In order to reduce the chanciness of the diagnosis results, ten times tenfold cross validation have been done and the average value is used to estimate the accuracy of the proposed method. To facilitate the display of experimental results, six working conditions are marked with number from 1 to 6. The results are shown as histogram in Fig. 8. Obviously, for the existing 6 working conditions, the test accuracy is significantly improved with the increase of manual features, and the highest accuracy was found at group of $(G_1, G_2, G_3, G_4)$. The other increased combinations cannot improve the diagnostic accuracy. This proves the proposed feature combination $(G_1, G_2, G_3, G_4)$ has the characteristic of completeness.

D. PREDICTED RESULTS OF PROPOSED METHOD

After the feature combination is determined, a confusion matrix is carried out to demonstrate the diagnosis results in Table 2. Table 2 shows that Gas locked and Parting rod achieve 100% accuracy, and even the worst Normal obtains the accuracy of 93.5%. However, the misclassification of Normal, Insufficient liquid supply, Gas affected and Standing valve leakage have reached 1.75%, 4.62%, 5.50% and 6.5%. Almost 1.25% of Normal are misclassified to Gas affected, 0.50% of Normal are misclassified to Standing valve leakage, 4.62% of Insufficient liquid supply are misclassified to Gas affected. Almost 2.87% of Gas affected are misclassified to Normal, 2.63% of Gas affected are misclassified to Insufficient liquid supply, 6.5% of Standing valve leakage are misclassified to Normal. In the process of actual oil production, the oil well must be shut down for maintenance immediately either fault Gas locked or Parting rod occurs. For these two fault types, the proposed method gives a completely accurate judgement. On the other hand, the change process of the other four working conditions is relatively slow. When one condition is transformed into another, the overlap phenomenon will occur because of the oil well load and other parameters slight changes, thus it causes the misclassification. These conditions (Insufficient liquid supply, Gas affected and Standing valve leakage) caused by slow processes will not immediately damage the operation of SRP and are acceptable within a certain threshold, so they need to be monitored in real-time.

E. COMPARISON RESULTS WITH OTHER METHODS

The purpose of this subsection is to compare the proposed DBLS with other classical methods. Considering the motor power curve as a type of 1-D time series signal, two conventional feature analysis methods (FFT and Wavelet transformation), pure CNN feature extractor and pure manual feature extractor are used together with the merging feature extractor proposed in this paper for comparison. All the four

![Comparison results using different manual features.](image-url)
comparison methods adopt BLS as classifier. The wavelet basis is Daubechies wavelet and four-layer wavelet decomposition is adopted. The comparison results for six working conditions is illustrated in Table 3. It shows that FFT, Wavelet and pure CNN perform poorly in some working conditions due to a lack of expert experience guidance. These analysis methods lose the mechanism information of the sucker rod pump system, and it is difficult to extract the clear fault features in the high frequency of the motor power in the case of high noise. And the pure manual features cannot extract more abstract features, resulting in the low accuracy of some working conditions. The feature extraction method proposed in this paper combines the advantages of the above methods, and shows the outstanding performance. Consequently, the accuracy of proposed DBLS is compared with the mainstream method in SRP fault diagnosis field, including ELM, SVM and HMM using the same input by CNN-based feature extractor proposed in this paper. The results of them are listed in Table 4, which shows that the proposed DBLS outperforms other three diagnosis methods.

The above experimental results illustrate the superiority of proposed DBLS method on SRPs fault diagnosis from three aspects. Firstly, the proposed merging feature extractor based on CNN extracts the abstract features of the data and then use manual features as the guidance, while absorbing the advantages of the two feature extraction methods, as proved in Table 3. Secondly, the structure of proposed DBLS framework provides constructive suggestions when the model accuracy needs to be improved. Moreover, when new features (new working conditions) appear and need to be added to the SRP system for adjustment, the DBLS framework adopts the incremental learning method, which avoids the retraining of the entire model and saves the model training time.

V. CONCLUSION

This paper aims to solving three problems in Sucker Rod Pumps (SRPs) diagnosis method. Firstly, we have labeled the motor power data with the dynamometer card data by analyzing the mechanism model. And a novel method has proposed to diagnose SRPs working conditions using motor power instead of dynamometer card. Secondly, under the circumstance of insufficient expert experience, the CNN-based feature extractor merging expert experience has shown its superior abilities, which can be illustrated in the experiment results. In addition, the incremental learning endows the diagnosis model strong ability to update the structure. Finally, the data supporting the experiments are collected from a self-developed device. And the results comparison with the other methods shows the effectiveness and superiority of the proposed method.

REFERENCES

[1] B. M. Wilamowski and O. Kaynak, “Oil well diagnosis by sensing terminal characteristics of the induction motor,” IEEE Trans. Ind. Electron., vol. 47, no. 5, pp. 1100–1107, Oct. 2000.
[2] J. F. Keating, R. E. Laine, and J. W. Jennings, “Society of petroleum engineers SPE production operations symposium—Oklahoma City, Oklahoma (1991-04-07) SPE production operations symposium—Application of a pattern-matching expert system to sucker-rod, dynamometer-card pattern recognition,” Tech. Rep.
[3] H. J. Derek, J. W. Jennings, and S. M. Morgan, “Society of petroleum engineers permian basin oil and gas recovery conference—[1988.03.10-1988.03.11] proceedings of permian basin oil and gas recovery conference—Sucker rod pumping unit diagnostics using an expert system,” in Proc. Permian Basin Oil Gas Recovery Conf., 1988.
[4] J. Wang and Z. Bao, “Study on pump fault diagnosis based on rough sets theory,” in Proc. 3rd Int. Conf. Innov. Comput. Inf. Control, 2008, p. 288.
[5] K. Li, X. Gao, Z. Tian, and Z. Qiu, “Using the curve moment and the PSO-SVM method to diagnose downhole conditions of a sucker rod pumping unit,” Petroleum Sci., vol. 10, no. 1, pp. 73–80, Mar. 2013.
[6] P. Xu, S. Xu, and H. Yin, “Application of self-organizing competitive neural network in fault diagnosis of sucker rod pumping system,” J. Petroleum Sci. Eng., vol. 58, nos. 1–2, pp. 43–48, Aug. 2007.
[7] K. Li, X.-W. Gao, H.-B. Zhou, and Z.-D. Tian, “Fault diagnosis for down-hole conditions in beam pumping units based on an improved fuzzy iterative self-organizing data analysis technique,” in Proc. 10th Int. Conf. Fuzzy Syst. Knowl. Discovery (FSKD), Jul. 2013, pp. 279–284.
[8] K. Li, X.-W. Gao, W.-B. Yang, Y.-L. Dai, and Z.-D. Tian, “Multiple fault diagnosis of down-hole conditions of sucker-rod pumping wells based on freeman chain code and DCA,” Petroleum Sci., vol. 10, no. 3, pp. 347–360, Sep. 2013.
[9] B. Zheng and X. Gao, “Sucker rod pumping diagnosis using valve working position and parameter optimal continuous hidden Markov model,” J. Process Control, vol. 59, pp. 1–12, Nov. 2017.
[10] A. Zhang and X. Gao, “Supervised dictionary-based transfer subspace learning and applications for fault diagnosis of sucker rod pumping systems,” Neurocomputing, vol. 338, pp. 293–306, Apr. 2019.
[11] B. Zheng and X. Gao, “Diagnosis of sucker rod pumping based on dynamometer card decomposition and hidden Markov model,” Trans. Inst. Meas. Control, vol. 40, no. 16, pp. 4309–4320, Dec. 2018.
[12] K. Li, Y. Han, and T. Wang, “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. Petroleum Sci. Eng., vol. 160, pp. 285–301, Jan. 2018.

J. Wei, X. Gao: Fault Diagnosis of SRP Based on Deep-Broad Learning Using Motor Data
J. Wei, X. Gao: Fault Diagnosis of SRP Based on Deep-Broad Learning Using Motor Data

[13] J. Wei and X. Gao, “Electric-parameter-based inversion of dynamometer card using hybrid modeling for beam pumping system,” Math. Problems Eng., vol. 2018, pp. 1–12, Jun. 2018.

[14] B. Zheng, X. Gao, and X. Li, “Fault detection for sucker rod pump based on motor power,” Control Eng. Pract., vol. 86, pp. 37–47, May 2019.

[15] T. Han, C. Liu, L. Wu, S. Sarkar, and D. Jiang, “An adaptive spatiotemporal feature learning approach for fault diagnosis in complex systems,” Mech. Syst. Signal Process., vol. 117, pp. 170–187, Feb. 2019. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0888327018304503, doi: 10.1016/j.ymssp.2018.07.048.

[16] T. Han, C. Liu, W. Yang, and D. Jiang, “A novel adversarial learning framework in deep convolutional neural network for intelligent diagnosis of mechanical faults,” Knowl.-Based Syst., vol. 165, pp. 474–487, Feb. 2019. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0950705118306142, doi: 10.1016/j.knosys.2018.12.019.

[17] T. Han, C. Liu, W. Yang, and D. Jiang, “Deep transfer network with joint distribution adaptation: A new intelligent fault diagnosis framework for industry application,” ISA Trans., vol. 97, pp. 269–281, Feb. 2020. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0019057819303489, doi: 10.1016/j.isatra.2019.08.012.

[18] H. Shao, H. Jiang, X. Zhang, and M. Niu, “Rolling bearing fault diagnosis using an optimization deep belief network,” Meas. Sci. Technol., vol. 26, no. 11, Nov. 2015, Art. no. 115002.

[19] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” Nature, vol. 521, no. 7553, p. 436, 2015.

[20] C. L. P. Chen and Z. Liu, “Broad learning system: An effective and efficient incremental learning system without the need for deep architecture,” IEEE Trans. Neural Netw. Learn. Syst., vol. 29, no. 1, pp. 10–24, Jan. 2018.

[21] L. Wen, X. Li, L. Gao, and Y. Zhang, “A new convolutional neural network-based data-driven fault diagnosis method,” IEEE Trans. Ind. Electron., vol. 65, no. 7, pp. 5990–5998, Jul. 2018.

[22] S. Kiranyaz, T. Ince, and M. Gabbouj, “Real-time patient-specific ECG classification by 1-D convolutional neural networks,” IEEE Trans. Biomed. Eng., vol. 63, no. 3, pp. 664–675, Mar. 2016.

[23] T. Ince, S. Kiranyaz, L. Eren, M. Askar, and M. Gabbouj, “Real-time motor fault detection by 1-D convolutional neural networks,” IEEE Trans. Ind. Electron., vol. 63, no. 11, pp. 7067–7075, Nov. 2016.

[24] O. Abdeljaber, O. Avci, S. Kiranyaz, M. Gabbouj, and D. J. Inman, “Real-time vibration-based structural damage detection using one-dimensional convolutional neural networks,” J. Sound Vib., vol. 388, pp. 154–170, Feb. 2017.

[25] D. Ma, J. Wang, Q. Sun, and X. Hu, “A novel broad learning system based leakage detection and universal localization method for pipeline networks,” IEEE Access, vol. 7, pp. 42345–42353, 2019.

[26] S. Gao, G. Guo, H. Huang, X. Cheng, and C. L. P. Chen, “An end-to-end broad learning system for event-based object classification,” IEEE Access, vol. 8, pp. 45974–45984, 2020.

[27] J. Fan, X. Wang, X. Wang, J. Zhao, and X. Liu, “Incremental wishart broad learning system for fast polsar image classification,” IEEE Geosci. Remote Sens. Lett., vol. 16, no. 12, pp. 1854–1858, Dec. 2019.

[28] B. Zheng, X. Gao, and R. Pan, “Sucker rod pump working state diagnosis using motor data and hidden conditional random fields,” IEEE Trans. Ind. Electron., vol. 67, no. 9, pp. 7919–7928, Sep. 2020.

JINGLIANG WEI received the B.S. degree from the Department of Information Science and Engineering, Northeastern University, Shenyang, China, in 2011, and the M.S. degree in control theory and control engineering from the Shenyang University of Technology, Shenyang, in 2014. He is currently pursuing the Ph.D. degree in control theory and control engineering with the School of Information Science and Engineering, Northeastern University. His research interests include diagnosis, soft sensor, prognostics and health management, and modeling of complex industry process.

XIANWEN GAO received the B.S. degree in automation from the Shenyang University of Chemical Technology, Shenyang, China, in 1978, and the M.S. and Ph.D. degrees in control theory and control engineering from Northeastern University, Shenyang, in 1993 and 1998, respectively. He is currently a Professor with Northeastern University. His research interests include modeling of complex industry process and intelligent control, sliding-mode control, and stochastic jump systems.