ISSN 2663-1563 Vol. 7, Issue 1: 17-22, DOI: 10.25236/AJAGE.2025.070103

Research on Intelligent Diagnosis Technology of Oil Well Condition Based on Big Data and Deep Learning Model OWDNet

Yuxin Wang

2970 International Dr. APT 109C, Ypsilanti, Michigan 48197, USA
yuxinece@gmail.com

Abstract: *Accurate diagnosis of oil well operating conditions is the key to achieve safe and efficient production and enhanced oil recovery. With the promotion of oilfield information construction, real-time collection of production dynamic monitoring data has covered a wide range, but how to effectively tap the potential value of massive data is still a difficult problem to be solved. Combined with the advanced advantages of big data technology and deep learning model, intelligent diagnosis technology is showing the potential to break through the limitations of traditional condition diagnosis methods. Based on more than 40 million sets of historical monitoring data covering multiple types of oil reservoirs, this study constructed a high-quality diagnostic sample set containing 5 classes of 37 working conditions, and designed a special convolutional neural network model OWDNet according to the characteristics of oil well working conditions. The model has a 26-layer network structure and more than 59 million learnable parameters, and after optimized training, it achieves an excellent performance of 99.7% training accuracy and 98.9% verification accuracy. In the actual oilfield application, the intelligent diagnosis system developed based on OWDNet has completed more than 5 million working conditions identification, and the field diagnosis accuracy rate has reached 90%, and the real-time alarm push has been significantly improved. As a result of the continuous operation of the system, well production management has been refined, and the proportion of continuously stable production Wells has increased from 68% to 88%. The research shows that the deep integration of deep learning and big data not only accelerates the intelligent transformation of oil fields, but also provides strong support for the construction of a more efficient and reliable oil field production management system, and sets a new benchmark for industry technology upgrading and innovative application.*

Keywords: *Oil well condition; Intelligent diagnosis; Big data analysis; Deep learning; Convolutional neural network; OWDNet model; Field production optimization*

1. Introduction

China has more than 100,000 oil Wells, more than 90 percent of which rely on rod pumping systems, which use surface pumping units to drive pumps kilometers below the surface to lift underground crude oil to the surface. This kind of system is prone to corrosion, wax formation, sand and other problems in the long run, which leads to abnormal production. Abnormal working conditions may not only cause significant output loss, but also cause serious safety hazards and environmental pollution events. It is of great significance to grasp the working conditions of oil well in real time and accurately to ensure the safety and efficiency of oil field production and to maximize the efficiency of oil production.

As a key technology in oil exploitation management, well condition diagnosis always faces great challenges. Because oil Wells are widely distributed, key equipment is buried deep underground, and the relationship between monitoring indicators and actual working conditions is complex and difficult to model, traditional diagnostic methods are difficult to deal with diversified and complicated working conditions. Traditional technology relies on manual experience analysis or single index judgment, and mostly focuses on the identification of common faults, which is difficult to meet the needs of multi-dimensional and multi-working condition accurate diagnosis, restricting the wide application in actual production.

With the continuous development of information technology, oilfield production systems are

ISSN 2663-1563 Vol. 7, Issue 1: 17-22, DOI: 10.25236/AJAGE.2025.070103

gradually equipped with a large number of sensors, which can monitor multi-dimensional data such as downhole temperature, pressure and current in real time, and form massive dynamic monitoring big data. This data provides a solid foundation for working condition diagnosis, but also puts forward higher requirements for data mining and utilization. Big data technology and deep learning algorithms have achieved breakthrough applications in many fields, such as image recognition, speech processing, automatic driving, etc., providing a new way to solve complex system problems. The oil industry has gradually realized the importance of intelligent transformation, and major enterprises have developed strategic plans based on big data and artificial intelligence to deal with complex problems in the production process and promote technological innovation in the industry.

In order to deal with the technical bottleneck of current working condition diagnosis, this study developed a set of intelligent working condition diagnosis system based on deep learning and big data based on the results of oilfield information construction. By integrating over 40 million sets of historical dynamic monitoring data, a high-quality diagnostic sample set containing 37 operating conditions of 5 categories was constructed, and a convolutional neural network model (OWDNet) was designed and optimized for the diagnosis of well operating conditions. Based on the efficient feature extraction ability and strong generalization performance, the model combines with the intelligent diagnosis system with continuous learning ability to realize the precision, automation and real-time diagnosis of working conditions. The field application results show that the system not only significantly improves the efficiency of oil well production management, but also provides a feasible path for the oil industry to realize technological change and production optimization under the background of intelligent upgrading.

2. Related research

In this paper, YWu's team proposed a method based on convolutional neural network to intelligently identify the degree of fluid insufficiency in oil Wells[1]. JB Wosowei, C Shastry described an efficient IoT-based monitoring system that uses ESP32 microcontrollers and sensors to monitor well pressure, temperature, level and flow in real time[2]. According to the wellbore and production characteristics of shale oil horizontal Wells, ZK Deng's team selected ESPCP as the main lifting technology[3]. Z Huang's team proposed a multivariate time series transformation method for time series data of oil well sensors. The time series data can be converted into image texture features to better distinguish the operating state of the oil well pump[4].

3. Deep learning model design and OWDNet architecture for well condition diagnosis

The purpose of well condition diagnosis is to evaluate the current running state of the well accurately by analyzing the indicator diagram of the well and using computer technology for automatic classification recognition, which is essentially a pattern recognition problem. Traditional pattern recognition methods perform well when dealing with small scale data, but with the explosion of oil field data and the diversification of working conditions, it is difficult to meet the practical application needs by relying solely on classical machine learning algorithms. Deep learning, especially convolutional neural networks, has shown excellent ability in processing large-scale and high-dimensional data, and has become a mainstream method to solve complex pattern recognition problems.

Compared with traditional machine learning methods, deep learning can automatically extract deep features from data through multi-layer network structure, reduce the complexity of manual feature selection, and significantly improve the processing ability of complex data. Especially in the diagnosis of oil well operating conditions, CNN can learn the dynamic characteristics of oil well operation from the indicator diagram through the convolutional layer, and then complete the classification task of working conditions. The advantage of CNN is its powerful feature extraction capability and multi-level learning structure, which enables the model to extract more accurate pattern information from large-scale data sets without human intervention.

OWDNet (Oil Well Diagnosis Network), a highly targeted deep learning architecture, was designed based on the characteristics of oil well condition diagnosis. Especially considering the timing and complexity of the indicator diagram, the convolutional neural network structure with strong adaptability is adopted. Compared with traditional networks, OWDNet is optimized in terms of the number of layers and parameter configuration, so that the model can not only handle large-scale data,

ISSN 2663-1563 Vol. 7, Issue 1: 17-22, DOI: 10.25236/AJAGE.2025.070103

but also maintain high recognition accuracy under various working conditions. OWDNet uses an adaptive learning mechanism to continuously optimize for increasing data volumes, ensuring that the model can continuously improve diagnostic performance as the field production environment changes. The design and implementation of OWDNet architecture aims to solve the complex problems faced in the diagnosis of oil well conditions through the advantages of deep learning, and realize the intelligent whole process from data acquisition to real-time diagnosis. This not only improves the accuracy and efficiency of well condition diagnosis, but also provides intelligent decision support for oil field production management.

3.1 Performance comparison between traditional machine learning and deep learning in oil well condition diagnosis

Oil well condition diagnosis is a classification and judgment task based on indicator diagram, and traditional machine learning methods usually rely on manual feature extraction for condition analysis. The traditional method, such as grey theory model, transforms the indicator graph into gray matrix and calculates the features such as mean value and variance for classification. The effectiveness of this method is highly dependent on the accuracy of feature selection, and the limitations of artificial feature extraction make this process particularly difficult in the face of complex oil well conditions. In the process of feature extraction, some key information may be ignored or lost, which makes it difficult to accurately identify and classify certain working conditions.

Unlike traditional methods, deep learning models can automatically learn effective features from raw data, avoiding the complexity and information loss of manual feature extraction. Deep neural networks can not only process high-dimensional data through multi-layer nonlinear transformation, but also capture potentially complex patterns in the data, which gives them obvious advantages in the face of diversity and complexity of well conditions. The core of deep learning is the ability to self-optimize the feature extraction process through large amounts of labeled data, improving diagnostic accuracy and the generalization ability of the system. Deep learning algorithms also face a high dependence on data volume and computing power, especially when a large amount of indicator diagram data is required for training, which requires the system to have a high data processing capacity.

With the continuous advancement of oilfield information construction, a large number of indicator diagram data of oil Wells have been automatically collected, greatly enriching the data resources that can be used to train deep learning models. The accumulation of a large amount of historical data provides sufficient training samples for the deep learning model, which enables the technology to show higher accuracy and robustness in the intelligent diagnosis of oil well conditions. The combination of big data and deep learning can not only optimize the limitations of traditional diagnostic technology, but also further improve the automation and intelligence level of oilfield production process, and promote the intelligent transformation of oilfield production.

3.2 Well condition diagnosis model optimization and architecture design based on deep learning

In the intelligent diagnosis of oil well condition, selecting the appropriate deep learning algorithm is the key to ensure the diagnosis accuracy and system robustness. Although traditional deep learning algorithms, deep confidence networks (DBN) and recurrent neural networks (RNN), have made significant progress in some areas, convolutional neural networks (CNNS) are undoubtedly the more ideal choice for well condition diagnosis tasks. CNN can automatically extract multi-level and rich feature information from the input image, which is especially suitable for processing complex image data such as oil well indicator diagram. Its structural design is inspired by biological vision systems, using convolutional layers to gradually extract local features, and pooling layers to effectively reduce feature dimensions while preserving key information. This adaptive feature extraction method can not only automatically find effective patterns, avoid the limitations of manual feature selection, but also greatly improve the accuracy and efficiency of working condition diagnosis.

The characteristics of well condition diagnosis determine the special requirements of CNN architecture. The network structure consists of five convolution layers and three pooling layers, which is designed to dig into the details of the well indicator diagram. Each convolution layer uses a small 3×3 convolution kernel, which is designed to capture fine spatial features in a local region. At the same time, a smaller convolution kernel helps to reduce the amount of computation and avoid over-fitting. The pooling layer adopts the maximum pooling operation of 2×2, which not only helps to reduce the dimension of features, but also further enhances the translation invariance of features and improves the

ISSN 2663-1563 Vol. 7, Issue 1: 17-22, DOI: 10.25236/AJAGE.2025.070103

generalization ability of the network. ReLU activation function is used in the middle layer of the network, and nonlinear transformation is introduced to enable the model to capture more complex feature interaction. With Softmax activation function, multiple classification problems can be effectively handled to achieve accurate classification of well conditions.

OWDNet's design advantage is that it is highly adaptive, able to cope with the variable operating conditions and complex data characteristics of the field production process. With the deepening of information construction of oil Wells and the increasing size of data samples, OWDNet can continuously optimize the model in a large-scale data environment to further improve the diagnostic accuracy. Considering the high dimensionality and diversity of well performance data, OWDNet ensures that the model can fully learn the characteristics of different operating conditions and effectively solve complex and diverse well performance identification problems through careful network structure design. This diagnosis technology based on deep learning provides solid technical support for intelligent management and production safety of oil fields, and promotes the development of intelligent diagnosis technology for oil well conditions. The formula of Softmax and ReLU activation function is as follows:

$$\mathrm{softmax}_i(x) = \frac{e^{x_i}}{\sum_{j=1}^{J} e^{x_j}} \qquad i=1,2,3,\ldots,J \tag{1}$$

$$\mathrm{relu}_i(x) = \max\left(0,\ e^{x_i}\right) \qquad i=1,2,3,\ldots,J \tag{2}$$

The total number of learnable parameters of OWDNet model is as high as 59 million. Although such a large scale of parameters can improve the model's expressibility and learning depth, it also brings significant computational challenges to the training process, and overfitting is prone to occur in high-dimensional data. To address these issues, five Dropout layers are introduced into the model, a design strategy designed to reduce excessive parameter dependence and reduce computational complexity by randomly discarding the participation of a subset of neurons. The Dropout layer simulates the integrated effects of multiple smaller models by randomly masking neurons, thereby improving the model's generalization ability.

This method not only reduces the computational burden, but also avoids the overfitting of training data to a certain extent and improves the adaptability of the model to unknown data in practical applications. Notably, the use of Dropout also supports the stability of the model and the speed of convergence during training, especially in the face of a large number of well performance data, to maintain high training efficiency and accuracy. The design of the Dropout layer is critical to improving OWDNet's performance in complex oilfield environments. By this means, the model can prevent overfitting while maintaining strong prediction ability and good robustness when dealing with high-dimensional and large-scale data sets, ensuring the accuracy and reliability of intelligent diagnosis of oil well conditions.

4. OWDNet model training and field application

4.1 OWDNet neural network model training and optimization

The core goal is to optimize the parameters in the convolutional layer and the fully connected layer so that the model can achieve the smallest possible error between the input data and the actual label. This process is not only dependent on the selection of the loss function, but is also influenced by the optimizer's strategy, which determines how the weights and biases in the network are updated in each training iteration. The training process usually starts from the input of data samples, gets the forecast output through forward propagation, and measures the difference between the prediction of the current model and the actual label through the loss function. Optimizers such as backpropagation algorithm and gradient descent adjust the weight of the network according to the size of the loss, so that the model can gradually reduce the error and improve the classification accuracy. The cross entropy loss function was chosen as the main loss measure. The cross-entropy function has the advantage of quantifying the distance between the probability distribution and the true label when dealing with multi-class classification problems.Cross-entropy assesses the classification effectiveness of a model by comparing the "proximity" of two probability distributions. Since the diagnosis of oil well conditions is usually faced with multiple types of conditions, cross entropy provides an error measurement method suitable for this kind of problems, which is helpful to guide the weight updating during training.

In order to further improve training efficiency and convergence speed, more complex optimization algorithms such as Adam optimizer may be used in the training process of OWDNet, which automatically adjust the learning rate during the update process and effectively deal with the problem of gradient disappearance or gradient explosion in deep learning. The success of neural network training depends not only on the choice of loss function and optimizer, but also on many factors such as network structure and data preprocessing. Combined with the characteristics of the oil well condition diagnosis task, a reasonable neural network training strategy can ensure the efficiency and generalization ability of the model, and improve its accuracy and stability in practical application.

4.2 OWDNet intelligent diagnosis application and implementation in oil field

A highly automated intelligent diagnosis system for oil well condition is constructed. The system integrates real-time data acquisition, processing and diagnostic push functions, and the architecture is shown in Figure 5. The design of the system includes two key workflows: (1) The condition diagnosis process, which ensures real-time monitoring information in the oilfield database through the continuous inflow of well sensor data. The system background will obtain the latest data from the database regularly, generate indicator diagram after standardization and preprocessing, and input it into the trained OWDNet model for working condition identification and classification. The diagnostic results are automatically transmitted to the field management system for analysis and decision making by field managers to further support the intelligent control of the well production process. ② The system has a continuous learning module, which can be manually corrected by oil field staff when misdiagnosis is found, and the corrected data is fed back to the database. When the corrected data accumulates to a certain amount, the OWDNet model will be trained increments on a regular basis to ensure that the model constantly updates and optimizes itself as oilfield conditions change. This incremental learning mechanism enables the system to adapt to complex and varied oil well working conditions and improves its diagnostic accuracy and adaptability.

As of October 2021, the intelligent well condition diagnosis system has processed more than 5 million condition diagnoses, and the field diagnosis accuracy rate has steadily increased to 90%. The average response time from data acquisition to diagnosis push is only 2 minutes, which significantly improves the detection and treatment efficiency of abnormal well conditions. The intelligent diagnostic system has improved field production control, increasing the proportion of continuously stable production Wells from 68% to 88%. This development not only provides solid technical support for the efficient and safe operation of the field, but also promotes the further improvement of the recovery rate of the well. Field managers save time on inspection and monitoring, allowing them to focus on more strategic tasks such as production optimization, risk management and decision support. The continuous optimization and intelligent operation of the system not only improves the operational efficiency of the oilfield, but also provides valuable experience for the digital transformation of future oilfield production.

5. Conclusion

Based on the information construction of S oilfield, this study collected the indicator diagram data of 1,930 oil Wells and constructed a working condition diagnosis sample set containing 40 million sets of data. By designing the specification of superposition indicator diagram and combining with expert experience, 37 common working conditions of 5 categories were summarized, and an iterative labeling strategy was proposed to improve the efficiency of data labeling. By comparing traditional machine learning methods with deep learning methods, Convolutional neural network (CNN) was used to build a 26-layer OWDNet model with 59 million learnable parameters. After training, the accuracy rate reached 99.7%. On this basis, the intelligent diagnosis system of oil well condition has been developed, and has completed 5 million diagnoses with an accuracy rate of 90%, which significantly improves the efficiency of oil well control and production stability, and promotes the efficient production and recovery of oil fields. In the future, the research will focus on expanding the sample base, optimizing the data distribution, and introducing more monitoring indicators to further improve the diagnostic accuracy and application range of the system.

ISSN 2663-1563 Vol. 7, Issue 1: 17-22, DOI: 10.25236/AJAGE.2025.070103

References

[1] Wu Y, Ren X, Cao Z,et al. Diagnosis and analysis of repeatable diseases based on deep intelligent model driven by big data of medical images[J]. Digital Communications and Networks, 2021, (2). DOI:10.1016/j.dcan.2021.04.004.

[2] Wosowei J B, Shastry C. An Efficient Intelligent Oil Well Monitoring System for Niger Delta Oil Fields[J]. International Journal of Advanced Research in Science, Communication and Technology, 2023. DOI: 10.48175/ijarsct-9563.

[3] Deng Z K, Chang P X, Zhou Z P,et al. Application of Intelligent Rodless Artificial Lift Technology in Shale Oil of Changqing Oilfield[C]. International Petroleum and Petrochemical Technology Conference. Springer, Singapore, 2023. DOI:10.1007/978-981-99-2649-7_47.

[4] Huang Z, Li K, Ke C,et al. An intelligent diagnosis method for oil-well pump leakage fault in oilfield production Internet of Things system based on convolutional attention residual learning[J]. Engineering Applications of Artificial Intelligence, 2023. DOI:10.1016/j.engappai.2023.106829.