International Journal of Science and Engineering Applications
Volume 15-Issue 05, 86 – 91, 2026, ISSN:- 2319 - 7560
DOI: 10.7753/IJSEA1505.1014

Fault Diagnosis of Pumping Units Based on an Improved LeNet Model

Sida Wang
School of Electric Information and Electrical Engineering
Yangtze University
Jingzhou, China

Ying Wang
School of Electric Information and Electrical Engineering
Yangtze University
Jingzhou, China

Yawen Dou
School of Electric Information and Electrical Engineering
Yangtze University
Jingzhou, China

Abstract: To address the problems of insufficient feature representation in traditional models and high computational cost in deep networks for working-condition diagnosis of sucker-rod pumping wells, this paper proposes an improved LeNet-based working-condition diagnosis method for the real-time diagnostic requirements of the host computer. The proposed method takes indicator diagrams as input and first performs preprocessing on the samples. Then, based on the traditional LeNet architecture, the Mish activation function, Triplet Attention mechanism, and adaptive global average pooling are introduced to enhance the model's ability to extract local features among similar working conditions, improve prediction accuracy and the interpretability of diagnostic results, reduce the number of parameters, and optimize the model structure. Experiments are conducted using a dataset collected from a sampled oilfield. The results show that the improved LeNet can identify eight common working conditions, achieving a final prediction accuracy of 97.69% and a single-sample processing time of 0.36 ms. Therefore, the proposed method can meet the requirements of fast and stable working-condition diagnosis on the host computer and has certain engineering application value.

Keywords: LeNet model; cross-dimensional attention; dynamometer card of pumping unit; fault diagnosis

1. INTRODUCTION

In the field of working-condition monitoring for sucker-rod pumping wells, indicator diagrams are plotted from the periodically collected polished-rod load and displacement data. These diagrams provide an important basis for fault diagnosis of the subsurface pump, sucker rod, and downhole supply–discharge relationship, and have significant application value in evaluating single-well production, extending equipment service life, and reducing production energy consumption. Therefore, intelligent working-condition diagnosis of sucker-rod pumping wells has become an important research topic in the digital development of oilfields.

Because sucker-rod pumping systems operate for long periods in complex and harsh downhole environments, downhole working conditions are characterized by frequent occurrence, concealment, and coupling. As a result, some working conditions exhibit highly similar graphical features. Traditional manual diagnosis is strongly influenced by individual experience and subjective judgment, making it difficult to accurately distinguish subtle differences among similar working conditions. Compared with manual methods, indicator-diagram diagnosis methods based on neural network models possess nonlinear fitting and adaptive feature learning capabilities, making them suitable for multi-class working-condition diagnosis tasks and providing technical support for the accurate identification of similar fault conditions.

In early studies, Zhang et al.[1] used Freeman chain codes and differential coding to extract contour features from indicator diagrams, and constructed a BP neural network with a simple structure, achieving a recognition accuracy of 94.79%. However, this method relies heavily on manually designed features, and the training process is prone to falling into local optima, resulting in insufficient generalization ability when dealing with similar samples.Dong et al.[2] further applied the LeNet convolutional neural network to indicator-diagram diagnosis, introducing Dropout and local response normalization layers to reduce the risk of overfitting, and achieved an average diagnostic accuracy of 94.68%. This study verified the feasibility of CNNs in indicator-diagram diagnosis. However, the traditional LeNet structure is relatively shallow, and its feature representation ability is weaker than that of deep networks, leaving room for improvement in handling complex working conditions.To address the difficulties in training deep neural networks and the limited number of indicator-diagram samples, Zhang et al.[3] proposed a working-condition diagnosis method based on CNN and transfer learning. In this method, pretrained AlexNet and GoogLeNet were used to extract deep image features, and operations such as mirror expansion were adopted to improve model stability. The experimental results showed that transfer learning can alleviate the difficulty of model training in small-sample scenarios; however, the model scale is large, leading to higher requirements for computational resources and the training environment.

With the widespread application of neural network methods, Ye et al.[4] constructed a CNN prediction model based on AlexNet-BN, in which batch normalization layers were used to replace LRN and Dropout. The model achieved a diagnostic accuracy of 96.05% in the recognition of 15 common working conditions. This study indicates that, compared with manually designed features, deep learning can effectively reduce errors caused by manual feature selection and is suitable for multi-class classification tasks.To address problems such as label overlap and feature flattening in compound working-condition diagnosis, Zhang W[5] proposed a diagnostic method for compound working conditions. Although this method enhanced the recognition capability for compound working conditions, its model structure was more complex and the labeling cost of sample data was relatively high.In addition, Li J[6] designed a working-condition diagnosis method by combining electrical parameter features and completed the fault identification task based on an improved LeNet-5

International Journal of Science and Engineering Applications
Volume 15-Issue 05, 86 – 91, 2026, ISSN:- 2319 - 7560
DOI: 10.7753/IJSEA1505.1014

network. This study verified the feasibility of collaborative diagnosis using electrical parameters, indicator diagrams, and CNNs. However, its final test accuracy was only 83.3%, indicating that the method still had limitations in terms of model accuracy and diagnostic stability.

A summary of previous studies shows that neural-network-based indicator-diagram working-condition diagnosis methods have made progress in feature extraction, transfer learning, and improved network structures. However, several limitations remain. First, some methods rely on manually designed features, making it difficult to fully adapt to graphical differences under different load scales. Second, deep networks usually have a large number of model parameters and high computational complexity, which is not conducive to fast and stable deployment on the host computer. Third, most existing methods are based on general CNNs or large-scale networks, while lightweight models specifically designed for practical working-condition diagnosis requirements are still lacking.

Based on the above analysis, this paper proposes a working-condition diagnosis method based on an improved LeNet model to meet the diagnostic requirements of sucker-rod pumping wells on the host computer. The proposed method supports the prediction and diagnosis of eight common working conditions and faults, including normal operation, sucker-rod breakage, and wax deposition. Taking indicator diagrams as the research object, the input samples are preprocessed through cropping, normalization, and other operations to obtain standardized inputs, thereby reducing the influence of differences in sample distribution on model training.In terms of model structure, the proposed method improves the traditional LeNet architecture by replacing the conventional ReLU function with the Mish activation function, which helps alleviate the problem of neuron inactivation. After deep feature extraction, a cross-dimensional attention mechanism is introduced to collaboratively model the feature information of indicator diagrams from the height, width, and channel dimensions, enabling the model to adaptively focus on key curve regions and reducing the interference caused by large blank backgrounds and random noise in the classification results.To address the problem that the flattening operation in traditional CNNs may lead to an increase in parameters, this paper adopts a global average pooling structure to compress deep feature maps into global semantic vectors. This reduces model complexity while improving the model's noise robustness and the interpretability of diagnostic results. Therefore, the proposed method provides a solution for sucker-rod pumping well working-condition diagnosis that achieves both high recognition accuracy and real-time operation.

This paper uses an indicator-diagram sample dataset collected and generated from the production site to conduct training and testing experiments for the improved LeNet model. In addition, several working-condition diagnosis models are reproduced on the host computer for comparative experiments, and the performance advantages of the proposed improved LeNet model are comprehensively analyzed.

2. Fault Diagnosis Method

2.1 Sample Preprocessing

The closed graph formed by the curve describing the relationship between polished-rod load and displacement in a pumping unit is called the polished-rod indicator diagram, which reflects the variation relationship between load and displacement during one stroke cycle. When convolutional neural networks are applied to working-condition diagnosis, the indicator-diagram samples need to be preprocessed first. In this study, eight working conditions of sucker-rod pumping wells are investigated, and some original indicator diagrams are shown in Figure 1.

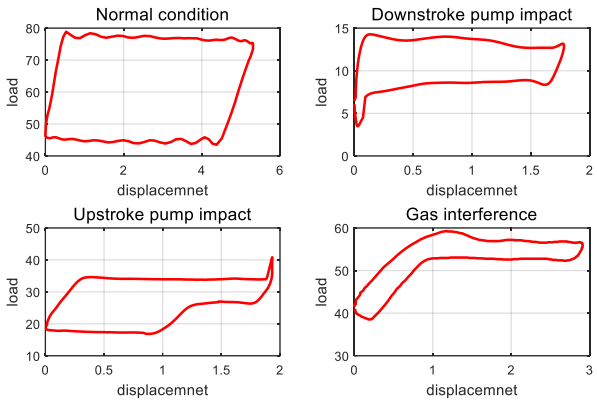

Figure 1:Partial Indicator Diagram Examples

It can be seen from the figure that the indicator-diagram samples contain structural noise such as coordinate axes and tick marks. These elements are fixed geometric structures in the image, and the model may treat them as stable features for extraction and learning, thereby causing misclassification. Therefore, in order to provide effective samples without ambiguity, the input images are first resized to a fixed size to ensure consistency in subsequent training and processing. Using the prior knowledge that the curve color in the original indicator diagrams is red, the RGB images are converted into the HSV color space, and the distribution threshold of red in the HSV space is determined for color segmentation. Information outside the threshold range is removed, and only the contour information of the indicator-diagram curves is retained. Subsequently, morphological opening is applied to remove isolated noise points, and dilation and thinning are used to connect possibly broken curves. Finally, the connected-component analysis method is employed to retain the largest connected region, and binarization is performed to obtain an image containing only the target curve contour, as shown in Figure 2.

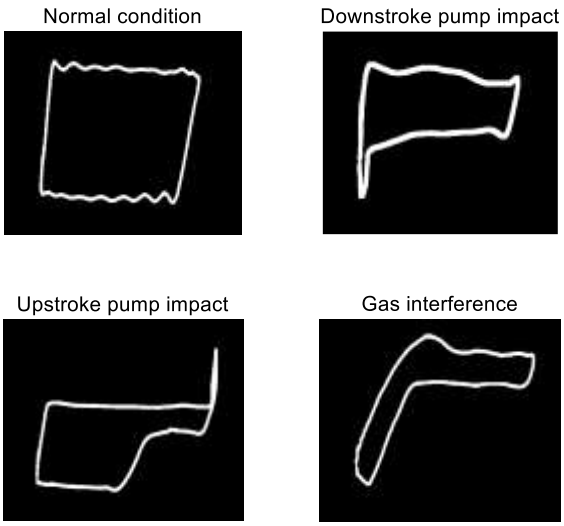

Figure 2:Preprocessing Results

By distinguishing color information, this method effectively solves the problem of irrelevant information such as coordinate axes being incorrectly detected by traditional grayscale

International Journal of Science and Engineering Applications
Volume 15-Issue 05, 86 – 91, 2026, ISSN:- 2319 - 7560
DOI: 10.7753/IJSEA1505.1014

segmentation methods, thereby improving the accuracy of curve extraction.

2.2 Establishing the Improved LeNet Model

For the working-condition recognition task of pumping-unit indicator diagrams, this paper improves the basic LeNet-CNN structure and designs an improved LeNet model based on multi-scale feature extraction and a cross-dimensional attention mechanism. The model takes normalized images as input and realizes the classification of complex working conditions through layer-by-layer feature extraction and semantic enhancement. The model structure is shown in Figure 3.

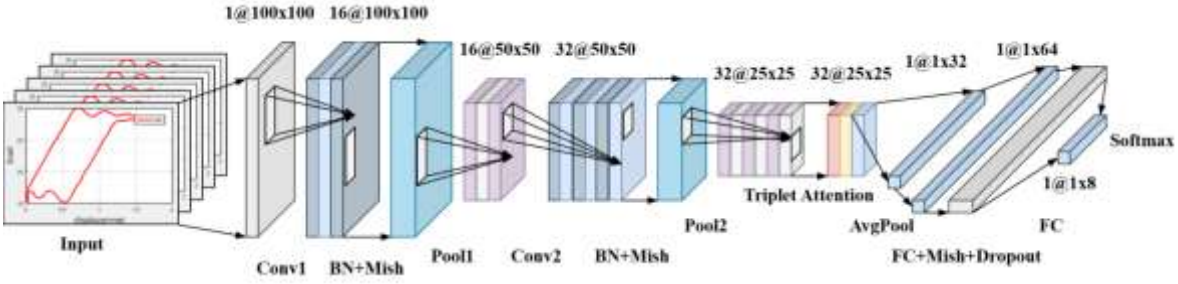

Figure 3:Architecture of the Improved LeNet Model

In the preprocessing stage, the original indicator diagrams are resized and normalized to obtain standard images with a size of 1×100×100, thereby reducing the distribution differences among different samples and providing a unified scale for subsequent feature extraction.

The first convolutional layer, BN layer, and Mish activation function together constitute the shallow feature extraction module. The first convolutional layer uses a 5×5 convolution kernel with a stride of 1 and a padding of 2, which keeps the feature map size unchanged while mapping the single-channel input to 16 feature channels, thereby effectively capturing local edges, curve variations, and basic morphological features of the indicator diagram. The BN layer is introduced to improve gradient propagation efficiency and enhance training stability. In addition, the Mish activation function is adopted instead of ReLU to strengthen the nonlinear representation capability of the model and improve its sensitivity to feature information.

After shallow feature extraction, a 2×2 max-pooling operation is introduced to reduce the feature map size from 100×100 to 50×50. This operation lowers the feature dimensionality and reduces computational cost while preserving the main structural information.

The second convolutional layer, BN layer, and Mish activation function together constitute the deep feature extraction module. The second convolutional layer also adopts a 5×5 convolution kernel, while expanding the number of channels to 32 to further enhance the semantic representation capability. Combined with the BN layer, Mish activation function, and another pooling operation, the feature map is compressed to 32×25×25. In this stage, the model focuses on learning the overall contour structure of the indicator diagram and the discriminative features of different working conditions, such as the inclination trend of the unloading curve and changes in the shape of the closed region, thereby improving its ability to distinguish complex working conditions.

After deep feature extraction, a cross-dimensional attention mechanism is introduced to model the global dependency relationships of features in parallel from three dimensions: height–channel, width–channel, and height–width. By identifying feature representations that remain stable across multiple dimensions, the model obtains the attention weight distribution specific to each working condition, thereby effectively enhancing feature representation capability and improving the interpretability of diagnostic results. Image overlay is adopted to present the attention visualization results. The attention heatmaps for normal condition, pump bumping at the top, pump bumping at the bottom, and gas interference are shown in Figure 4.

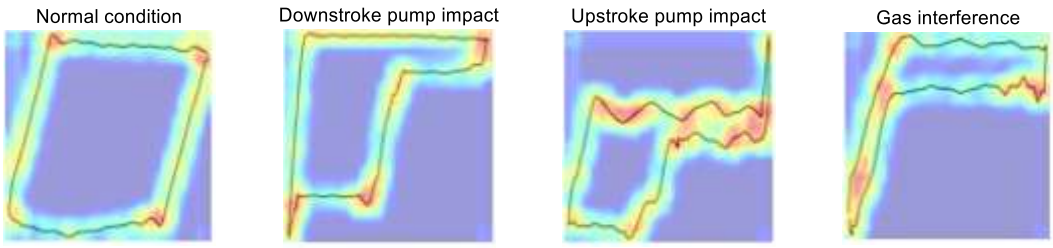

Figure 4:Attention Heat Maps of Four Working Conditions

In the feature compression stage, adaptive global average pooling is used instead of the traditional direct flattening operation to process the 32×25×25 feature maps. Specifically, the mean value of the 25×25 feature response in each channel is first calculated, and then the result is transformed by the Flatten operation into a 1×32 feature vector. This significantly reduces the number of parameters and improves the model's robustness to image translation and positional variations.

In the fully connected layer, the model maps the 32-dimensional feature vector to a 64-dimensional high-level semantic space. The Mish activation function is then used to further enhance the nonlinear representation capability, and the Dropout value is set to 0.2 to alleviate overfitting. Finally, the Softmax function outputs the probability weights of the eight diagnostic categories, enabling the identification of the working conditions of sucker-rod pumping wells.

International Journal of Science and Engineering Applications
Volume 15-Issue 05, 86 – 91, 2026, ISSN:- 2319 - 7560
DOI: 10.7753/IJSEA1505.1014

By introducing global information integration through the cross-dimensional attention mechanism, the discriminative capability of the improved LeNet is effectively enhanced, providing a reliable network basis for the diagnosis of complex working conditions.

3. Experiments and Analysis

3.1 Dataset Description

At an oil production site, polished-rod load and displacement data were collected using an indicator diagram tester. The indicator diagrams obtained from several pumping units over a period of time were screened and categorized, resulting in eight common working conditions of sucker-rod pumping wells, including normal operation, wax deposition, and pump bumping at the top. A total of 7,203 representative samples were selected, and the dataset was divided into training and testing sets at a ratio of 7:3 to complete the training and testing of the improved LeNet model. The mapping relationship between training labels and working-condition types was jointly determined based on equipment operating status and manual diagnosis results. The sample distribution and label combinations of the dataset are recorded in Table 1.

Table 1:Table of Dataset Sample Distribution

Working Condition	Label	Training Set	Test Set	Total Samples
Normal	A01	774	332	1106
Upstroke Pump Impact	A02	725	314	1039
Downstroke Pump Impact	A03	771	340	1111
Rod Breakage	A04	447	191	638
Wax Deposition	A05	493	211	704
Standing Valve Leak	A06	485	213	698
Gas Interference	A07	625	273	898
Plunger Disengagement	A08	706	303	1009

3.2 Experimental Environment

The improved LeNet model was trained on a Windows 11 operating system equipped with an NVIDIA GeForce RTX 4060 Laptop GPU and an Intel Core Ultra 5 125H @ 3.60 GHz CPU. The experimental environment was configured with Anaconda3 and CUDA 13.0. The model was implemented using the Python 3.13.5 programming language and the PyTorch-based inference engine.

3.3 Training Experiment of the Improved LeNet

The training process adopts the Adam adaptive gradient optimization method based on backpropagation. The number of training epochs is set to 100, and the batch size is set to 64. In each epoch, the training set is divided into 81 batches. The cross-entropy function is selected as the loss function, and class weights are introduced to reduce the influence of sample imbalance on the training results. The Adam optimizer is used with an initial learning rate of 0.001. In addition, a StepLR learning rate scheduler is introduced, which decays the learning rate to 0.5 times its original value every 10 epochs, so as to improve the convergence stability in the later stage of training.

To observe the model training process, this paper records the training loss, test loss, training accuracy, and test accuracy at the end of every 10 epochs, and plots the training history curves accordingly, as shown in Figure 5.

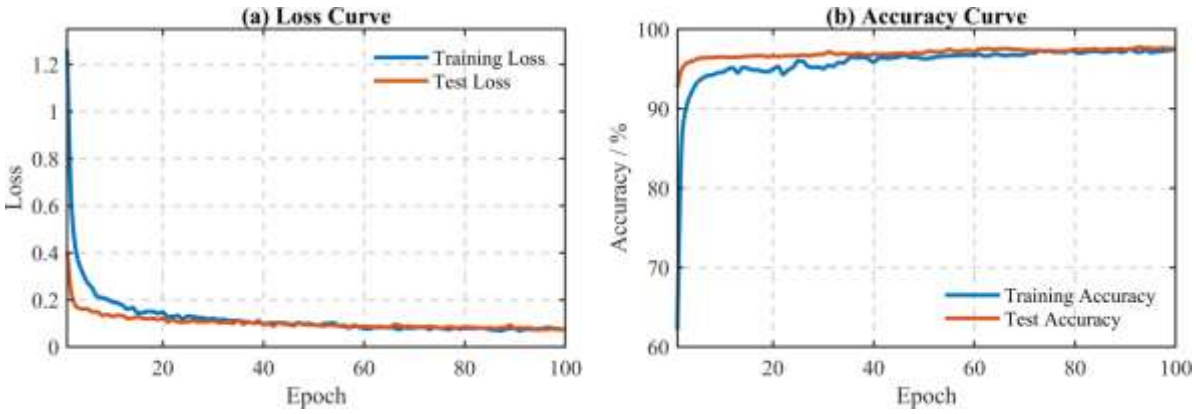

Figure 5:Training History Curve

From the training history curves, it can be observed that both the training loss and test loss decrease rapidly in the early stage and then gradually stabilize at relatively low levels, indicating that the model can effectively learn the features of indicator diagrams. In

International Journal of Science and Engineering Applications
Volume 15-Issue 05, 86 – 91, 2026, ISSN:- 2319 - 7560
DOI: 10.7753/IJSEA1505.1014

the accuracy curves, both the training accuracy and test accuracy increase rapidly with the number of epochs during the first 10 epochs, and then remain at a high level in the subsequent training process. The test accuracy exceeds 97% after 75 epochs and finally reaches 97.88%.

To further verify the recognition performance of the improved LeNet model on the test dataset, precision, recall, and F1-score are introduced as evaluation metrics. Precision is denoted as Precision, recall is denoted as Recall, and F1-score is denoted as F1. Their calculation formulas are as follows:

$$Precision = \frac{TP}{TP + FP} \quad (1)$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (2)$$

$$F1 - score = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (3)$$

Here, TP denotes the number of positive samples correctly predicted as positive, FP denotes the number of negative samples incorrectly predicted as positive, and FN denotes the number of positive samples incorrectly predicted as negative. In the multi-class classification task, each working condition is regarded as the positive class in turn, while the remaining classes are treated as negative classes, and the Precision, Recall, and F1-score are calculated for each working-condition category.

The classification report of the improved LeNet model on the test set is shown in Table 2. According to the table, the model achieves good overall performance in recognizing the eight types of indicator-diagram working conditions, with the precision, recall, and F1-score of each category remaining at relatively high levels. In particular, the F1-scores of all eight working conditions are above 0.96, and no category shows significantly lower recognition performance, indicating that the model has relatively stable classification performance across different working conditions.

Table 2:Classification Report Results of the Test Set

Working Condition	Precision	Recall	F1-score
Normal	0.98	0.98	0.98
Upstroke Pump Impact	0.97	0.96	0.97
Downstroke Pump Impact	0.98	0.99	0.99
Rod Breakage	0.96	0.97	0.97
Wax Deposition	0.99	0.98	0.99
Standing Valve Leak	0.98	0.98	0.98
Gas Interference	0.97	0.97	0.97
Plunger Disengagement	0.99	0.98	0.99

3.4 Contrast Experiment

To further verify the effectiveness of the improved LeNet model, this paper uses the same dataset to conduct prediction experiments on several neural network models, including the original LeNet[7], intelligent BP network[8], SRPNet[9], and VGG16[10], on the host computer platform. A comprehensive comparative analysis is carried out using evaluation indicators such as the total parameters, computational operations, single-sample processing time, and prediction precision. The structural comparison results of different models are shown in Table 3.

Table 3:Results of Contrast Experiment

Model Type	Total parameters/K	Computational Operations/M	Processing Time ms/sample	Prediction Precision
Improved LeNet	14.99	37.2	0.36	97.69%
Original LeNet	30.93	20.71	1.23	93.51%
BP Network	0.918	0.09	3.93	86.69%
SRPNet	669.48	807.35	9.53	94.63%
VGG16	14722.89	4710.96	7.44	92.77%

As shown in Table 3, the intelligent BP network has the smallest number of parameters and computational operations. However, due to its simple structure and low computational complexity, its feature extraction capability is relatively limited, and its prediction precision is only 86.69%. In contrast, the original LeNet model extracts features through convolutional operations and enhances the perception of image regions, improving the prediction precision to 93.51%. However, it still has room for further optimization, with 30.93 K parameters and a single-sample processing time of 1.23 ms per sample. The proposed

improved LeNet model reduces the total number of parameters to 14.99 K and the single-sample processing time to 0.36 ms per sample, while increasing the prediction precision to 97.69%. This indicates that the introduction of the Mish activation function, cross-dimensional attention mechanism, and adaptive global average pooling improves both the feature representation capability and recognition accuracy of the model.The model scale and computational operations of SRPNet and VGG16 are significantly larger. In particular, VGG16 has 14722.89 K parameters and 4710.98 M computational operations, resulting in high computational cost and greater requirements for hardware resources. SRPNet also reaches 807.35 M computational operations and has a single-sample processing time of 9.53 ms per sample. However, the prediction precision of both models is lower than that of the improved LeNet.Overall, the improved LeNet model achieves a good balance between classification performance and computational efficiency. Compared with the original LeNet, the improved LeNet reduces the number of parameters from 30.93 K to 14.99 K while increasing the prediction precision from 93.51% to 97.69%. Compared with more complex models such as VGG16 and SRPNet, the improved LeNet achieves higher prediction precision with a smaller model scale and lower computational cost. Compared with the intelligent BP network, although the improved LeNet has a larger number of parameters and computational operations, it significantly improves the recognition accuracy of indicator-diagram working conditions. Therefore, the proposed improved LeNet model can more effectively extract the morphological features of indicator diagrams and is suitable for fast and accurate identification of oil-well working conditions.

4. Conclusion

This paper proposes an improved LeNet model for the working-condition recognition task of indicator diagrams. Based on the traditional LeNet architecture, the Mish activation function, cross-dimensional attention mechanism, and adaptive global average pooling are introduced to enhance the model's ability to extract local features among similar working conditions, improve prediction accuracy, and increase the interpretability of diagnostic results. Meanwhile, adaptive global average pooling is adopted to reduce feature dimensionality, decrease the number of model parameters, and optimize the model structure.

The results of the comparative experiments show that the improved LeNet achieves a good balance between accuracy and complexity. The model has 14.99 K parameters, 37.2 M computational operations, and a single-sample processing time of 0.356 ms, while its prediction precision reaches 97.69%. The experimental results indicate that the improved LeNet further improves recognition accuracy while reducing the number of parameters and inference time, demonstrating good overall performance.

The improved LeNet model proposed in this paper can effectively address the problems of insufficient feature representation in traditional models and high computational cost in deep networks for working-condition diagnosis tasks. It has the advantages of high accuracy, a small number of parameters, and fast inference speed. Therefore, it is suitable for oil-well indicator-diagram working-condition recognition, pumping-unit operating-state monitoring, and rapid on-site fault diagnosis, and has certain engineering application value.

5. ACKNOWLEDGMENTS

This work would not have been possible without the guidance and advice of my supervisor. I also sincerely thank my team members for their dedicated collaboration, and my family for their constant encouragement and support.

6. REFERENCES

[1] Zhang, L., Du, Q., Liu, T., and Li, J. 2020. A Fault Diagnosis Model of Pumping Unit Based on BP Neural Network. 2020. International Conference on Networking and Network Applications.

[2] Dong, Y., Niu, L., Bai, Y., Wang, L., and Liu, Y. 2024. A Study on the Diagnosis of the Working Conditions of a Traveling Beam Pumping Unit Based on Artificial Intelligence. In Innovative Design and Intelligent Manufacturing.

[3] Zhang, R., Wang, L., and Chen, D. 2021. An Intelligent Diagnosis Method of the Working Conditions in Sucker-Rod Pump Wells Based on Convolutional Neural Networks and Transfer Learning. Energy Engineering, 118(3).

[4] Ye, Z., and Yi, Q. 2022. Working-condition diagnosis of a beam pumping unit based on a deep-learning convolutional neural network. Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science, 236(5): 2559–2573.

[5] Zhang, W. 2025. Research on Composite Working Condition Diagnosis Method of Oil Well Dynamometer Cards Based on Deep Learning. Qufu Normal University.

[6] Li, J. 2018. Research on Working Condition Diagnosis of Pumping Wells Based on Electrical Parameters Using Machine Learning. China University of Petroleum.

[7] Ye, Z., Yi, Q., and Luo, L. 2023. Research on Working Condition Diagnosis of Beam Pumping Unit Based on LeNet Model. Journal of Southwest Petroleum University (Science & Technology Edition), 45(6): 164–174.

[8] Zhang, N. 2017. Research on Dynamometer Card Fault Diagnosis Based on Intelligent BP Neural Network. Lanzhou University of Technology.

[9] He, J., Chen, P., Zhao, L., and Miao, K. 2023. Fault Diagnosis of Pumping Unit Based on Improved Prototypical Network Method. Oil Drilling & Production Technology, 45(3): 312–318.

[10] Zhu, D. 2022. Research on Comprehensive Analysis System of Oil Wells Based on Neural Network and Dynamometer Card. Xi'an Shiyou University.