

Research on Intelligent Recognition of Calibration Charts Based on an Improved MobileNetV3 Model

Guoqing Cai

School of Electric Information and Electrical Engineering

Yangtze University

Jingzhou, China

Abstract: Load diagrams directly characterize the operational status of pumping wells and serve as a core diagnostic tool for identifying malfunctions. To address the limitations of traditional load diagram recognition—which relies on manual expertise, suffers from low efficiency, and lacks sufficient accuracy—this paper proposes an intelligent load diagram recognition method based on an enhanced MobileNetV3 architecture. This approach replaces the SE module in the original MobileNetV3 network with a CBAM attention mechanism, thereby enhancing feature expression capabilities across both channel and spatial dimensions. Using a real-world indicator diagram dataset from an oilfield, classification experiments were conducted on nine typical operating conditions. Experimental results demonstrate that the improved MobileNetV3_CBAM model outperforms other comparison models in accuracy, precision, recall, and F1 score metrics. Specifically, precision reaches 95.4%, with overall recognition accuracy improving by approximately 1.6% compared to the original model. The research findings demonstrate that the proposed improvement method effectively enhances the accuracy and robustness of load curve recognition, providing reliable technical support for pumping unit fault diagnosis and intelligent oilfield development.

Keywords: MobileNetV3; CBAM attention mechanism; load curve recognition; deep learning

1. Introduction

The load-displacement curve is a crucial characteristic curve reflecting the relationship between the load and displacement of a pumping unit over time. By analyzing this curve, one can effectively evaluate the operational status of the pumping unit, optimize production parameters, and achieve rapid fault diagnosis [1]. However, traditional load-displacement curve recognition primarily relies on manual experience, which not only involves substantial labor and low efficiency but also yields results susceptible to human factors. This approach struggles to meet the real-time and accuracy requirements of intelligent oilfields.

In recent years, with the advancement of deep learning technology, convolutional neural networks (CNNs) have demonstrated superior feature extraction capabilities in image recognition and are widely

applied in industrial inspection and fault diagnosis tasks [2]. The lightweight MobileNetV3[3] network, with its high computational efficiency and low parameter count, has become a key model for embedded devices and edge computing. However, its built-in SE attention mechanism focuses solely on channel information, limiting its ability to extract spatial features.

To address these limitations, this paper proposes an enhanced version of MobileNetV3 by replacing the SE module with the CBAM attention mechanism. This modification strengthens feature representation capabilities in both the channel and spatial dimensions. Experimental validation on a real-world gauge chart dataset demonstrates that the proposed model maintains lightweight characteristics while improving recognition accuracy and generalization ability. This approach offers an effective new strategy for

automated gauge chart recognition and intelligent monitoring of pumping units.

2. Materials and Methods

2.1 Data Collection

This paper utilizes actual production test data from a domestic oilfield to preprocess displacement and load information and plot load-displacement curves. Nine typical abnormal operating conditions were selected for study, including: normal operation, gas interference, insufficient fluid supply, pump-up collision, pump-down collision, fixed valve leakage, slow-closing floating valve, plunger disengagement, and floating valve leakage. The classification results are shown in Figure 1.

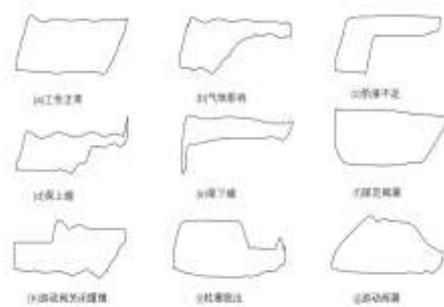


Figure 1: Classification of Work-Energy Diagrams

2.2 Data Augmentation

Since oilfield pumping units operate under normal conditions most of the time, fewer abnormal test curve images are obtained. To avoid excessive normal-condition samples causing uneven data distribution and prevent model overfitting during training, data augmentation techniques such as binarization, rotation, scaling, and flipping were applied to expand the remaining eight categories of test curves. After augmentation, each category now contains 500 images. The classification and labeling of the indicator diagram categories are shown in Table 1. The generated indicator diagrams were divided into a training set and a test set in an 8:2 ratio.

Fault Type	Quantity	Label
Operating normally	500	0
Gas Effects	500	1
Insufficient fluid supply	500	2
Pump collision	500	3
Bottom-end collision	500	4
Fixed valve leakage	500	5
The floating valve closes slowly.	500	6
Piston rod dislodged	500	7
Leak in the floating valve	500	8

Table 1: Indicator Diagram Dataset

2.3 MobileNetV3 Network

MobileNetV3 is a novel lightweight network architecture proposed by the Google team in 2019 [4]. It inherits the separable convolutions from MobileNetV1 and the linear bottleneck-inverted residual structure from MobileNetV2. Building upon this foundation, MobileNetV3 further optimizes the architecture by introducing the SE module within blocks to enhance feature expression capabilities and adopting the novel activation function h-swish to replace ReLU6. These improvements significantly reduce computational complexity and memory consumption. To accommodate diverse application scenarios and computational resource constraints, MobileNetV3 offers both Large and Small variants. The Large variant is selected for this study.

2.4 SE Attention Mechanism

Attention mechanisms play a crucial role in deep learning [5]. The SE attention mechanism is a lightweight channel attention module whose core idea is to adaptively reweight the channel dimension of feature maps to enhance feature representation capabilities. This mechanism achieves channel-adaptive weighting through three stages: compression, activation, and recalibration. Compared to traditional methods, the SE mechanism offers greater

expressiveness and flexibility in the channel dimension while maintaining low computational overhead, enabling seamless integration into various CNNs to boost performance. However, its primary focus on channel-dimension modeling limits its ability to capture spatial information interactions.

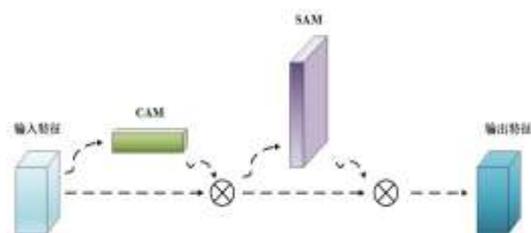
3. Improved MobileNetV3 Network Approach

Traditional performance chart recognition relies on manual expertise, resulting in high labor intensity and relatively low efficiency and accuracy. The development of intelligent oilfields demands higher standards for the efficiency and precision of performance chart recognition, rendering manual identification methods inadequate for practical needs. To enhance recognition efficiency and accuracy, this paper introduces the CBAM attention mechanism based on the MobileNetV3 model, replacing the original SE attention mechanism to improve overall network performance.

3.1 CBAM Attention Mechanism

Attention mechanisms are computational models that simulate how the human visual system processes information. By assigning different weights to various inputs, they optimize information processing efficiency, thereby enhancing the overall performance of neural networks, as shown in Figure 2. CBAM (Convolutional Block Attention Module) is a lightweight, simple, and efficient attention module specifically designed for convolutional neural networks (CNNs). It can be seamlessly integrated into most CNN architectures, significantly improving network performance.

Figure 2: CBAM Attention Module Structure Diagram



Channel attention highlights category-relevant features by weighting feature channels; spatial attention precisely locates target regions by weighting spatial positions. CBAM consists of a Channel Attention

Module (CAM) and a Spatial Attention Module (SAM) connected in series. Input feature maps first undergo CAM processing to adjust channel weights, then pass through SAM to optimize spatial position weights, ultimately yielding adjusted feature maps.

The channel attention mechanism focuses on the importance of each channel within the feature map. By assigning weights to different channels, it highlights key information while suppressing redundant features. Within the Channel Attention Module (CAM), the input feature undergoes global average pooling and max pooling to extract global information. Finally, a Sigmoid activation function generates channel weights, which are multiplied with the original feature on a per-channel basis to achieve adaptive feature optimization. The structural diagram of the channel attention module is shown in Figure 3.

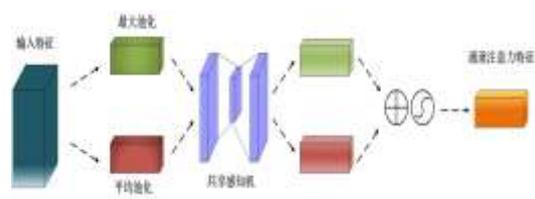


Figure 3: CAM Attention Module Structure Diagram

As shown in Figure 4, the spatial attention mechanism further enhances the expression of key information by focusing on the importance of different spatial locations within feature maps. It first concatenates the input feature maps after applying global average pooling and max pooling. These concatenated maps are then fused and reduced in dimension through a 2D convolutional layer. A Sigmoid activation function is applied to generate a spatial weight map, which is finally multiplied element-wise with the original feature maps to produce weighted features.

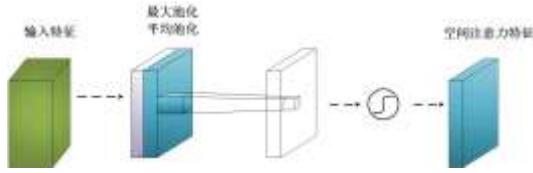


Figure 4: SAM Attention Module Structure Diagram

4. Experiments and Analysis of Results

4.1 Experimental Environment and Parameter Settings

The experimental environment consists of a 64-bit Windows 11 operating system, a 12th-generation Intel Core i7 processor, an NVIDIA GeForce RTX 4060 8GB GPU, 16GB RAM, Python 3.8, PyCharm, PyTorch 2.0.1, and CUDA 11.8.

4.2 Evaluation Indicators

In the classification experiment of the performance diagram presented in this paper, quantitative metrics are employed to evaluate the performance of the trained model. These metrics include: Accuracy (Acc), Precision (Pre), Recall (Rec), and F1 Score, calculated as follows:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (2)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (3)$$

$$F_1 = 2 \times \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \quad (4)$$

In the formula: TP denotes the number of labeled samples correctly detected by the model; FP denotes the number of unlabeled samples falsely classified as labeled by the model; FN denotes the number of labeled samples not detected by the model, i.e., the number of missed detections; TN denotes the number of unlabeled samples correctly identified as such by the model.

4.3 Analysis of Results

To comprehensively evaluate the overall performance of the models, this paper selected six models for comparative experiments: AlexNet[6], GoogleNet[7], ShuffleNetV2[8], MobileNetV2, MobileNetV3, and the improved MobileNetV3_CBAM. By comparing the recognition accuracy and computational efficiency of each model, the performance of the improved model can be intuitively demonstrated.

The experimental results are shown in Table 2. The improved MobileNetV3_CBAM model achieves the best performance across all metrics, including accuracy, precision, recall, and F1 score. Specifically, the model achieves a precision of 95.4%, representing improvements of 9.1%, 6.9%, 7.4%, 4.2%, and 2.1% over AlexNet, GoogleNet, ShuffleNetV2, MobileNetV2, and MobileNetV3, respectively. The results demonstrate that incorporating the CBAM attention mechanism effectively enhances the model's feature extraction capabilities and generalization performance, thereby significantly improving the accuracy and stability of detonator diagram recognition.

Network Model	Accur acy Rate (%)	Precisi on (%)	Rec all Rate (%)	F1 Sco re (%)
AlexNet	84.0	86.3	85.1	85.7
GoogleNet	86.2	88.5	87.3	87.9
ShuffleNetV2	87.1	88.0	86.5	87.2
MobileNetV2	91.6	91.2	90.9	91.0
MobileNetV3	93.5	93.3	92.8	93.0
MobileNetV3_ CBAM	95.1	95.4	95.0	95.2

Table 2 Comparison of Model Results

5. Conclusion

This paper proposes an intelligent identification method for gauge diagrams based on an improved MobileNetV3 architecture. By incorporating the CBAM attention mechanism to replace the original SE module within the MobileNetV3 network, the enhanced MobileNetV3_CBAM model achieves optimal performance across key metrics including accuracy, precision, recall, and F1 score. Specifically, it achieves an accuracy of 95.1%, precision at 95.4%, recall at 95%, and an F1 score of 95.2%. This enhanced model provides an efficient and feasible deep learning solution for the automatic recognition and fault diagnosis of pumping unit gauging charts, demonstrating practical value in advancing the intelligent monitoring of oilfield equipment.

6. References

- [1] Huang Lihua, Shang Yufan, Gao Guogang, et al. Intelligent Monitoring System for Pumping Units Based on Alibaba Cloud IoT [J]. Machinery Manufacturing and Automation, 2025, 54(04): 294-297. DOI:10.19344/j.cnki.issn1671-5276.2025.04.053.
- [2] Rawat W, Wang Z. Deep convolutional neural networks for image classification: A comprehensive review[J]. Neural computation, 2017, 29(9): 2352-2449.
- [3] Howard A, Sandler M, Chu G, et al. Searching for mobilenetv3[C]//Proceedings of the IEEE/CVF international conference on computer vision. 2019: 1314-1324.
- [4] Howard A G, Zhu M, Chen B, et al. Mobilenets: Efficient convolutional neural networks for mobile vision applications[J]. arXiv preprint arXiv:1704.04861, 2017.
- [5] Liu Jianwei, Liu Junwen, Luo Xionglin. Research Progress on Attention Mechanisms in Deep Learning [J]. Journal of Engineering Sciences, 2021, 43(11): 1499-1511. DOI: 10.13374/j.issn2095-9389.2021.01.30.005.
- [6] Duan, Y. X., Li, Y., Sun, Q. F., et al. Improved AlexNet Model and Its Application in Well Performance Diagram Classification. Computer Applications and Software, 2018, 35(07): 226-230+272.
- [7] Swapna M, Sharma Y K, Prasad B M G. CNN Architectures: AlexNet, LeNet, VGG, GoogleNet, ResNet [J]. International Journal of Engineering Science, 2020, 8(6): 953-960.
- [8] Ma N, Zhang X, Zheng H T, et al. Shufflenet v2: Practical guidelines for efficient cnn architecture design[C]//Proceedings of the European conference on computer vision (ECCV). 2018: 116-131.