

WENCONG XIAO<sup>1</sup>, GAIPIN CAI<sup>2</sup>

## Ball mill load identification method based on IRF-Net with multi-signal time-frequency images

### Introduction

A ball mill is key equipment in the ore dressing process (Bortnowski et al. 2021; Yang et al. 2024), which improves the recovery and yield of useful minerals by grinding the materials, but the grinding operation often consumes a large amount of energy (Gupta et al. 2020), so it is important to research the load identification method of the mill to make the ball mill always maintain in the normal state for the improvement of grinding energy efficiency.

Current methods of mill load identification primarily utilize various indirect detection techniques. Wang et al. (2021) used an integrated empirical modal decomposition method to decompose the grinding signals, reconstructed the grinding sound signals by suitable IMF components and extracted the box fractal dimension features, and proposed an optimized

✉ Corresponding Author: Gaipin Cai; e-mail: 3359973059@qq.com

<sup>1</sup> School of Mechanical and Electrical Engineering, Jiangxi University of Science and Technology, China; ORCID iD: 0009-0004-1609-2971; e-mail: 2468298330@qq.com

<sup>2</sup> School of Mechanical and Electrical Engineering, Jiangxi University of Science and Technology; Jiangxi Mechanical and Electrical Engineering Technology Research Center of Mining and Metallurgy, China; ORCID iD: 0009-0006-8880-6616; e-mail: 3359973059@qq.com



© 2025. The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution-ShareAlike International License (CC BY-SA 4.0, <http://creativecommons.org/licenses/by-sa/4.0/>), which permits use, distribution, and reproduction in any medium, provided that the Article is properly cited.

limit learning machine method to effectively identify the mill loads. Cai et al. (2019) proposed a method based on improved empirical wavelet transform, multi-scale fuzzy entropy, and adaptive evolutionary particle swarm optimization probabilistic neural network classification for ball mill load identification. Huang et al. (2020) proposed improved integrated multi-classifier modeling, which sequentially fuses the recognition results of multiple classifiers and multiple sensors to be applied to measure ball mill loads.

The above machine learning method using one-dimensional signals can realize the recognition of ball mill loads, but it is limited by the manual extraction of features and also fails to consider the influence of time domain and frequency domain information at the same time. With the development of deep learning methods, neural networks can automatically learn features from data without the need for manual extraction (Hoang et al. 2020; Shi et al. 2020). Xu et al. (2022) demonstrated the time series of ball mill acceleration vibration signals in the form of two-dimensional images, fed into a VGG19 network model to achieve end-to-end classification of operating conditions. Meanwhile, converting a one-dimensional signal into a time-frequency graph analysis allows for the simultaneous acquisition of features in both the time and frequency domains. Shao et al. (2019) proposed a deep neural network for fault diagnosis of induction, bearings, and gearboxes by converting raw data into images to obtain time-frequency distributions through wavelet transform. Wang et al. (2019) the raw acceleration signals were converted into time-frequency images using eight methods, such as STFT and instantaneous frequency, and trained the model with a neural network structure modified by AlexNet (Krizhevsky et al. 2012). Kong et al. (2023) proposed a deep neural network based on the Mel spectrogram of grinding sound signals as input and a Resnet network (He et al. 2016) to recognize mill load.

The model's accuracy can be improved due to the fusion of multiple signals to provide different information (Ali et al. 2019). In this paper, by combining CVR and MBVR, which are two load parameters that affect the operating state of ball mills, and considering to reduce the complexity of model training, we propose an improved residual fusion network for ball mill load identification. By comparing with mainstream networks (VGG (Xu et al. 2022), CNN (Alzubaidi et al. 2021), AlexNet (Krizhevsky et al. 2012), GoogLeNet (Ak et al. 2022) and ResNet (He et al. 2016)), we conduct experiments, along with ablation experiments, to verify the superiority of the method in this paper. The main contributions of this paper are as follows.

1. In this paper, the ball mill cylinder vibration signal and acoustic signal in the grinding process are collected simultaneously, and the one-dimensional raw signal is converted into a two-dimensional time-frequency map for fusion training.
2. This study proposes an improved residual fusion network method for identifying load states with a multi-signal merging approach under different CVR and different MBVR. The improved residual fusion network obtains higher accuracy than individual signals and other multi-signal approaches.
3. A comparison was made with the proposed method through classical deep learning methods and ablation experiments to validate the better results of the proposed model.

## 1. Related work

### 1.1. Continuous wavelet transform

Continuous Wavelet Transform (CWT) (Zhang et al.2021) is a time-frequency localized analysis method in which both time and frequency windows can be varied. This property makes the CWT adaptive to the signal, overcoming the disadvantage of the STFT, which cannot be localized in the time and frequency domains.

This paper focuses on the ball mill's cylinder vibration signal and acoustic signal processing, using CWT for time-frequency analysis. The key of CWT lies in the selection of the wavelet basis function.

For an arbitrary space, the CWT of a function  $f(t)$  is defined as:

$$CWT(a, T) = \frac{1}{\sqrt{a}} \int f(t) \varphi\left(\frac{t-T}{a}\right) dt \quad (1)$$

$$\varphi_{a,T} = \frac{1}{\sqrt{a}} \varphi\left(\frac{t-T}{a}\right) dt, \quad a, T \in \mathbb{R}, \quad a \geq 0 \quad (2)$$

- $a$  – the scale factor,
- $T$  – the translation factor,
- $\varphi_{a,T}$  – the wavelet basis function.

Since Morlet wavelet has a good balance between time and frequency localization, this paper proposes to use Morlet wavelet, the specific wavelet basis function is “cmor3-3”.

### 1.2. Residual networks

An increase in the number of neural network layers can make the network difficult to train, affect the training accuracy, and lead to network degradation. To solve this problem, He et al. (2016) proposed a residual network (ResNet) structure, which utilizes ResNet using shortcut connections to achieve a stacked composition of multiple residual modules with constant mapping of the network layers.

The residual unit focuses on the neural network locally by fitting a residual mapping: as in Figure 1, assuming  $X$  is the original input and  $H(X)$  is the desired ideal mapping, the residual unit is needed to fit the residual mapping  $F(X)$ , a mapping that is often easier to optimize in reality. In the residual cell, inputs can be propagated forward faster through cross-layer data lines. It effectively reduces the mapping learning difficulty and speeds up the convergence of the model.

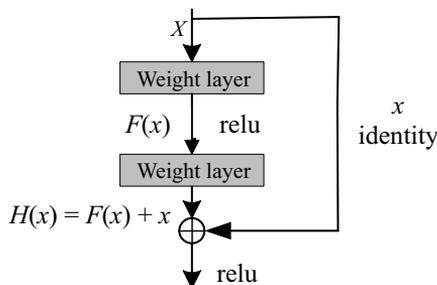


Fig. 1. Structure of residual unit

Rys. 1. Struktura jednostki rezydualnej

### 1.3. Depthwise separable convolution

Depthwise separable convolution (DSC) (Liu et al. 2022) utilizes the reduction of computational complexity and number of parameters to improve the performance of neural networks. Firstly, through the depthwise convolution operation, the convolution operation of the single-channel convolution kernel is carried out on each channel of the input samples to obtain the same number of feature maps as the number of input channels to significantly reduce the parameter size and the amount of operations in the network; secondly, the pointwise convolution operation utilizes the conventional convolution kernel size of  $1 \times 1$  to combine the information of different channels at the same pixel position and generate the final required feature maps to achieve the purpose of combining the information of different channels. Secondly, in pointwise convolution, a conventional convolution operation with a kernel size of  $1 \times 1$  is used to combine the information of different channels at the same pixel position and generate the final required feature map to realize the purpose of combining the information of different channels.

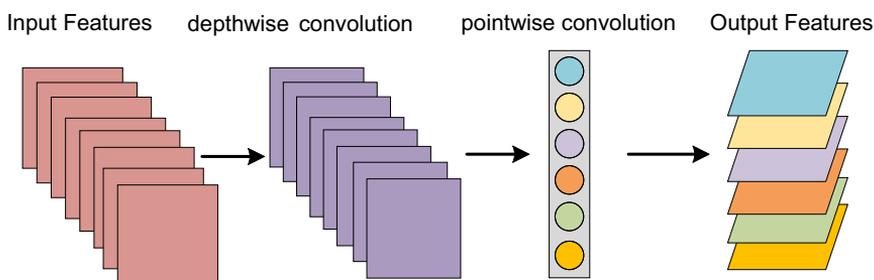


Fig. 2. Depthwise separable convolutional structure

Rys. 2. Rozdzielna struktura spłotowa

At feature input, are used as the height and width of the map, the number of channels is  $M$ , and the size of the convolution kernel is  $K$ . At output feature, the number of channels is  $N$ , and the height and width of the feature map after deep convolution, and the parametric quantities of the ordinary convolution and the DSC as  $C_1 = H \times W \times M \times K \times K \times N$ ,  $C_2 = H \times W \times M \times K \times K \times M \times N$ , respectively, the resulting ratio is:

$$\frac{C_2}{C_1} = \frac{H \cdot W \cdot M \cdot (K^2 + N)}{H \cdot W \cdot M \cdot N \cdot K \cdot K} = \frac{1}{N} + \frac{1}{K^2} \quad (3)$$

Equation (3) shows that compared with the standard convolution, DSC can reduce computational complexity and the number of required parameters by splitting the convolution into depthwise convolution and pointwise convolution. This alleviates the overfitting phenomenon and improves the model generalization ability.

#### 1.4. Hardswish activation function

The addition of Hardswish activation function can enhance the network nonlinear ability, which improves the expressive ability of the neural network and effectively alleviates the problem of neuronal necrosis of ReLU activation function (Jin et al. 2023). It has the advantage of good numerical stability and fast computational speed. It can also extract the main information of the maximum pooling layer in the image while reducing the number of operations and can be used as a segmentation function to reduce the number of memory accesses.

Therefore, in order to improve the ability of ball mill load identification, after improving the residual network by using DSC, the Hardswish activation function is used to replace the ReLU activation function in the network and combined with Dropout to further improve the robustness and generalization ability of the model. The Hardswish activation function  $f(x)$  is:

$$f(x) = x \frac{\min(\max(0, x + 3), 6)}{6} \quad (4)$$

$f(x)$  generates a nonlinear feature mapping by truncating and normalizing the input  $x$  to  $1/6$  in the range  $[0, 6]$  after shifting the input  $x$  by 3, thus filtering the input data  $x$  and effectively controlling its transmission from the bottom to the top.

## 2. Improvement of mill load identification by residual network

This section first describes the acquisition and conversion of raw data during mill operation. In order to compare the network performance under single and multiple signals,

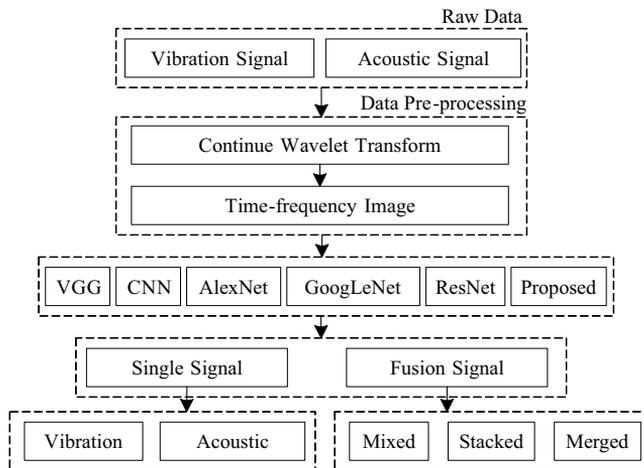


Fig. 3. Flowchart for comparison of load recognition methods

Rys. 3. Schemat blokowy dla porównania metod rozpoznawania obciążenia

the improved residual network under different signal input modes is established, and the feasibility of the improved residual fusion network under the proposed merging mode is subsequently verified in comparison with other classical network models. The processing flow of the signals in this paper is shown in Figure 3, where cylinder vibration and acoustic signals are collected simultaneously during ball mill operation, and the raw time-domain signals are converted into time-frequency images by CWT, and then VGG, CNN, AlexNet, GoogLeNet, ResNet, and the improved residual network model in this paper are trained using single, and multiple signals and the results are evaluated.

### 2.1. Signal Acquisition and Conversion Process of Experimental Ball Mill

Using an experimental ball mill with a motor power of 0.75 kW, the operating conditions of the ball mill were classified into nine categories based on nine operating conditions combining low, medium, and high CVR and low, medium, and high MBVR. Acceleration

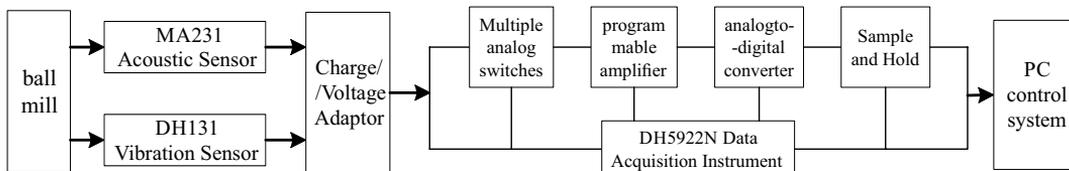


Fig. 4. Signal Acquisition Flow

Rys. 4. Przepływ akwizycji sygnału

sensors were utilized to collect mill cylinder vibration signals, and sound sensors were used to collect acoustic signals at the same time, with a sampling rate of 20,000 Hz for both sensors. The complete acquisition process is shown in Figure 4. The raw time domain signal was composed of one sample per 2048 data points with a sliding window step of 1,500, and the raw signal was converted into a time-frequency image using CWT. Each category was randomly divided into 80% for training data and 20% for test data.

## 2.2. Training models

In order to compare the results using single and multiple signals, different improved residual networks are proposed by improving the ResNet18 model through 3 aspects: DSC, dropout, and Hardswish activation function. In the single-signal approach, the network is trained separately and independently using mill cylinder vibration signals and acoustic signals, the structure of which is shown in Figure 5. First, a single-signal time-frequency image is an input, and features are initially extracted using a depth-separable convolutional layer, which significantly improves computational efficiency, reduces the number of parameters, and enhances the stability and performance of the model through a combination of batch normalization, depth-separable convolutional kernel, batch normalization, and Hardswish activation function while maintaining effective feature extraction. Next, the preliminary extracted features are sequentially fed into the improved residual blocks IRBB1 and IRBB2 to achieve further efficient and stable feature extraction, and the feature maps are converted into one-dimensional vectors through the global average pooling layer. To prevent overfitting of the network, a dropout layer is added after the GAP in which 20% of the neurons are randomly closed. Finally, the load identification results are output through a fully connected layer with a softmax activation function.

For the multi-signal approach using both vibration and acoustic signals, different fusion methods are proposed to train the improved residual network, and the model results are shown in Figure 5. The network structure of the first two methods is similar to the single-signal model, with the difference being that the inputs to the model are different. In Figure 6(a), the input to the network is a time-frequency map mixing both vibrational and acoustic signal data. In contrast, in Figure 6(b), the input to the network is a stacked input of the two types of time-frequency map data, i.e., the color channels are stacked together,

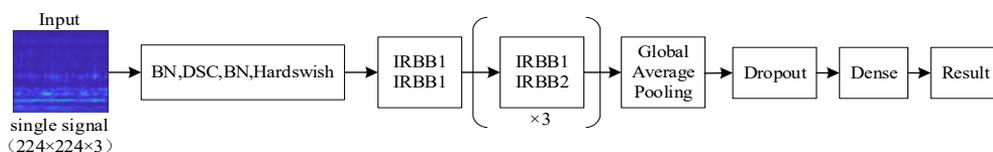


Fig. 5. Single-signal model training

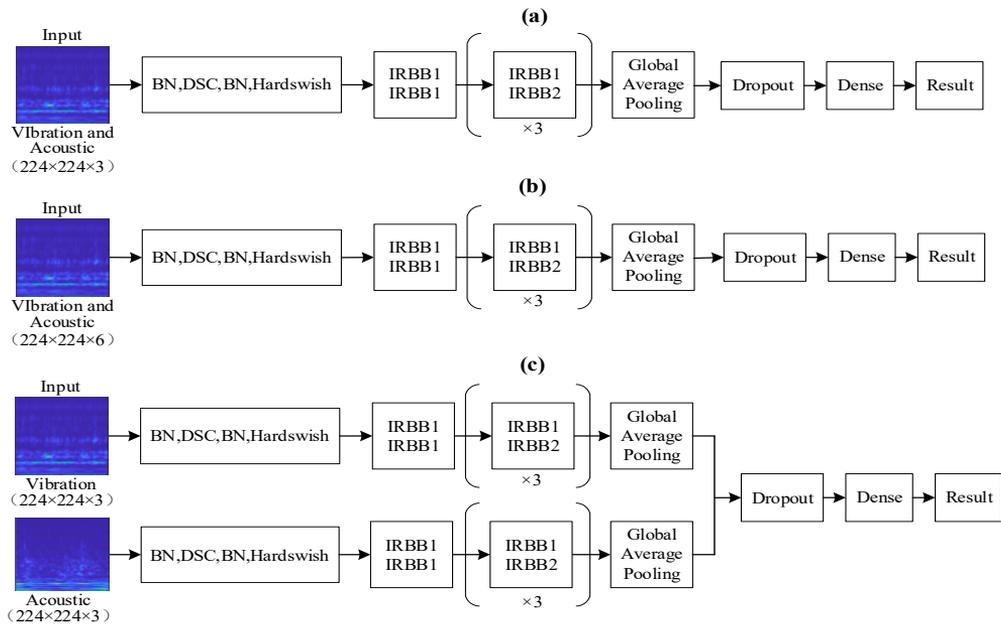


Fig. 6. Multi-signal model training  
(a) mixed signals, (b) stacked signals, (c) merged signals

Rys. 6. Trening modelu wielosygnałowego  
(a) sygnały mieszane, (b) sygnały skumulowane, (c) sygnały połączone

allowing the third channel to be expanded from three to six. Unlike the mixed and stacked methods, the merged signal approach involves feeding the vibrational and acoustic signals into separate branches for model training. As shown in Figure 6(c), the depth-separable convolutional layer and the improved residual block extract the features of the two types of time–frequency maps, respectively, and merge the two networks after the GAP layer in order to train the remaining neurons and output the recognition results.

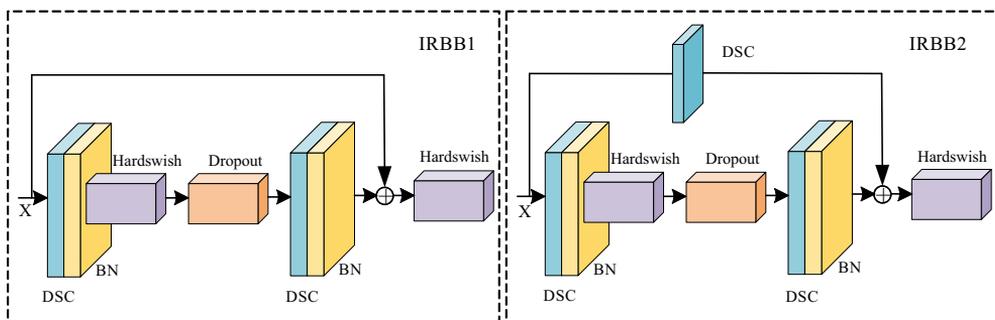


Fig. 7. IRBB structure

Rys. 7. Struktura IRBB

The structural composition of IRBB1 and IRBB2 is shown in Figure 7. The two improved residual modules have the same main path. They are composed of DSC, bulk normalization, and Hardswish activation function modules, as well as a Dropout layer, where the DSC and Hardswish activation functions are used to improve the computational efficiency and model performance. In contrast, the Dropout layer improves the network's generalization performance. The difference between the modules is that IRBB1 contains two DSCs and connects the inputs to the outputs directly via jump connections, whereas IRBB2 contains three DSCs, and its jump connections have to go through one DSC before adding the outputs.

### 3. Experimentation and analysis

This section discusses the experimental results of mill load identification using vibration and acoustic signals. First, the experimental ball mill's cylinder vibration and acoustic signals during the operation were collected and saved. The saved one-dimensional raw signal data were converted into two-dimensional time-frequency diagrams using CWT, and the results were compared for different load categories and signals. Second, different improved residual networks for single and multi-signal as well as classical classification neural network structures (VGG, CNN, GoogLeNet, AlexNet, and ResNet) are proposed to compare the accuracy of the proposed models. Finally, ablation experiments are conducted on the improved module of the proposed model, and the effectiveness of the recognition method is further measured by four metrics: accuracy, precision, recall, and F1 score.

#### 3.1. Ball Mill Variable Load Experiment

In this paper, the experimental ball mill is selected as the experimental object, and CVR and MBVR are used as variables to carry out the variable-load multi-condition experiments. The M12 ball bearing is fixed under the ball mill barrel, and its contact with the outside



Fig. 8. Experimental setup

Rys. 8. Konfiguracja eksperymentalna

of the ball mill barrel for rolling friction to conduct the vibration signals generated by the impact parts of the bottom of the ball mill, and the linear vibration sensor is magnetically absorbed under the ball bearing to receive the vibration signals of the barrel, and the MA231 sound sensor is arranged in a fixture at a distance of 30–50 cm from the ball mill. The installation of the experimental ball mill and ball bearing is shown in the Figure 8. Grinding media selection of  $\varphi$  30 mm,  $\varphi$  40 mm,  $\varphi$  50 mm, size of the steel ball, and the use of tungsten

Table 1. Parameters of ball mill variable load experiment

Tabela 1. Parametry eksperymentu ze zmiennym obciążeniem młyna kulowego

| CVR | MBVR | Number of steel balls |                 |                 | Material with proportion 1:2:3:4 (kg) |         |          |           |
|-----|------|-----------------------|-----------------|-----------------|---------------------------------------|---------|----------|-----------|
|     |      | $\varphi$ 30 mm       | $\varphi$ 40 mm | $\varphi$ 50 mm | +3–6 mm                               | +6–9 mm | +9–13 mm | +13–18 mm |
| 0.2 | 0.1  | 27                    | 23              | 17              | 0.178                                 | 0.356   | 0.534    | 0.712     |
|     | 0.3  | 19                    | 16              | 12              | 0.369                                 | 0.737   | 1.106    | 1.475     |
|     | 0.5  | 15                    | 12              | 9               | 0.482                                 | 0.964   | 1.445    | 1.927     |
|     | 0.7  | 12                    | 10              | 8               | 0.571                                 | 1.143   | 1.714    | 2.285     |
|     | 0.9  | 11                    | 9               | 7               | 0.645                                 | 1.291   | 1.936    | 2.582     |
|     | 1.1  | 10                    | 8               | 6               | 0.711                                 | 1.421   | 2.132    | 2.842     |
| 0.3 | 0.1  | 38                    | 32              | 25              | 0.252                                 | 0.504   | 0.756    | 1.008     |
|     | 0.3  | 26                    | 22              | 17              | 0.512                                 | 1.024   | 1.536    | 2.048     |
|     | 0.5  | 20                    | 17              | 13              | 0.645                                 | 1.291   | 1.936    | 2.581     |
|     | 0.7  | 16                    | 13              | 10              | 0.726                                 | 1.453   | 2.179    | 2.905     |
|     | 0.9  | 13                    | 11              | 9               | 0.781                                 | 1.561   | 2.342    | 3.123     |
| 0.4 | 0.1  | 51                    | 43              | 33              | 0.336                                 | 0.672   | 1.008    | 1.344     |
|     | 0.3  | 34                    | 29              | 22              | 0.683                                 | 1.365   | 2.048    | 2.731     |
|     | 0.5  | 26                    | 22              | 17              | 0.860                                 | 1.721   | 2.581    | 3.442     |
|     | 0.7  | 21                    | 18              | 14              | 0.968                                 | 1.937   | 2.905    | 3.874     |
|     | 0.9  | 18                    | 15              | 11              | 1.041                                 | 2.082   | 3.123    | 4.164     |
| 0.5 | 0.1  | 64                    | 54              | 41              | 0.420                                 | 0.840   | 1.260    | 1.680     |
|     | 0.3  | 43                    | 36              | 28              | 0.853                                 | 1.707   | 2.560    | 3.414     |
|     | 0.5  | 33                    | 28              | 21              | 1.076                                 | 2.151   | 3.227    | 4.302     |
|     | 0.7  | 26                    | 22              | 17              | 1.211                                 | 2.421   | 3.632    | 4.842     |
|     | 0.9  | 22                    | 19              | 14              | 1.301                                 | 2.602   | 3.904    | 5.205     |
|     | 1.1  | 19                    | 16              | 12              | 1.366                                 | 2.733   | 4.099    | 5.466     |

ore with a density of 2,500 kg/m<sup>3</sup> as the experimental ore material, before the experiment will be fed into the particle size screening for four particle size classes: +3–6 mm, +6–9 mm, +9–13 mm, and +13–18 mm.

As shown in Table 1, the experiments were set up with CVR of 0.2, 0.3, 0.4, 0.5, and MBVR of 0.1, 0.3, 0.5, 0.7, 0.9, and 1.1 in a total of 22 sets of experiments, which included nine working conditions with combinations of low, medium, and high CVR and low, medium, and high MBVR. Ball mill CVR in [0.3, 0.4] within the normal, less than 0.3 is too low, easy to empty smash, easy to damage the cylinder in the long run, more significant than 0.4 is too high, the impact of insufficient. MBVR in [0.5, 0.7] when regular, less than 0.5 too low, too few steel balls, the ore is difficult to grind, greater than 0.7 is too high, too many steel balls will lead to the impact of the drop movement being too large, so that the material will reach a much smaller particle size than the target particle size. The impact on the ball mill body is greatest when the CVR is high and the MBVR is low. When the MBVR of 0.1 and 1.1 for the extreme conditions of extreme ball and material, respectively. Set the CVR of low for A1, normal for A2, too high for A3, the same MBVR of low in high set for B1, B2, B3, 9 kinds of conditions were A1B1, A1B2, ..., A3B2, A3B3. In each set of experiments, the cylinder vibration and acoustic signals were collected and saved using the instrument while the ball mill was in operation.

### 3.2. Time-frequency image generation

The raw vibration and acoustic signals acquired are converted into time-frequency images by CWT, and each two-dimensional map after conversion shows the results for one second, and the pixels are all 224\*224. As shown in Figure 9, the vibration signal has continuous fluctuations at multiple frequencies from 0 to 10,000 Hz when the CVR is 50% and the MBVR is 0.9. In comparison, the acoustic signal only varies from 0 to 5,000 Hz, which is a more noticeable difference. However, not all the differences between all categories of time-frequency diagram features can be observed by the naked eye, for example, in the case of the three categories of vibration signals at different MBVR with a CVR of 50% the differences are not obvious on the image display. Therefore, we analyze the images using the feature extraction capability of the improved residual network and use these image data for subsequent model performance evaluation.

### 3.3. Identification results

Single-signal and multi-signal fusion methods are used to classify and recognize the features processed by DWT. To validate the performance of the proposed method, the models under single-signal and multi-signal fusion are trained using the mainstream networks VGG, CNN, GoogLeNet, AlexNet, and ResNet. The resulting accuracies are compared with the

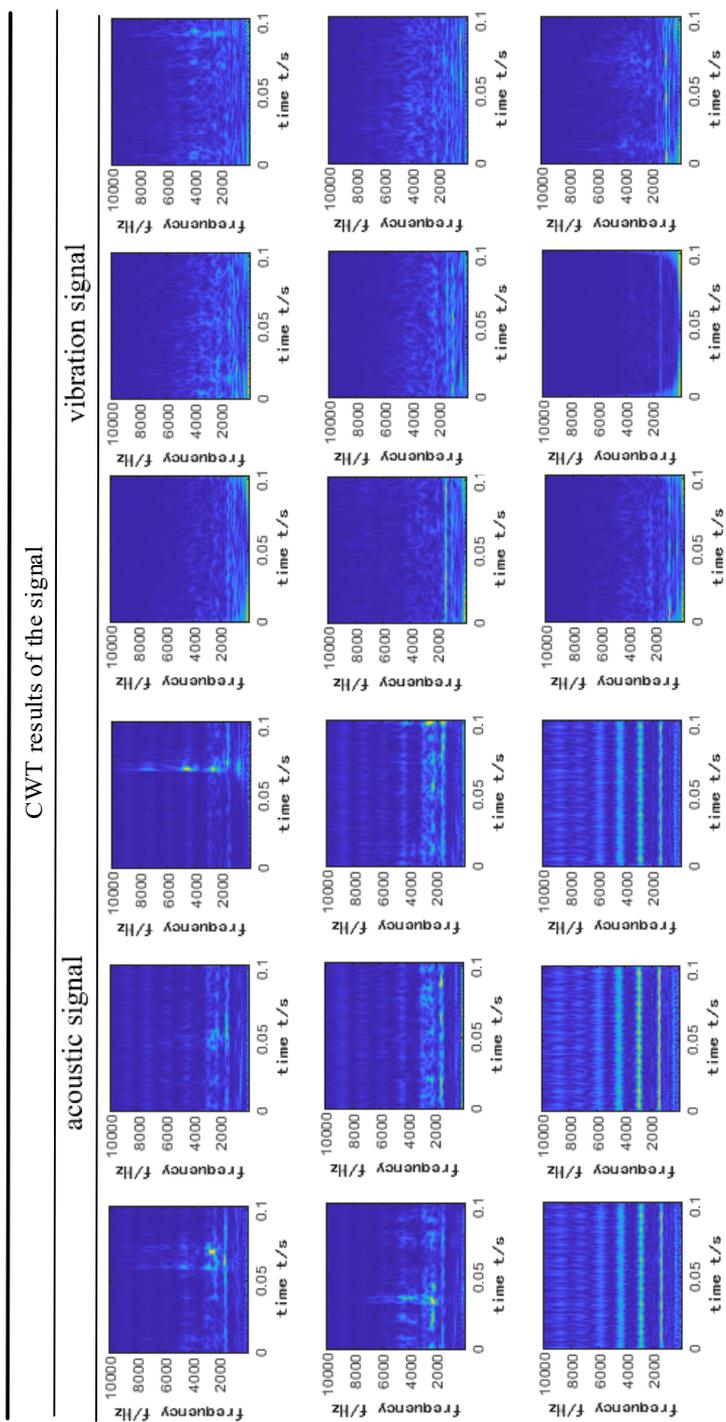


Fig. 9. Pre-processing results of vibration and acoustic data for different classifications

Rys. 9. Wyniki wstępnego przetwarzania danych wibracyjnych i akustycznych dla różnych klasyfikacji

results of the improved residual network, and the average recognition accuracy is chosen as the final metric term for evaluation to obtain the best results.

The results obtained by the different methods in the single-signal model training are shown in Figure 10. In the figure, VGG got poor results in two sets of experiments using vibration signal and acoustic signal datasets as training sets, CNN, GoogLeNet and AlexNet recognition accuracies reached more than 84%. In contrast, ResNet and the improved residual network of this paper reached more than 90% in both sets of experiments, with the improved network of this paper having better results. The results of the single-signal model training show that the improved residual network model can improve the accuracy of ball mill load recognition when trained on the experimental dataset.

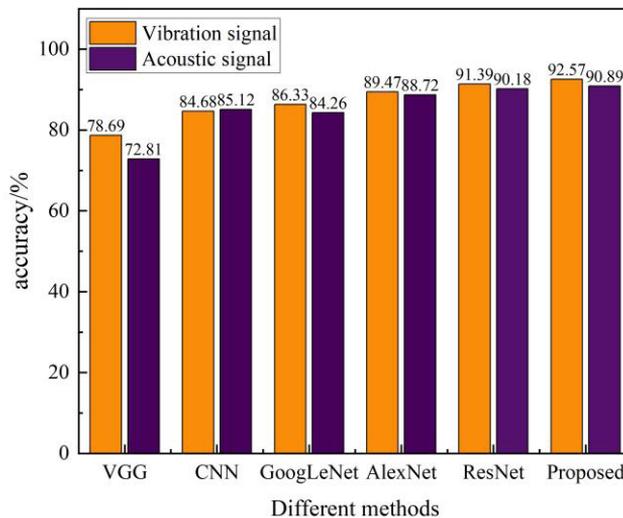


Fig. 10. Accuracy of single-signal model training

Rys. 10. Dokładność treningu modelu pojedynczego sygnału

In order to obtain higher load recognition accuracy, a combination of signal fusion and improved residual network structure (mixed, stacked, and merged methods) is considered and compared with the single signal approach, and the results obtained are shown in Table 2. In a single method, better results were obtained by improving the residual network. For fused signals, if two signals are fused while using the traditional ResNet network, the accuracy can be higher than the single signal using the stacked and merged methods. However, using the improved residual network, it can be found that it exhibits higher accuracy on both fusion methods, both showing better results than the single signal and the mixed signals only, with the merged ones showing the best results. A comparison of the results shows that the proposed improved residual fusion network works better than other methods, with

Table 2. Accuracy results for different methods (%)

Tabela 2. Wyniki dokładności dla różnych metod (%)

|                  | VGG   | CNN   | AlexNet | GoogLeNet | ResNet | Proposed |
|------------------|-------|-------|---------|-----------|--------|----------|
| Single vibration | 78.69 | 84.68 | 86.33   | 89.47     | 91.39  | 92.57    |
| Single acoustic  | 72.81 | 85.12 | 84.26   | 88.72     | 90.18  | 90.89    |
| Mixed            | 70.23 | 83.18 | 86.51   | 83.89     | 87.24  | 91.34    |
| Stacked          | 83.57 | 90.25 | 89.81   | 82.47     | 92.46  | 95.18    |
| Merged           | 85.34 | 90.74 | 92.46   | 91.79     | 94.41  | 98.33    |

an average recognition accuracy of up to 98.33%. In contrast, the best results of VGG, CNN, AlexNet, GoogLeNet, and ResNet are 85.34%, 90.74%, 92.46%, 91.79%, and 94.41%, respectively. It can be seen that the learned multi-source information feature vector using the improved residual fusion network retains the most compelling features, and the proposed method is useful for recognition performance.

### 3.4. Ablation experiments

In order to further explore the efficiency of the IRF-Net, six sets of ablation experiments are designed to analyze the effects of DSC, dropout, and Hardswish activation functions on the mill load identification results. The obtained experimental results are shown in Table 3, where “√” means this item is used and “×” means this item is not used.

Experiment 1 represents the conventional Resnet18 network, which has an accuracy of 94.41% in recognizing mill loads. To verify the effectiveness of DSC, experiment 2 replaces the ordinary convolution in the original Resnet18 model with DSC, and the resulting accuracy

Table 3. Results of ablation experiments

Tabela 3. Wyniki eksperymentów z ablacją

| Model                        | Experiment 1 | Experiment 2 | Experiment 3 | Experiment 4 | Experiment 5 | Experiment 6 |
|------------------------------|--------------|--------------|--------------|--------------|--------------|--------------|
| DSC                          | ×            | √            | ×            | ×            | ×            | √            |
| Dropout                      | ×            | ×            | √            | ×            | √            | √            |
| Hardswish                    | ×            | ×            | ×            | √            | √            | √            |
| accuracy/%                   | 94.41        | 94.75        | 95.68        | 95.18        | 97.03        | 98.33        |
| Duration of training round/s | 28           | 19           | 31           | 33           | 34           | 26           |

is 94.75%, an improvement of 0.34% over the unimproved model. While experiment 3 adds only the Dropout layer to the original model, the recognition accuracy can be increased to 95.68%, and the model’s generalization ability is significantly improved. In investigating the contribution of the Hardswish activation function to the model, experiment 4 replaced all the ReLU activation functions of the original network with the Hardswish activation function, which resulted in an increase in the model identification accuracy of 0.77 percentage points because it produces a strong regularization effect and is continuously derivable everywhere. The above shows that improvements to all three components individually are beneficial in improving the recognition accuracy of the load. Therefore, experiment 5 considers the combination of Dropout and Hardswish activation functions, retaining the ordinary convolution, and the accuracy obtained from recognition is improved to 97.03%; finally, based on experiment 5, the introduction of depthwise separable convolution yields a the recognition accuracy of 98.33% and the time taken for the training round is 26 s, which proves that the combination of DSC, Dropout, and Hardswish activation function is effective and reliable for improving the residual fusion network in improving the recognition of ball mill loads.

In order to observe more concretely the recognition effect of the merging method in this paper on each class of loads, the confusion diagrams obtained on the test set are shown in Figure 11. The states of six of these loads were accurately identified. A small

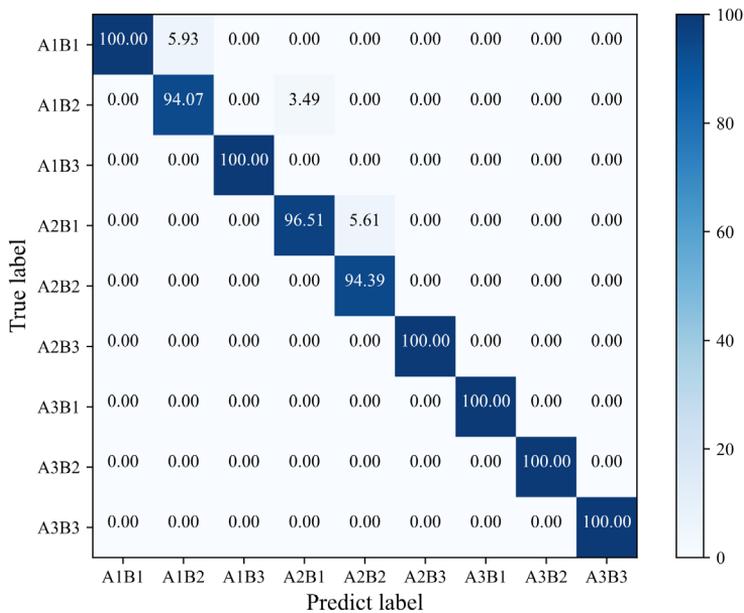


Fig. 11. Test set recall confusion matrix

Rys. 11. Macierz pomyłek przywoływania zestawu testowego

number of samples were incorrectly predicted, so in the final load identification results, the identification accuracy of this paper's method did not reach 100%, which is consistent with the results in the table above. However, considering that this confusion matrix can only show the statistical results on the recall rate of each category, which may not be sufficient to evaluate the performance of the improved residual fusion network model, four metrics were used for the measurement, as shown in Table 4.

Table 4. Indicators related to the consolidated methodology

Tabela 4. Wskaźniki związane ze skonsolidowaną metodologią

|           | A1B1  | A1B2  | A1B3 | A2B1  | A2B2  | A2B3 | A3B1 | A3B2 | A3B3 | Average |
|-----------|-------|-------|------|-------|-------|------|------|------|------|---------|
| Accuracy  | 100   | 94.06 | 100  | 96.51 | 94.39 | 100  | 100  | 100  | 100  | 98.33   |
| Precision | 99.21 | 93.65 | 100  | 97.22 | 95.01 | 100  | 100  | 100  | 100  | 98.34   |
| Recall    | 100   | 94.07 | 100  | 96.51 | 94.39 | 100  | 100  | 100  | 100  | 98.33   |
| F1-score  | 99.58 | 94.13 | 100  | 96.91 | 94.32 | 100  | 100  | 100  | 100  | 98.33   |

As can be seen from the relevant indicators of the merging method, the results of the four indicators are balanced in each category, and the average value of all of them reaches 98.33%, which means that the residual fusion network can effectively identify each state of the ball mill load.

## Conclusions

This study proposes an improved residual fusion network method for ball mill load identification under multiple signals. Firstly, the vibration signals and acoustic signals during the operation of ball mills with different CVR and MBVR are collected. The time-frequency diagrams of the raw signals are obtained by using CWT, and the model is trained using the IRF-Net based on multi-signal merging. The effectiveness of the method is verified by comparing it with the classical algorithm and the improved residual network with different signal input modes and combining it with ablation experiments. The experimental results show that the proposed method takes 26s for one round of training, and the recognition accuracy reaches 98.33%, which is better than other comparison networks, realizing the reduction of model recognition time and the improvement of recognition accuracy. Meanwhile, the average values of the proposed model in Precision, Recall, and F1-score metrics are 98.34%, 98.33%, and 98.33%, respectively, which verifies that the method of this paper has good robustness in ball mill load identification. Using time-frequency graphs

as network inputs, the model achieves automatic feature extraction in the time domain and frequency, obtaining higher recognition accuracy.

*This research was funded by National Natural Science Foundation of China (52364025).*

*The Authors have no conflict of interest to declare.*

## REFERENCES

- Ak et al. 2022 – Ak, A., Topuz, V. and Midi, I. 2022. Motor imagery EEG signal classification using image processing technique over GoogLeNet deep learning algorithm for controlling the robot manipulator. *Biomedical Signal Processing and Control* 72(6), DOI: 10.1016/j.bspc.2021.103295.
- Ali et al. 2019 – Ali, M.Z., Shabbir, M.N.S.K., Liang, X., Zhang, Y. and Hu, T. 2019. Machine learning-based fault diagnosis for Single- and multi-faults in induction motors using measured stator currents and vibration signals. *IEEE Transactions on Industry Applications* 55(3), pp. 2378–2391, DOI: 10.1109/TIA.2019.2895797.
- Alzubaidi et al. 2021 – Alzubaidi, L., Zhang, J., Humaidi A.J., Ayad Al-Dujaili, A., Duan, Y., Al-Shamma, O., Santamaria, J., Fadhel, M.A., Al-Amidie, M. and aith Farhan, L. 2021. Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. *Journal of big Data* 8, DOI: 10.1186/s40537-021-00444-8.
- Bortnowski, P., Gladysiewicz, L., Król, R. and Ozdoba, M. 2021. Energy Efficiency Analysis of Copper Ore Ball Mill Drive Systems. *Energies* 14(6), DOI: 10.3390/en14061786.
- Cai et al. 2019 – Cai, W.H., Liu, A., Bing, D. and Luo, C. 2019. Load state identification method for ball mills based on improved EWT, multiscale fuzzy entropy and AEPPO\_PNN classification. *Processes* 7(10), DOI: 10.3390/pr7100725.
- Gupta, V.K. 2020. Energy absorption and specific breakage rate of particles under different operating conditions in dry ball milling. *Powder technology* 361, pp. 827–835, DOI: 10.1016/j.powtec.2019.11.033.
- He et al. 2019 – He, K.M., Zhang, X.Y., Ren, S.Q. and Sun, J. 2016. Deep residual learning for image recognition. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 770–778, DOI: 10.1109/CVPR.2016.90.
- Hoang, D.T. and Kang, H.J. 2020. A motor current signal-based bearing fault diagnosis using deep learning and information fusion. *IEEE Transactions on Instrumentation and Measurement* 69(6), pp. 3325–3333, DOI: 10.1109/TIM.2019.2933119.
- Huang et al. 2020 – Huang, P., Sang, G., Miao, Q., Ding, Y. and Jia, M. 2020. Soft measurement of ball mill load based on multi-classifier ensemble modelling and multi-sensor fusion with improved evidence combination. *Measurement Science and Technology* 32(1), DOI: 015105.10.1088/1361-6501/aba885.
- Jin et al. 2024 – Jin, X., Jiang, J., Li, Y. and Wang, Z. 2024. Improved ShuffleNetV2 for Action Recognition in BPPV Treatment. *Biomedical Signal Processing and Control* 88, DOI: 10.1016/j.bspc.2023.105601.
- Kong et al. 2023 – Kong, Y., Wang, X., Zhou, J. and Qin, L. 2023. A Mill Load Identification Method Based on Deep Neural Network. China Automation Congress (CAC). IEEE, pp. 6747–6752, DOI: 10.1109/CAC59555.2023.10452121.
- Krizhevsky et al. 2012 – Krizhevsky, A., Sutskever, I. and Hinton, G.E. 2012. ImageNet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems* 25(2), pp. 84–90, DOI: 10.1145/3065386.
- Liu et al. 2022 – Liu, D., Wang, L., Du, Y., Cong, M. and Li, Y. 2022. 3-D prostate MR and TRUS images detection and segmentation for puncture biopsy. *IEEE Transactions on Instrumentation and Measurement* 71, DOI: 10.1109/TIM.2022.3192292.
- Shao et al. 2019 – Shao, S., McAleer, S., Yan, R. and Baldi, P. 2019. Highly-Accurate Machine Fault Diagnosis Using Deep Transfer Learning. *IEEE Transactions on Industrial Informatics* 15(4), pp. 2446–2455, DOI: 10.1109/TII.2018.2864759.

- Shi et al. 2020 – Shi, Y., Deng, A., Deng, M., Zhu, J., Liu, Y. and Cheng, Q. 2020. Enhanced lightweight multiscale convolutional neural network for rolling bearing fault diagnosis. *IEEE Access* 8, pp. 217723–217734, DOI: 10.1109/ACCESS.2020.3041735.
- Wang et al. 2019 – Wang, J., Mo, Z., Zhang, H. and Miao, Q. 2019. A deep learning method for bearing fault diagnosis based on time-frequency image. *IEEE Access* 7, pp. 42373–42383, DOI: 10.1109/ACCESS.2019.2907131.
- Wang et al. 2021 – Wang, X., Sun, K., Zhang, H., Xiong, W. and Yang, C. 2021. Mill load identification method for ball milling process based on grinding signal. *IFAC-PapersOnLine* 54(21), pp. 7–12, DOI: 10.1016/j.ifacol.2021.12.002.
- Xu et al. 2022 – Xu, H., Wang, T., Zou, W.J., Zhao, J.J., Tao, L. and Zhang, Z.J. 2022. Ball mill load status identification method based on the convolutional neural network, optimized support vector machine model, and intelligent grinding media. *Chinese Journal of Engineering* 44(11), pp. 1821–1831, DOI: 10.13374/j.issn2095-9389.2022.03.06.001.
- Yang, L. and Yang, H. 2024. Load identification method of ball mill based on the CEEMDAN-wavelet threshold-PMMFE. *Gospodarka Surowcami Mineralnymi – Mineral Resources Management* 40(2), pp. 163–180, DOI: 10.24425/gsm.2024.150823.
- Zhang et al. 2021 – Zhang, K., Wang, J., Shi, H., Xiaochen Zhang, X. and Tang, Y. 2021. A fault diagnosis method based on improved convolutional neural network for bearings under variable working conditions. *Measurement* 182, DOI: 10.1016/j.measurement.2021.109749.

#### BALL MILL LOAD IDENTIFICATION METHOD BASED ON IRF-NET WITH MULTI-SIGNAL TIME-FREQUENCY IMAGES

#### Keywords

time-frequency image, residual networks, depthwise separable convolution, mill signals

#### Abstract

Accurately identifying the load status of the ball mill during the grinding process is conducive to improving the overall production efficiency and ensuring the safe operation of the entire grinding process. In this study, ball mill loads were classified into nine categories based on charge volume ratio (CVR) and material-to-ball volume ratio (MBVR). Different sensors are utilized to collect cylinder vibration and acoustic signals in the grinding process, respectively, and the raw data are converted into time-frequency images by continuous wavelet transform. In this paper, the ResNet18 model is improved from three aspects, namely, depthwise separable convolution (DSC), dropout layer, and Hardswish activation function, and an improved residual fusion network (IRF-Net) based on the merging of two time-frequency image signals is proposed for load recognition. In order to validate the performance of the proposed model, time-frequency images of the acquired data are analyzed, single and multiple signals are used as network inputs, respectively, compared with other classical models, and ablation experiments are performed on the different modules of the improvement. The results show that the improved residual fusion network achieves the best results in recognition with an accuracy of 98.33%, demonstrating good load recognition. The IRF-Net-based multi-signal time-frequency diagram identification method can be utilized to make a sound judgment on the load status of the mill.

**METODA IDENTYFIKACJI OBCIĄŻENIA MŁYNA KULOWEGO  
OPARTA NA SIECI IRF-NET Z WIELOSYGNAŁOWYMI OBRAZAMI CZASOWO-CZĘSTOTLIWOŚCIOWYMI**

**Słowa kluczowe**

obraz czasowo-częstotliwościowy, sieci rezydualne,  
splot separowalny głębokościowo, sygnały młyna

**Streszczenie**

Dokładne określenie stanu obciążenia młyna kulowego podczas procesu mielenia sprzyja poprawie ogólnej wydajności produkcji i zapewnia bezpieczną pracę całego procesu mielenia. W tym badaniu obciążenia młyna kulowego zostały sklasyfikowane do dziewięciu kategorii na podstawie stosunku objętości wsadu (CVR) i stosunku objętości materiału do kuli (MBVR). Różne czujniki są wykorzystywane do zbierania drgań cylindra i sygnałów akustycznych w procesie mielenia, odpowiednio, a surowe dane są konwertowane na obrazy czasowo-częstotliwościowe za pomocą ciągłej transformacji falkowej. W tym artykule model ResNet18 został ulepszony pod trzema względami, a mianowicie: poprzez zastosowanie splotu separowalnego głębokościowo (DSC), warstwy dropout i funkcji aktywacji Hardswisha, a ulepszona sieć fuzji resztkowej (IRF-Net) oparta na połączeniu dwóch sygnałów obrazu czasowo-częstotliwościowego jest proponowana do rozpoznawania obciążenia. Aby zweryfikować wydajność proponowanego modelu, analizowane są obrazy czasowo-częstotliwościowe pozyskanych danych, pojedyncze i wielokrotne sygnały są używane jako wejścia sieciowe, odpowiednio, w porównaniu z innymi klasycznymi modelami, a eksperymenty ablacji są przeprowadzane na różnych modułach ulepszenia. Wyniki pokazują, że ulepszona sieć fuzji resztkowej osiąga najlepsze wyniki w rozpoznawaniu z dokładnością 98,33%, co świadczy o dobrym rozpoznawaniu obciążenia. Metodę identyfikacji wielosygnałowego diagramu czasowo-częstotliwościowego opartą na IRF-Net można wykorzystać do rzetelnej oceny stanu obciążenia młyna.

