Ball Mill Grinding Condition Classification Method Based on Triaxial Vibration Spectrograms and Deep Attention Networks

Dianyuan JU [1], Rongfeng ZHANG [3,], Xiaohong WANG [1,2], Zhao LIU [2], Bing HUANG [2], Hongliang YU [2]*

[1] Shandong Provincial Key Laboratory of Preparation and Measurement of Building Materials, University of Jinan, Jinan, China
[2] School of Electrical Engineering, University of Jinan, Jinan, China
[3] School of Chemistry and Chemical Engineering, University of Jinan, Jinan 250022, Shandong, China

202311100023@stu.ujn.edu.cn, {*cse_zhangrf, cse_wxh, cse_liuz, cse_huangb, cse_yuhl}@ujn.edu.cn

Submitted March 27, 2025 / Accepted July 16, 2025 / Online first August 4, 2025

Abstract. *Accurate detection of mill conditions during the cement grinding process directly impacts the quality of particle size distribution and energy consumption per ton of cement. This paper proposes a mill condition classification method based on three-axis wireless vibration sensing and deep feature learning to address issues such as distortion of mill condition characterization caused by power grid disturbances in traditional electrical power methods and sound transmission attenuation in mill sound methods. First, three-axis wireless vibration sensors were installed on the mill shell to collect three-dimensional vibration signals. After filtering and outlier removal, the Fast Fourier Transform (FFT) was applied to generate frequency-domain energy distribution images, creating a vibration spectrum dataset with physical interpretability. Next, a deep dilated separable convolution and multi-head attention fusion network model is proposed. In this model, dilated convolution captures multi-scale frequency domain features using adjustable dilation rates, and the multi-head attention mechanism dynamically adjusts the weight distribution of key frequency bands, enabling adaptive extraction of global frequency domain correlations and local resonance features. Experimental results show that the use of three-dimensional vibration signals to characterize mill conditions improves accuracy by 10% compared to one-dimensional signals. Classification accuracy increased by 6.7% compared to traditional convolutional neural networks, and by 7.6% and 5.5% compared to linear models and machine learning methods, respectively.*

Keywords

cement grinding, mill grinding conditions, vibration signal, frequency domain features

1. Introduction

The detection of mill grinding conditions is a pivotal technology in the cement production process [1]. Given that cement grinding determines the particle size and performance of the final product [2], its quality and efficiency directly impact the structural strength and durability of buildings [3], [4]. Consequently, with the advancement of industrial automation and data-driven technologies in recent years, the detection of mill grinding conditions has garnered extensive research attention aimed at optimizing the cement production process [5], [6].

Mill condition detection methods vary widely to accommodate different production environments and monitoring needs. One common approach involves measuring electrical power to detect grinding conditions. For instance, Wang et al. proposed a machine learning framework that includes a Training Data Generator (TDG) and a Look-back Optimizer (LBO) [7]. This framework identifies working conditions by analyzing fluctuations in current loads and changes in power consumption. The electrical power signal represents the energy transfer in the system, where fluctuations in the power grid voltage directly influence the motor's input power. The nonlinear relationship between changes in the mill's mechanical load and electrical power further exacerbates errors. Another method, proposed by Kalantari et al., involves detecting mill conditions using particle size maps and acoustic signal data [8]. This approach was practically implemented and tested on a ball mill at the Lakan lead and zinc processing plant, yielding promising results. Additionally, acoustic signal analysis is an effective method for detecting grinding conditions. Xu et al. introduced a method that integrates online correlation analysis, Principal Component Analysis (PCA), and adaptive K-means clustering to extract features from acoustic signals [9]. The acoustic signal analysis method extracts internal state information from the

DOI: 10.13164/re.2025.0564

mill's sound spectrum. When sound waves propagate from inside the mill shell to the external receiving device, the material friction sound undergoes exponential attenuation in the air, as described by the formula $A(f) = e^{-afd}$(where a is the attenuation coefficient and d is the propagation distance), resulting in the loss of high-frequency features. Additionally, background noise from sources such as fans and conveyors overlaps with the frequency band of the mill's sound signal, with the signal-to-noise ratio (SNR) typically below 10 dB, leading to environmental noise contamination.

The mill's vibration signal has been recognized as a critical indicator of its operational condition [10], [11]. Previous studies have demonstrated that collecting and processing vibration signals to assess load status provides an effective framework for identifying ball mill conditions [12]. Numerous researchers have analyzed vibration signals to derive features indicative of the mill's operating state [13], [14]. The following sections will discuss the algorithms for extracting and modeling mill vibration signals from two aspects.

Traditional signal feature extraction algorithms have been widely used in the processing of mill vibration signals, but there are still significant limitations on their effectiveness [15], [16]. For this reason, many scholars have improved these algorithms [17]. For instance, Shi et al. first converted the vibration signals to multi-order frequency spectrums (MFS) by fractional Fourier transform (FrFT) to construct a mill load soft sensor model. In another study, Behera et al. obtained vibration signals from accelerometers mounted on the mill shaft and transformed the time-domain signals into frequency-domain signals using the Fast Fourier Transform (FFT) [18]. Tang et al. proposed a method involving the selective integration of multi-source information, which significantly enhanced the generalization ability of the model by utilizing a subset of useful features from the vibration frequency spectrum [19]. However, most existing research, such as sampling from the mill-bearing pedestal, has predominantly analyzed and processed vibration information in a single dimension. This approach fails to capture the richer motion and vibration information of the ball mill in three-dimensional space, thereby limiting the comprehensiveness of the analysis.

Research in modeling algorithms based on vibration signals has been extensively explored by numerous scholars. Tang et al. introduced an updated Kernel Partial Least Squares (KPLS) model [20]. The model parameters and input variables were optimized using a Genetic Algorithm (GA). Although this method demonstrated high accuracy and predictive performance in laboratory-scale mills, its efficacy in the challenging environment of industrial mills remains uncertain. There is another approach involves modeling mill shell vibration signals. This method calculates the Power Spectral Density (PSD) of the vibration signal using FFT and extracts the mass and center frequency of minor peaks in the spectrum as features. For instance, Zhao et al. utilized a Support Vector Machine (SVM) to construct a soft measurement model, optimizing the SVM parameters and input variables via a GA, thereby enhancing the model's predictive performance [21]. Additionally, Liu et al. proposed a new method that can address the poor performance of traditional PCA and spectral feature selection techniques in multi-scale situations [22]. They combined Empirical Mode Decomposition (EMD), PCA, and optimal feature extraction methods to extract, select, and model signals across different frequency scales. Despite the theoretical advantages, this method requires validation in an industrial on-site mill. Also, Huang et al. proposed a deep migration learning-based method for soft measurement of ball mill loads under variable operating conditions [23]. However, deep convolutional networks are prone to overfitting and demand high computational resources. This study was also limited to single-source domain vibration signals from an experimental-scale mill. And, Sener et al. have trained a Deep Multi-Layer Perceptron (DMLP) algorithm by using the vibration labels marked by analyzing frequency-domain data obtained from FFT [24]. The algorithm selects time-domain signal features, such as root mean square, gap factor, skewness, crest factor, and shape factor, as inputs to detect vibrations. Nevertheless, multilayer perceptron training is hindered by a large number of parameters, slow convergence, and susceptibility to overfitting. Notably, Li et al. proposed a three-stage modeling framework using a Convolutional Neural Network (CNN) to predict ball mill performance [25]. This approach employs a discrete element method to generate data, including measurable variables such as Acoustic Emission (AE) signals, power consumption, and grinding rate. It compares pre-trained and untrained models through transfer learning. However, traditional neural networks have high computational demands, necessitating the introduction of alternative methods to reduce computational effort. Table 1 compares the differences between existing research and this paper in terms of signal acquisition methods and model construction.

The innovations of this study were mainly reflected in the following two aspects.

- By synchronously collecting axial, tangential, and radial vibration signals from the mill shell using triaxial wireless vibration sensors, a frequency-domain energy spectrum with industrial noise suppression is constructed for multidimensional vibration feature analysis.

- This study proposed a mill condition classification model that integrated deep dilated separable convolution and multi-head attention. The deep convolutional layers enabled independent learning of the three-dimensional vibration information, allowing the model to finely understand the unique process patterns of each dimension and avoid information confusion. The dilated convolutional layers dynamically adjusted the receptive fields to capture cross-band spectral features while ensuring lightweight computation. The multi-head attention layer was used to extract global dependencies across time steps and key frequency-domain features from the mill vibration signals.

	Existing research	The proposed method
Signal acquisition	Electrical signal [7], one-dimensional sound signal [8], [9], one-dimensional vibrationa signal of the bearing [13], [14].	The three-dimensional axial, tangential, and radial vibration signals of the mill shell.
Model construction	The modeling is primarily based on using a single machine learning method in experimental-scale mills [17], [18].	Feature extraction and neural network modeling are combined on industrial-scale mills.

Tab. 1. Comparison of existing related work.

The rest of the paper is organized as follows. The experimental process and milling machine vibration signal processing methods are introduced in Sec. 2. Then, in Sec. 3, the modeling method proposed in this paper is presented and analyzed. In Sec. 4 and Sec. 5, the effectiveness of the algorithm is verified by comparative experiments and the conclusion is given respectively.

2. Experimental and Signal Analysis

We selected a ball mill from a cement plant, with dimensions of 14.5 m × 4.2 m, to obtain the three-dimensional vibration response of the ball mill under various operating conditions. Initially, we installed a wireless vibration sensor on the surface of the mill cylinder. The sensor was secured using an adapter tool attached to screws on the cylinder's surface. The three-dimensional vibration sensor used in the experiment is a triaxial accelerometer with wireless communication, with a sensitivity of 0.5 mg, a frequency response range from 10 Hz to 10 kHz, an error range of less than 5%, a dustproof rating of IP67, and dimensions of ∅56 × 84 mm. The installation position and the shape of the sensor are illustrated in Fig. 1(a) and (b).

To capture vibration information in three axial directions, we oriented the sensor's Z-axis along the mill barrel axis, the Y-axis along the barrel's tangent, and the X-axis perpendicular to the barrel. The sensor, as the transmitter, collects the vibration signals and transmits them via ZigBee communication at a working frequency of 2.4 GHz to the wireless edge terminal, which serves as the receiver. The communication distance on site is less than 5 meters. The wireless edge terminal model is VS3000EDGE, with a ceramic antenna type, and a transmission rate of 256 kbps, as shown in Fig. 1(c) and (d). This setup allows us to obtain raw time-domain vibration data.

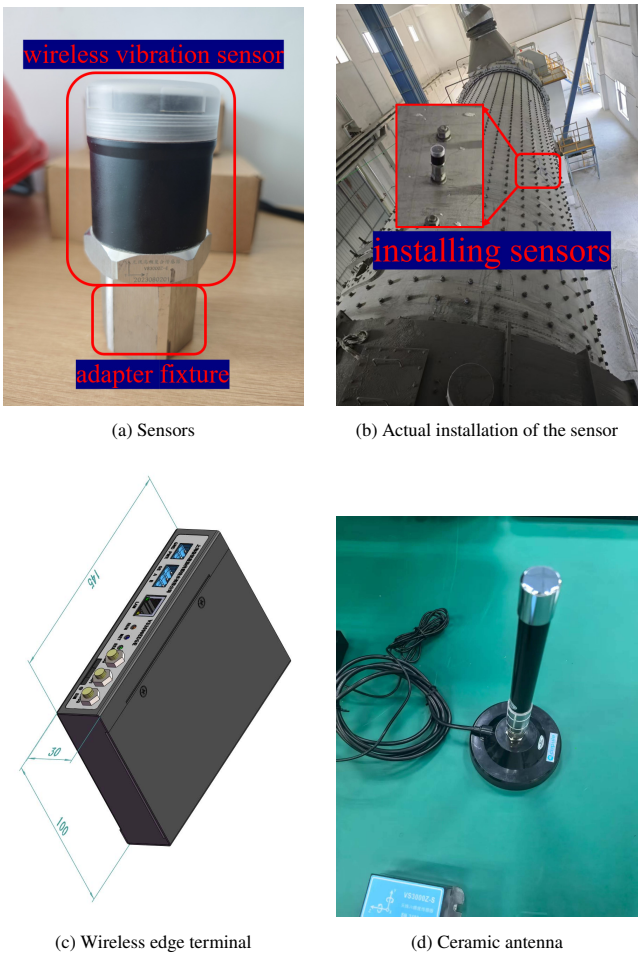
2.1 Vibration Signal Acquisition

In this experiment, we employed a sampling frequency of 1600 Hz. Since the bandwidth of the industrial mill we used is approximately 700 Hz, Shannon's sampling theorem dictates that the sampling frequency should be more than twice the signal bandwidth to ensure accurate recovery of the original information. To account for the delay in wireless signal transmission, we set the sampling duration to 4 seconds. Consequently, each sampling session yields 6400 data points for the X-axis, Y-axis, and Z-axis, with an interval of 0.000625 seconds between each point. The mill conditions are categorized into underload, normal load, and overload, with a total of 2100 rounds of data collected to form the dataset.

The time-domain information for each working condition is illustrated in Fig. 2.

2.2 Frequency-domain Transform

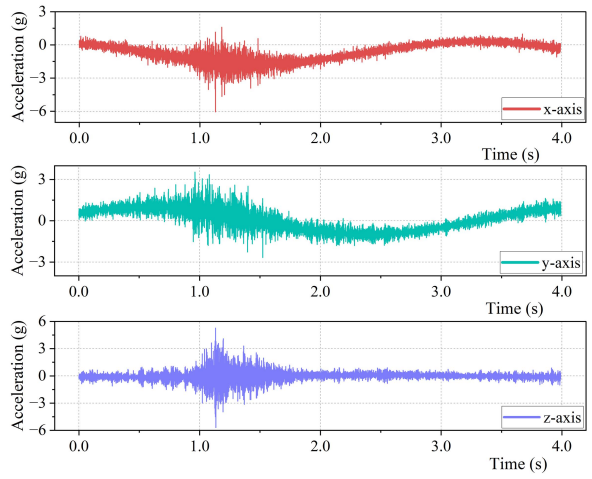
The time-domain vibration signals received by sensors contain extensive information about the vibration conditions of ball mill grinding. However, due to the presence of noise and the specific frequency range of these signals, interpreting the vibration information directly from the time-domain is challenging. Therefore, it is necessary to transform the time-domain vibration signals into the frequency-domain to facilitate interpretation.



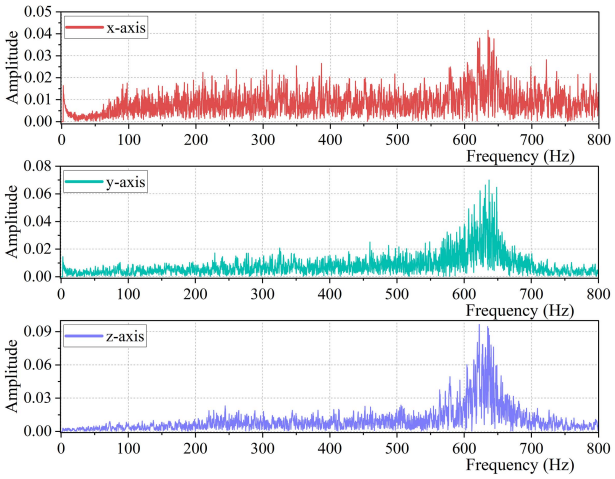
(a) Sensors (b) Actual installation of the sensor

(c) Wireless edge terminal (d) Ceramic antenna

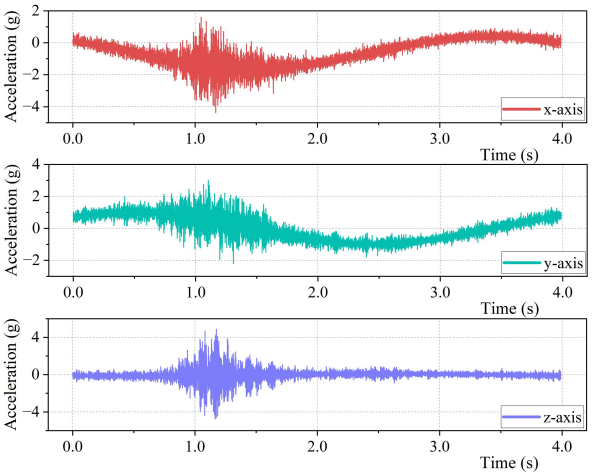
Fig. 1. Experimental equipment and installation.



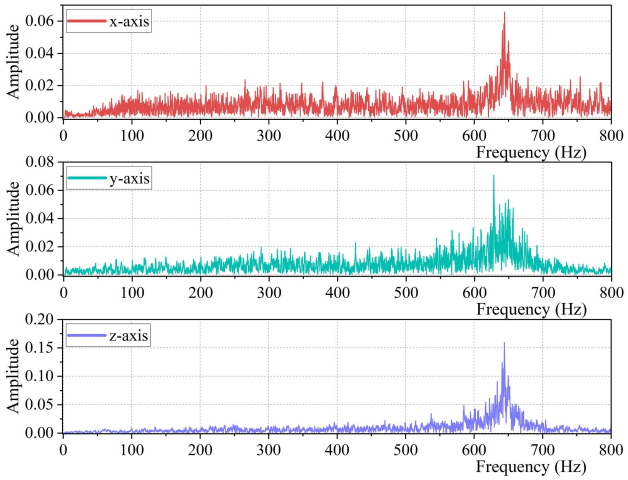
(a) Underload information in time-domain



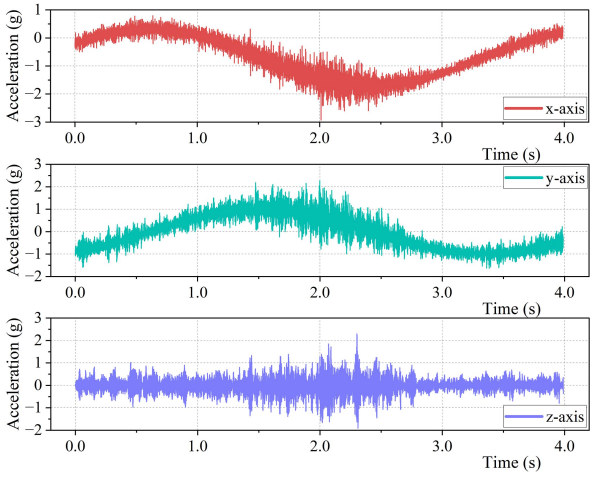
(b) Underload information in frequency-domain



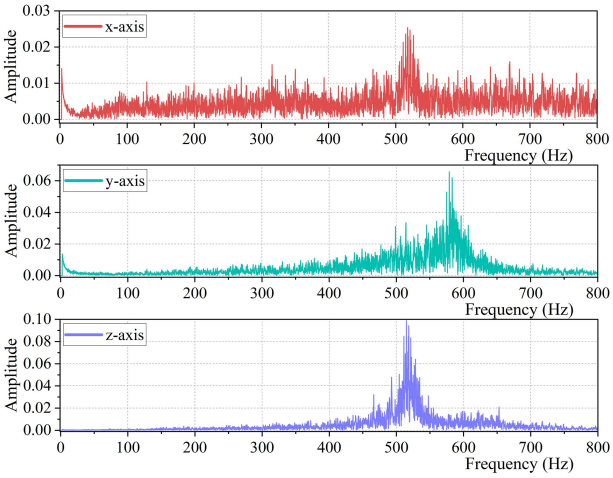
(c) Normal load information in time-domain



(d) Normal load information in frequency-domain



(e) Overload information in time-domain



(f) Overload information in frequency-domain

Fig. 2. Time-domain and frequency-domain information of the vibration signal.

The frequency content in the sampled signal can be calculated using the Discrete Fourier Transform (DFT), as shown in (1) [26].

$$X(f) = \sum_{n=0}^{N-1} x(n) \mathrm{e}^{-\mathrm{j}2\pi fn/N}, f = 0, 1, 2, \ldots, N-1 \quad (1)$$

where n is the number of samples, f is the sampling frequency, j is the imaginary unit, satisfying $\mathrm{j}^2 = -1$ and $\mathrm{e}^{(\cdot)}$ is the exponential operator.

Therefore, the frequency-domain amplitude of the sampled vibration signal can be calculated by (2) and the $X_\mathrm{A}(f)$ in this equation can be calculated efficiently using the Fast Fourier Transform (FFT) [27].

$$X_\mathrm{A}(f) = \frac{2}{N} \left| \sum_{n=0}^{N-1} x(n) \mathrm{e}^{-\frac{\mathrm{j}2\pi fn}{N}} \right|, f = 0, 1, 2, \ldots, \frac{N}{2} - 1 \quad (2)$$

where $X_\mathrm{A}(f)$ is a complex vector of length N, and the sum of its real and imaginary parts is denoted as $X_\mathrm{R}(f) + \mathrm{j}X_\mathrm{I}(f)$. The absolute value of $X_\mathrm{A}(f)$ is shown as (3):

$$|X_\mathrm{A}(f)| = \sqrt{[X_\mathrm{R}(f)]^2 + [X_\mathrm{I}(f)]^2} \quad (3)$$

and $|X_\mathrm{A}(f)|$ includes the amplitude of the positive and negative sampling frequencies. The corresponding spectrum of $x(n)$ is denoted as the first $N/2$ points of $|X_\mathrm{A}(f)|$ with a scale factor of $2/N$.

After performing the calculation, we obtained the distribution of the total energy of the time-domain signal in the frequency-domain. This approach allows for the analysis of any specific frequency band, thereby overcoming the challenge of selecting a single dominant frequency [18]. The transformed frequency-domain information for each operating condition is illustrated in Fig. 2.

2.3 Analysis of Mutual Information Value

This study simplified the model by using depthwise separable convolutions. The efficiency of depthwise separable convolutions assumes that features from different channels are independent, without explicitly modeling the mutual information (MI) between them. MI quantifies the dependence between two random variables, and $\mathrm{I}(X;Y) \approx 0$ indicates that X and Y are approximately independent. In general, to ensure the effectiveness of depthwise separable convolutions, it is required that $\mathrm{I}(X;Y) < 0.1$. This study used information-theoretic tools to compute the mutual information values between pairs of dimensions in the three-dimensional vibration signals. The results were presented in Tab. 2. The table showed that the mutual information values between the dimensions of the three-dimensional vibration signals were all below 0.086, confirming the applicability of depthwise separable convolution to simplify the model.

	X-axis	Y-axis	Z-axis
X-axis	1	0.05025	0.05009
Y-axis	0.05025	1	0.08628
Z-axis	0.05009	0.08628	1

Tab. 2. Statistical analysis of mutual information values between dimensions.

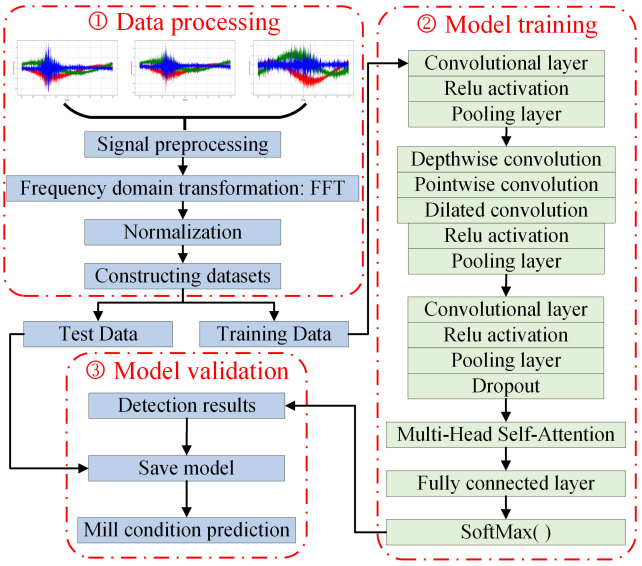


Fig. 3. Modeling framework.

3. Proposed Method

This section introduced the key components of the proposed Depthwise Separable Dilated Multi-Head Attention CNN (DM-CNN) algorithm, specifically the depthwise separable dilated convolution module and the multi-head attention mechanism. The model framework was illustrated in Fig. 3. The signal underwent preprocessing steps, including FFT frequency domain transformation and normalization, to construct training and testing datasets. In the training phase, convolutional layers, activation functions, and pooling layers were applied to extract features. The depthwise separable dilated convolution layer extracted frequency domain features from the signal using convolutional kernels and pooling operations. Dilated convolution also expanded the receptive field, allowing the model to capture signal variations over longer time intervals. The multi-head attention mechanism focused on different parts of the input features, capturing important patterns or anomalous fluctuations, which enhanced the model's predictive ability for specific operational conditions. The model further processed the data through pooling layers, dropout layers, and fully connected layers. Finally, the SoftMax function was applied for classification prediction, generating and storing the model results. In the validation phase, testing data were used to evaluate the model's performance and predict operational conditions.

3.1 Deeply Separable Dilated Convolutional

Depth separable dilated convolution is an advanced technique in CNNs that integrates depthwise, pointwise, and dilated convolutions. Depthwise separable convolution breaks down the standard convolution into spatial (per-channel) and pointwise (channel fusion) convolutions, reducing the parameter count by a factor of approximately K^2 compared to traditional convolution, where K is the kernel size. This reduction in parameters significantly decreases model complexity. Dilated convolution controls the sparse sampling interval of the kernel using the dilation rate, thereby expanding the receptive field without increasing the parameter count. It effectively captures long-range dependencies in frequency-domain images, such as wideband harmonic associations and resonance peak patterns. The module design naturally supports parallel computation, making it ideal for tasks like mill condition classification, which require high real-time performance.

3.2 Multi-head Attention

The multi-head attention mechanism learns the importance weights of frequency-domain images across different subspaces in parallel [28]. It dynamically focuses on key frequency bands strongly correlated with load conditions, such as high-frequency ranges dominated by material impact and low-frequency ranges dominated by barrel vibration, facilitating feature interaction across subspaces. Additionally, it adaptively adjusts the attention weights of the frequency-domain energy distribution, suppressing irrelevant noise, such as equipment vibrations and sensor interference, while enhancing the discriminative representation of load features. Furthermore, the attention mechanism compensates for the limitations of convolution operations in modeling global context. Meanwhile, the local features extracted by convolution provide spatial prior knowledge to the attention mechanism. Together, they enhance the model's ability to analyze complex frequency-domain patterns synergistically.

4. Experiment

This section evaluated the effectiveness of the proposed grinding condition classification method through several experiments. Experimental data were obtained from a cement plant in China by installing a three-axis wireless vibration sensor on the ball mill. A total of 2,100 data points, including underload, normal load, and overload conditions, were collected. The data were divided into training, validation, and test sets at an 80%, 10%, and 10% ratio, respectively, to create the grinding condition classification dataset. The neural network structure for all experiments was implemented using PyTorch 2.0.0 and Python 3.9.19. Model training was performed on an NVIDIA Quadro RTX 5000 GPU with CUDA 11.7 support. Model parameter settings were optimized through multiple training sessions using an exhaustive search method, considering hardware constraints, to determine the convolutional network parameters in the DM-CNN model, as detailed in Tab. 3. Specifically, "in_c" refers to the number of input channels, "out_c" to the number of output channels, "kernel" to the size of the convolution kernel, "stride" to the step size, "padding" to the number of zero-padding layers, "groups" to the number of independent convolution groups in the depthwise separable convolution layer, and "dilation" to the dilation rate. For the multi-head attention layer, key parameters were set as follows. "embedding_dim" was equal to 256, indicating the embedding dimension, and "n_head" was equal to 4, representing the number of attention heads.

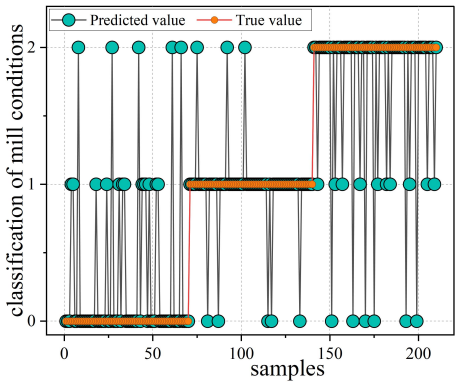
The loss function used in all networks was the mean square error (MSE) with L2 regularization. The training process was configured with a maximum of 300 epochs and a batch size of 16. The Adam optimizer was applied for optimization. A dynamic learning rate schedule was employed, starting with an initial learning rate of 0.002, a decay period of 30 epochs, and a decay factor of 0.9. Training was halted when the absolute difference in validation loss stayed below 0.0001 for 20 consecutive iterations.

4.1 Verification of the Validity of Three-dimensional Vibration Signals

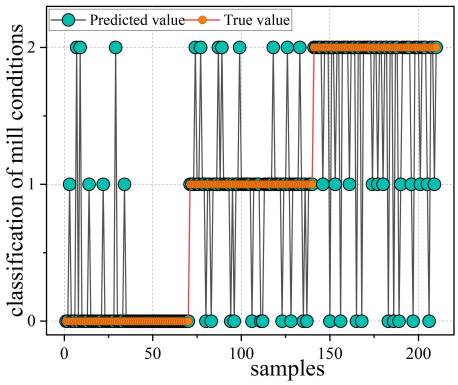
This section evaluated the effectiveness of the proposed three-dimensional vibration signal modeling approach. The detailed procedure was as follows. First, the acquired three-dimensional vibration signals were decomposed into one-dimensional components to simulate traditional one-dimensional vibration signals along three axes (denoted as X-axis, Y-axis, and Z-axis). Next, both acquired three-dimensional and one-dimensional datasets were put into model training and conducted mill condition predictions. Finally, a comparative analysis of the results was performed. The performance of classification predictions was illustrated in Fig. 4, with a detailed statistical data of classification predictions shown in Fig. 5, where Class 0 represented underload, Class 1 corresponded to normal load, and Class 2 denoted overload.

Parameters	in_c	out_c	kernel	stride	padding	groups	dilation
Convolutional Layer	3	16	5	2	1	-	-
Depthwise Separable Convolutional Layer	16	16	3	1	1	16	-
Pointwise Convolutional Layer	16	32	1	1	0	-	-
Dilated Convolutional Layer	32	64	3	1	2	-	2

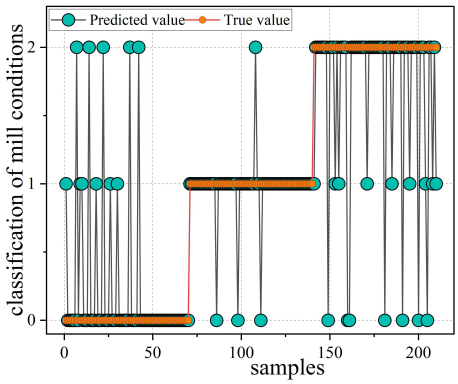
Tab. 3. Convolutional network parameter settings.



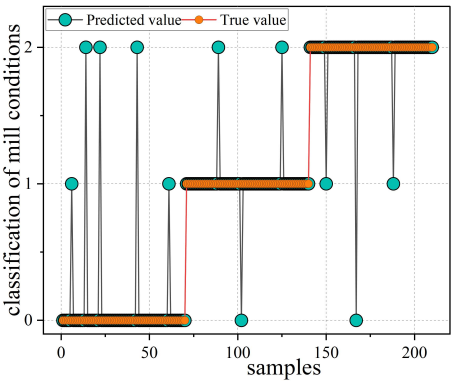
(a) Single x-axis classification



(b) Single y-axis classification



(c) Single z-axis classification



(d) Classification results of three-dimensional vibration signals

Fig. 4. Classification performance of one-dimensional and three-dimensional vibration signals.

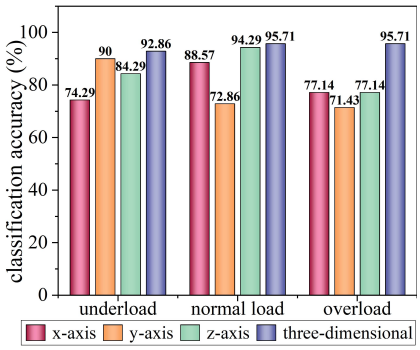


Fig. 5. Comparison of classification accuracy.

The experimental results indicated that the average prediction accuracies based on one-dimensional vibration signals were 80%, 78.10%, and 85.24%, respectively. In contrast, the average prediction accuracy of the three-dimensional vibration signal reached 94.29%. Notably, three-dimensional vibration signals perform significantly better than one-dimensional vibration signals in classifying overload grinding conditions. Upon further evaluation, the average prediction accuracies of the three-dimensional vibration signal model for underload, normal load, and overload conditions were 92.86%, 94.29%, and 95.71%, respectively.

The results show that the mill's vibration signals are fundamentally a superposition of multiple excitation sources. In particular, three-dimensional vibration signals offered a more comprehensive depiction of the mill's operating conditions. These included tangential eccentric vibrations of the mill shell and axial impact vibrations caused by the grinding media. Therefore, by integrating dynamic features across multiple axes, three-dimensional vibration signals allow for a more precise characterization of the mill's operational state.

4.2 Model Evaluation

This section presents a comprehensive evaluation of the DM-CNN model's overall performance, category differentiation, and robustness in grinding condition classification, based on three metrics: F1-score, confusion matrix, and AUC-ROC curve. Specifically, the F1-score indicates classification robustness under imbalanced mill status categories. The calculation formula of *F1-score* is shown as (4), where *TP* is True Positive, *FP* is False Positive, and *FN* is False Negative. The confusion matrix displays misclassification details and error distribution for each mill status, and the AUC-ROC curve quantifies the model's generalization ability concerning threshold robustness. The model parameters remain consistent with those in the previous section.

$$
\begin{aligned}
F1\text{-}score &= \frac{2 \times Precision \times Recall}{Precision + Recall}, \\
Precision &= \frac{TP}{TP + FP}, \\
Recall &= \frac{TP}{TP + FN}.
\end{aligned}
\tag{4}
$$

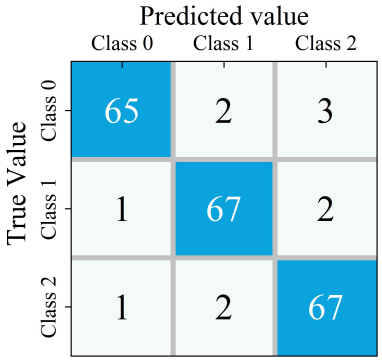


Fig. 6. Confusion matrix of DM-CNN.

mill condition	*TP*	*FP*	*FN*	*Precision*	*Recall*	*F1-score*
Class 0	65	2	5	0.97	0.93	0.95
Class 1	67	4	3	0.94	0.96	0.95
Class 2	67	5	3	0.93	0.96	0.94

Tab. 4. F1-score of DM-CNN.

The confusion matrix for the model is shown in Fig. 6. The definitions of Class 0, Class 1, and Class 2 were consistent with those in the previous section. Overall, the model achieved an accuracy of 94.76%. However, the model showed a false negative issue in detecting underloaded grinding conditions, $FN_{\text{Class0}} = 5$, which was higher than in other conditions. Additionally, the model showed false positive detections in overloaded grinding conditions, $FP_{\text{Class2}} = 5$, which was also higher than in other conditions. This suggests that, under balanced sampling, there is still room for improvement in the model's classification boundaries. The boundary for underloaded grinding conditions is stricter, while that for overloaded conditions is more lenient.

The F1-scores for each category were calculated based on the confusion matrix, and the results are shown in Tab. 4. $Recall_{\text{Class0}} = 0.93$ and $Precision_{\text{Class2}} = 0.93$ reflect similar issues with the model's classification boundaries. The F1-scores for the three mill conditions indicated that the model performed excellently overall, with some subtle differences. Specifically, the $F1\text{-}score_{\text{Class2}} = 0.94$ was lower than that for the other categories, and there was a slight imbalance between Categories $Recall_{\text{Class2}} = 0.96$ and $Precision_{\text{Class2}} = 0.93$. Therefore, adjusting the classification threshold for overloaded mill conditions is necessary to further improve performance.

The AUC-ROC curve evaluated the model's performance across various classification thresholds by illustrating the relationship between the true positive and false positive rates. The AUC, representing the area under the ROC curve, quantifies the model's ability to differentiate categories, with higher values closer to 1 indicating superior performance. This study adapted the method by treating one class as the true positive rate and the other two as the false positive rate, yielding three distinct curves. Figure 7 presented the AUC-ROC curves for DM-CNN.

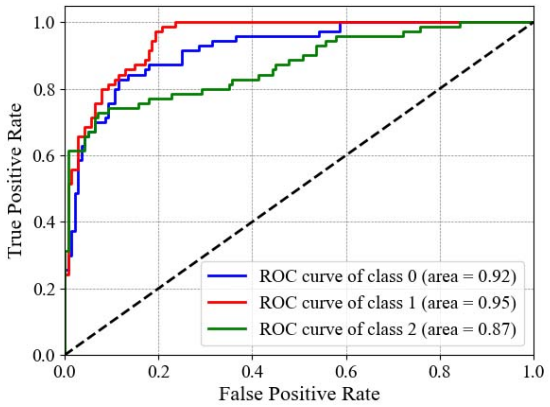


Fig. 7. AUC-ROC curve of DM-CNN.

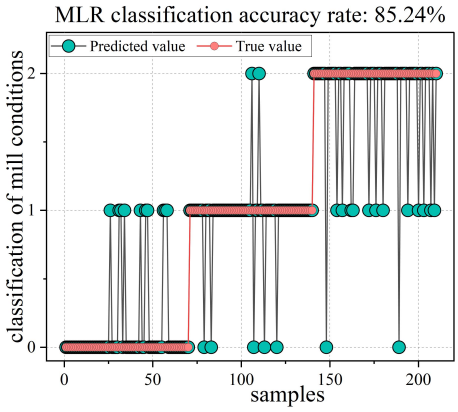
Figure 7 shows the ROC curve of the DM-CNN model, which is located near the upper-left region. Additionally, the AUC values of the DM-CNN model under various grinding conditions and thresholds approached 1. These results demonstrate that the DM-CNN model is highly effective in distinguishing between positive and negative samples.

4.3 Comparison with Other Methods

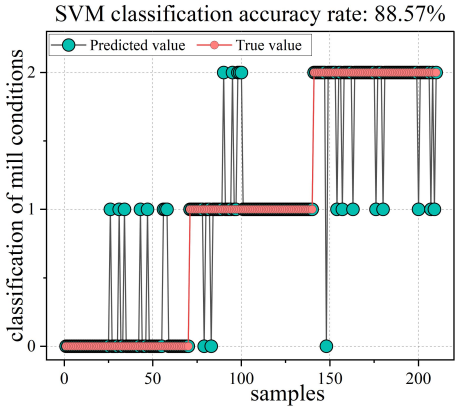
To comprehensively evaluate the DM-CNN model's efficacy, this section benchmarked it against representative linear (Multinomial Logistic Regression, MLR), non-linear (Support Vector Machine, SVM), and ensemble learning (eXtreme Gradient Boosting, XGBoost) classifiers. Model parameter settings were optimized through multiple training sessions using an exhaustive search method, and the model parameters used in this section were determined as follows. The MLR was implemented using a single fully connected neural network, with input and output dimensions based on the mill condition classification dataset. The learning rate matched that of the DM-CNN model. The SVM employed a radial basis function kernel, with the penalty coefficient set to $C = 1.0$ and the kernel coefficient set to $gamma = scale$. The XGBoost had a learning rate of $lr = 0.1$, with the maximum tree depth set to $max\_d = 6$. The row and column sampling ratios were both set to 0.8, the minimum leaf node sample weight $min\_c\_w = 1$, and the L2 regularization term $lambda = 1$. The models' classification performance was shown in Fig. 8.

As shown in Figs. 8 and 4(d), the DM-CNN model achieves the highest overall accuracy across various grinding conditions, outperforming MLR, SVM, and XGBoost by 9.05%, 5.72%, and 5.24%, respectively. Secondly, detailed analysis revealed that the MLR model performed poorly in distinguishing overload mill conditions. This was primarily due to its linear architecture, which could not effectively capture the nonlinear relationships in the three-dimensional vibration frequency-domain features. Although the SVM model employed a kernel-based nonlinear classification method, its accuracy in identifying underload and overload mill conditions remained lower than that of DM-CNN (Fig 4(d)). The XGBoost ensemble learning method, utiliz-

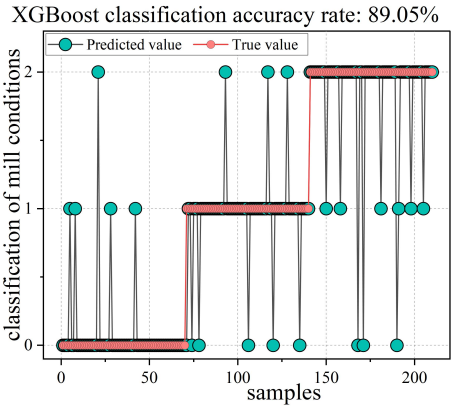
ing gradient boosting mechanisms and regularization techniques, demonstrated high accuracy and generalization in other fields. However, due to the non-stationarity and complexity of the mill's three-dimensional vibration signals, as well as the loss of spatial structural information, the classification accuracy of XGBoost was limited to 89.05%. In summary, DM-CNN successfully performs adaptive weight adjustment and multi-scale frequency feature extraction by integrating deep separable dilated convolutions with multi-head attention layers. This enables the effective recognition and classification of mill conditions in the specific industrial context of cement grinding.



(a)



(b)



(c)

Fig. 8. Classification methods visualization.

5. Conclusion

To classify grinding conditions in the cement industry, this study employed a novel three-dimensional wireless vibration sensor to collect comprehensive vibration data from the ball mill. To address low classification accuracy in existing models using three-dimensional vibration data, this study introduced a novel deep learning approach that combines depthwise separable dilated convolution with a multi-head attention mechanism. The method employs dilated convolution to expand the receptive field, capturing key patterns that reflect operational characteristics of the mill, thereby addressing the ball mill's vibration signal spectra and long-range dependencies. A lightweight model based on depthwise separable convolution is used to meet the real-time deployment requirements of the equipment. Moreover, the method utilizes a multi-head attention mechanism to adaptively select discriminative sensor feature combinations, effectively suppressing noise interference. The experimental results show that, compared to individual modules, this complementary ensemble method that integrates mill three-dimensional vibration information can more comprehensively characterize the mill's condition and achieve higher classification accuracy.

The proposed method demonstrated effective mill condition classification but lacks validation across diverse ball mill types and material conditions due to single-source industrial data. Future research will prioritize multi-plant validation to enhance model generalizability and operational robustness.

The simulated dataset and model validation scheme related to this study can be obtained by contacting the author.

References

[1] LIU, Y., YAN, G., XIAO, S., et al. A multi-task model for mill load parameter prediction using physical information and domain adaptation: validation with laboratory ball mill. *Minerals Engineering*, 2025, vol. 222, p. 1–12. DOI: 10.1016/j.mineng.2024.109148

[2] ALTUN, O. Energy and cement quality optimization of a cement grinding circuit. *Advanced Powder Technology*, 2018, vol. 29, no. 7, p. 1713–1723. DOI: 10.1016/j.apt.2018.04.006

[3] SOUTSOS, M., BOYLE, P., VINAI, R., et al. Factors influencing the compressive strength of fly ash based geopolymers. *Construction and Building Materials*, 2016, vol. 110, p. 355–368. DOI: 10.1016/j.conbuildmat.2015.11.045

[4] JANKOVIC, A., VALERY, W., DAVIS, E. Cement grinding optimisation. *Minerals Engineering*, 2004, vol. 17, no. 11–12, p. 1075–1081. DOI: 10.1016/j.mineng.2004.06.031

[5] ZHOU, M., ZHU, Z., HU, F., et al. An industrial load classification method based on a two-stage feature selection strategy and an improved MPA-KELM classifier: A Chinese cement plant case. *Electronics*, 2023, vol. 12, no. 15, p. 1–22. DOI: 10.3390/electronics12153356

[6] SCHNEIDER, M., HOENIG, V., RUPPERT, J., et al. The cement plant of tomorrow. *Cement and Concrete Research*, 2023, vol. 173, p. 1–11. DOI: 10.1016/j.cemconres.2023.107290

[7] WANG, X., YAO, Z., PAPAEFTHYMIOU, M. A real-time electrical load forecasting and unsupervised anomaly detection framework. *Applied Energy*, 2023, vol. 330, p. 1–13. DOI: 10.1016/j.apenergy.2022.120279

[8] KALANTARI, S., MADADI, A., RAMEZANI, M., et al. Controlling the ground particle size and ball mill load based on acoustic signal, quantum computation basis, and least squares regression, case study: Lakan lead-zinc processing plant. *International Journal of Industrial Electronics Control and Optimization*, 2023, vol. 6, no. 3, p. 205–218. DOI: 10.22111/ieco.2023.45981.1488

[9] XU, X., WEN, H., LIN, H., et al. Online detection method for variable load conditions and anomalous sound of hydro turbines using correlation analysis and PCA-adaptive-K-means. *Measurement*, 2024, vol. 224, p. 1–10. DOI: 10.1016/j.measurement.2023.113846

[10] WANG, H., XIONG, D., DUAN, Y., et al. Advances in vibration analysis and modeling of large rotating mechanical equipment in mining arena: A review. *AIP Advances*, 2023, vol. 13, no. 11, p. 1–13. DOI: 10.1063/5.0179885

[11] YANG, L., YANG, H. Load identification method of ball mill based on the CEEMDAN-wavelet threshold-PMMFE. *Gospodarka Surowcami Mineralnymi – Mineral Resources Management*, 2024, vol. 40, no. 2, p. 1–18. DOI: 10.24425/gsm.2024.150823

[12] TANG, J., ZHAO, L., ZHOU, J., et al. Experimental analysis of wet mill load based on vibration signals of laboratory-scale ball mill shell. *Minerals Engineering*, 2010, vol. 23, no. 9, p. 720–730. DOI: 10.1016/j.mineng.2010.05.001

[13] TANG, W., ZHANG, F., LUO, X., et al. Method of vibration signal processing and load-type identification of a mill based on ACMD-SVD. *Gospodarka Surowcami Mineralnymi – Mineral Resources Management*, 2023, vol. 39, no. 1, p. 217–233. DOI: 10.24425/gsm.2023.144626

[14] DAS, S., DAS, D., BEHERA, S., et al. Interpretation of mill vibration signal via wireless sensing. *Minerals Engineering*, 2011, vol. 24, no. 3–4, p. 245–251. DOI: 10.1016/j.mineng.2010.08.014

[15] SU, Z., WANG, P., YU, X., et al. Experimental investigation of vibration signal of an industrial tubular ball mill: Monitoring and diagnosing. *Minerals Engineering*, 2008, vol. 21, no. 10, p. 699–710. DOI: 10.1016/j.mineng.2008.01.009

[16] TANG, J., CHAI, T., YU, W., et al. Feature extraction and selection based on vibration spectrum with application to estimating the load parameters of ball mill in grinding process. *Control Engineering Practice*, 2012, vol. 20, no. 10, p. 991–1004. DOI: 10.1016/j.conengprac.2012.03.020

[17] SHI, J., SI, G., LI, S., et al. Feature extraction based on the fractional Fourier transform for vibration signals with application to measuring the load of a tumbling mill. *Control Engineering Practice*, 2019, vol. 84, p. 238–246. DOI: 10.1016/j.conengprac.2018.11.012

[18] BEHERA, B., MISHRA, B., MURTY, C. Experimental analysis of charge dynamics in tumbling mills by vibration signature technique. *Minerals Engineering*, 2007, vol. 20, no. 1, p. 84–91. DOI: 10.1016/j.mineng.2006.05.007

[19] TANG, J., CHAI, T., YU, W., et al. Modeling load parameters of ball mill in grinding process based on selective ensemble multisensor information. *IEEE Transactions on Automation Science and Engineering*, 2012, vol. 10, no. 3, p. 726–740. DOI: 10.1109/TASE.2012.2225142

[20] TANG, J., YU, W., ZHAO, L., et al. Modeling of operating parameters for wet ball mill by modified GA-KPLS. In *Proceedings of the Third International Workshop on Advanced Computational Intelligence (IWACI)*. Suzhou (China), 2010, p. 282–287. DOI: 10.1109/IWACI.2010.5585151

[21] ZHAO, L., TANG, J., YU, W., et al. Modelling of mill load for wet ball mill via GA and SVM based on spectral feature. In *Proceedings of the 2010 IEEE Fifth International Conference on Bio-Inspired Computing: Theories and Applications (BIC-TA)*. Changsha (China), 2010, p. 874–879. DOI: 10.1109/BICTA.2010.5645241

[22] LIU, Z., CHAI, T., YU, W., et al. Multi-frequency signal modeling using empirical mode decomposition and PCA with application to mill load estimation. *Neurocomputing*, 2015, vol. 169, p. 392–402. DOI: 10.1016/j.neucom.2014.08.087

[23] HUANG, P., GUO, J., SANG, G., et al. Soft measurement of ball mill load under variable working conditions based on deep transfer learning. *Measurement Science and Technology*, 2022, vol. 33, no. 7, p. 1–14. DOI: 10.1088/1361-6501/ac5c92

[24] SENER, B., SERIN, G., GUDELEK, M., et al. Intelligent chatter detection in milling using vibration data features and deep multi-layer perceptron. In *Proceedings of the 2020 IEEE International Conference on Big Data*. Atlanta (USA), 2020, p. 4759–4768. DOI: 10.1109/BigData50022.2020.9378223

[25] LI, Y., BAO, J., CHEN, T., et al. Prediction of ball milling performance by a convolutional neural network model and transfer learning. *Powder Technology*, 2022, vol. 403, p. 1–9. DOI: 10.1016/j.powtec.2022.117409

[26] YESIL, S., YILMAZ, A. Identification of the linear systems of the Wiener Hammerstein RF power amplifier model using DFT analysis. *Radioengineering*, 2024, vol. 33, no. 2, p. 265–273. DOI: 10.13164/re.2024.0265

[27] LIU, Y., RAO, X., HU, J., et al. A weak target detection algorithm IAR-STFT based on correlated K-distribution sea clutter model. *Radioengineering*, 2023, vol. 32, no. 1, p. 33–43. DOI: 10.13164/re.2023.0033

[28] YUAN, S., YOUREN, W., JIE, Y., et al. Context aware multimodal fusion YOLOv5 framework for pedestrian detection under IoT environment. *Radioengineering*, 2025, vol. 34, no. 1, p. 118–131. DOI: 10.13164/re.2025.0118

About the Authors . . .

Dianyuan JU was born in 1994. He has been a Ph.D. candidate at University of Jinan since 2023. His research focuses on cement industry process control and signal processing.

Rongfeng ZHANG (corresponding author) was born in 1992. He received his Ph.D. from University of Jinan in 2023. His research interests include industrial process control and automated control.

Xiaohong WANG was born in 1963. He received his Ph.D. from Northeastern University in 2005. His research interests include industrial process control and industrial artificial intelligence.

Zhao LIU was born in 1987. He received his Ph.D. from Shenyang Institute of Automation Chinese Academy of Sciences in 2016. His research interest is industrial process control.

Bing HUANG was born in 1991. He received his Ph.D. from University of Jinan in 2021. His research interest is automated control.

Hongliang YU was born in 1976. He received his Ph.D. from Northeastern University in 2007. His research interest is industrial process control.