LCRTR-Net: A low-cost real-time recognition network for rail corrugation in railway transportation

Xueyang Tang[a], Xiaopei Cai[a], Yuqi Wang[a], Yue Hou[b]

[a]School of Civil Engineering, Beijing Jiaotong University, Beijing, China

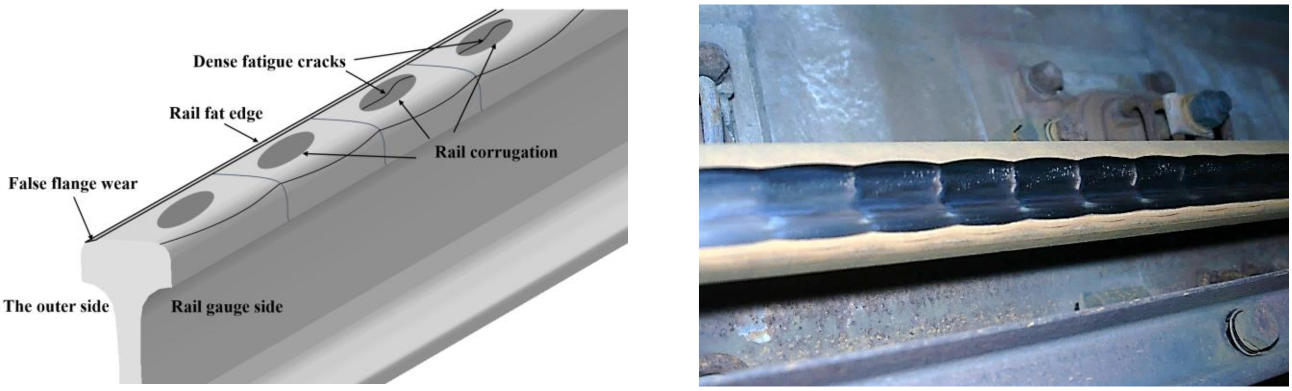
[b]Department of Civil Engineering, Faculty of Science and Engineering, Swansea University, Swansea, UK

Abstracts: Rail corrugation has a significant impact on the safety of high-speed railway operations, making its identification particularly important. Traditional manual inspection methods are infeasible for large-scale identification within limited time frames, while existing methods based on machine vision or axle box acceleration face challenges such as high costs, complex equipment installation and maintenance, as well as difficulties in achieving real-time performance. To address these challenges, this study proposes an innovative low-cost real-time recognition network (LCRTR-Net), which utilizes accelerometers installed on the underside of the train body and combines wavelet packet decomposition with dilated causal convolution in a residual neural network. Specifically, the approach first extracts the latent features of train body acceleration caused by rail corrugation through wavelet packet decomposition and reconstruction. Next, dilated causal convolution is employed to capture the temporal causal relationships and long-term dependencies of these latent features. Finally, the integration of residual connections further enhances the feature extraction performance and computational efficiency of LCRTR-Net. Experimental results demonstrate that LCRTR-Net exhibits significant generalization ability and real-time performance, achieving an average recognition accuracy exceeding 97.0%, with a recognition time of only 0.17 ms per rail corrugation sample, significantly outperforming existing rail corrugation recognition methods. This indicates that LCRTR-Net has broad application potential in practical railway operations. Future research directions will focus on unsupervised or few-shot learning algorithms and multi-sensor integration to further improve recognition accuracy and real-time performance, promoting the practical application of this technology.

Keywords: High-speed railway, Rail corrugation recognition, Car body acceleration, Wavelet packet decomposition and reconstruction, Dilated causal convolution

1 Introduction

Rail corrugation, as illustrated in Fig. 1, is a prevalent track defect in high-speed railways [1-2]. It is a primary contributor to severe vibrations, significantly compromising passenger comfort [3-4]. Regular grinding of corrugated rail sections is essential, emphasizing the need for accurate detection prior to maintenance [5]. Therefore, the precise and efficient detection of rail corrugation remains a critical challenge.



(a) Rail corrugation characteristic diagram

(b) Field rail corrugation condition

Fig. 1 Rail corrugation.

Detection methods for rail corrugation can be categorized into several types based on their underlying principles: manual detection, string measurement, inertial reference, and machine vision. The manual detection method requires inspectors to use tools such as rail corrugation inspection rulers [6-7] or inspection trolleys [8] for on-site measurements. However, this approach

---

* Corresponding authors.
E-mail addresses: xpcai@bjtu.edu.cn (X. Cai), 22110383@bjtu.edu.cn (Y. Wang)

is inefficient and impractical for large-scale assessments due to its reliance on human resources and potential for human error.

String measurement relies on a moving coordinate system relative to the rail; however, variations in rail height lead to changes in the reference frame, resulting in non-constant transfer functions (the ratio of measured value to actual value) and introducing inevitable measurement errors [9-10]. This variability can significantly compromise the accuracy of the results, particularly in tracks with uneven surfaces. The inertial reference method characterizes rail corrugation through the quadratic integration of axle box acceleration. Still, it is prone to interference from wheel conditions, complicating the elimination of wheel fault influences [11-12]. This method also assumes that the axle box acceleration is solely indicative of rail conditions, neglecting the potential influence of environmental factors such as track alignment and surface conditions, which can lead to misleading results. While the machine vision method offers high measurement accuracy, its implementation costs are prohibitively high [13-16]. Furthermore, machine vision systems often require ideal lighting and unobstructed views of the rail, which may not be feasible in all operational environments.

Previous studies have made significant contributions to rail corrugation detection. For instance, Sunaga et al. [17-18] proposed that the probability density function of axle box acceleration follows a log-normal distribution. However, this model does not account for potential variations in operational conditions that could affect the distribution, thereby limiting its applicability in diverse scenarios. Potter et al. [19] highlighted the user-friendly nature of axle box acceleration detection equipment installed on trains, making it a practical auxiliary tool for assessing rail corrugation. However, the practical effectiveness of this equipment in various environments and conditions has not been thoroughly validated. Additionally, Hou et al. [20] suggested using fuzzy approximation theory for predicting rail corrugation. This approach, while innovative, may struggle with the inherent uncertainty and variability of rail conditions, leading to less reliable predictions in dynamic environments. Roppongi et al. [21] identified rail corrugation through vertical axle box acceleration detection, defining filtering bands for locating corrugation, yet they did not adequately account for environmental factors affecting detection accuracy, such as track geometry and external vibrations. Coudert [22] estimated track irregularity using axle box acceleration; however, this method struggles in high-frequency noise environments, which are common in operational railways, reducing its effectiveness. Hopkins et al. [23-24] improved detection accuracy by processing axle box vibration acceleration with wavelet transforms, but this approach suffers from high computational complexity and insufficient real-time performance, making it less practical for immediate applications. Gomes et al. [25] analyzed axle box acceleration and extracted corrugation information using one-third octave band and wavelet packet methods; however, their approach is inefficient for processing large-scale data due to the extensive computational resources required, further hindering real-time analysis.

The frequency response function method proposed by Wei et al. [26] is effective but requires further verification under complex track conditions, where varying speeds and environmental conditions could significantly impact its reliability. Kojima et al. [27] analyzed the power spectral density of the bogie acceleration, linking railway corrugation to high-frequency vibrations; however, the limitation of this method lies in its sensitivity to high-frequency noise, which can lead to inaccurate measurement results, especially in noisy environments. Xie et al. [28] and Zhou et al. [29] implemented railway corrugation detection using numerical simulation models. Although these models are theoretically effective, their reliance on complex simulated environments makes real-time monitoring in practical applications challenging, as they do not reflect the variability of real-world conditions accurately. J. Li [30] diagnosed railway corrugation using the energy and spectrum of axle box acceleration; however, this method has limited adaptability under varying train speeds and track conditions, making it susceptible to environmental interference and noise. H. Li [31] demonstrated the effectiveness of axle box vibration signals in characterizing railway corrugation through simulation analysis, suggesting the use of techniques like de-bias filtering and Simpson's numerical integration for processing axle box acceleration signals. However, these processing steps increase computational complexity, limiting the potential for real-time applications and making it less suitable for on-the-fly assessments. Tankaka et al. [32] proposed a combined approach using axle box and car body acceleration for railway corrugation detection, analyzing data components within a fixed wavelength range and employing a "spatially weighted level" index. Nonetheless, this method may face challenges in data fusion during practical applications, affecting detection accuracy and increasing the likelihood of erroneous readings. Wang [33] associated effective axle box acceleration values and vibration response ratios with the degree of railway corrugation, proposing a rapid detection scheme for high-speed railways; however, it relies on specific sensor configurations and environmental conditions, limiting its universality and applicability across various railway systems.

In light of these limitations in traditional detection methods, recent studies have begun to integrate machine learning approaches into rail corrugation detection. For example, Jiang et al. [34] achieved a recognition rate of 93.82% using empirical mode decomposition and BP neural networks, but their method is highly data-dependent, requiring large and diverse datasets for optimal performance, which may not be readily available in practical scenarios. Zhou et al. [35] employed support vector machines for track defect recognition, achieving an accuracy of 94.67%, yet they did not fully utilize time series information, which could enhance the detection capabilities by providing context over time. Zhao et al. [36] utilized Naive Bayesian and deep convolutional neural networks for recognition, achieving high accuracy; however, their model training and optimization processes are complex and resource-intensive, demanding substantial computational power and time. Zhang [37] identified rail corrugations using wavelet packet analysis combined with support vector machine (SVM) methods based on axle box acceleration, but this approach can be limited by the computational intensity and time requirements associated with the wavelet packet analysis. Xiao et al. [38] proposed an SVM-based method for diagnosing rail corrugation using axle box acceleration, achieving an accuracy exceeding 93%. While this approach demonstrates a degree of effectiveness in detecting rail corrugation, it is limited by its reliance on complex feature extraction processes, which can be computationally intensive and time-consuming, and its performance can be sensitive to the selection of parameters and the quality of the training data, potentially leading to suboptimal results in varying operational conditions. Similarly, the method introduced by Xie et al. [39-40], based on axle box acceleration and convolutional neural networks, improved accuracy but still faces challenges in real-time detection, particularly due to the significant computational demands of deep learning models. Chen et al. [41] proposed an improved YOLO V3 algorithm for identifying surface defect images of rails, achieving an average accuracy of 87.41%, although there is still room for improvement in its accuracy, particularly in diverse environmental conditions and varying defect types.

Fig. 2 shows the results of the bibliometric analysis of publications in the field of rail corrugation detection over the past five years using VOSviewer software. Collectively, these studies highlight the advancements made in rail corrugation detection; however, they also emphasize the need for more efficient, adaptable, and real-time capable methods that can operate under varying environmental and operational conditions.

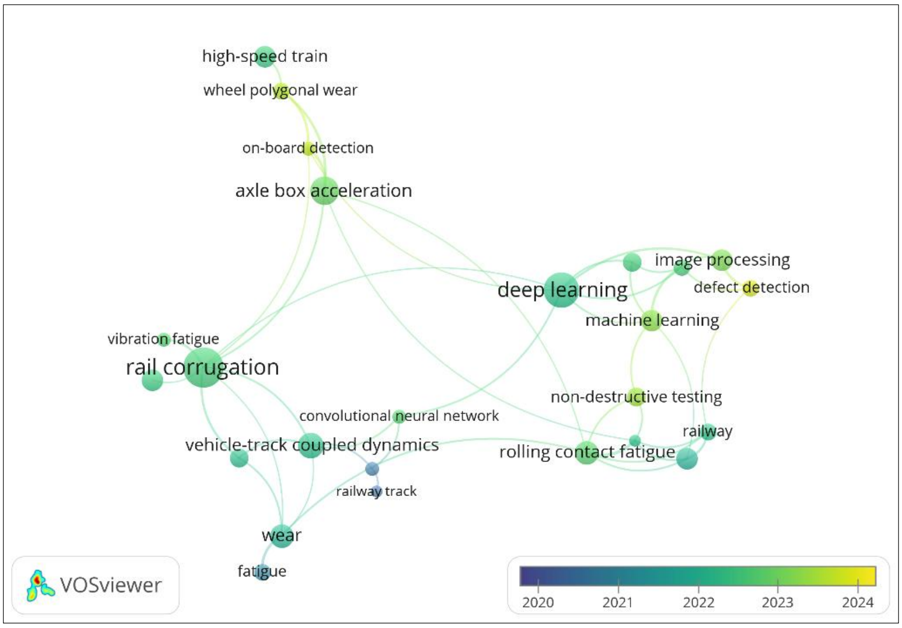


Fig. 2 Bibliometric analysis of publications in the field of rail corrugation detection over the past five years using vosviewer software.

The existing methods for detecting rail corrugation mainly face the following issues: First, traditional manual inspection methods are not only time-consuming but also susceptible to human factors, leading to insufficient detection accuracy. Second, axle box acceleration sensor-based detection methods have poor adaptability in complex environments and often fail to provide real-time data. Against this backdrop, we propose a novel rail corrugation detection method—Low-Cost Real-Time Recognition Network (LCRTR-Net)—designed to overcome the shortcomings of both traditional methods and machine learning techniques. As shown in Fig. 3, this research leverages the widespread availability of car body acceleration sensors and integrates advanced deep learning techniques to achieve more efficient real-time detection. Compared to the studies by Xie et al. [28, 39, 40] and Zhou et al. [35], LCRTR-Net innovates in model architecture and data processing, better adapting to complex railway environments while enhancing detection accuracy and real-time performance. Furthermore, the positioning of car body acceleration sensors provides greater flexibility in installation, and data collection is relatively straightforward, making our detection method more practical and feasible for real-world applications. Thus, this study not only addresses gaps in existing research but also offers new insights for the online real-time identification of rail corrugation.

The main contributions of this study can be summarized as follows:

(1) Low-cost data acquisition method: This study is the first to use car body acceleration to detect rail corrugation. Unlike traditional methods that rely on axle box acceleration sensors or visual sensors installed on specific inspection vehicles, LCRTR-Net allows sensors to be directly mounted on the chassis of normally operating vehicles. This approach reduces data acquisition costs and provides a feasible solution for large-scale online monitoring, advancing railway maintenance technology.

(2) Real-time data processing technology: The wavelet packet decomposition and reconstruction modules in LCRTR-Net effectively extract potential features of car body acceleration related to rail corrugation, while the dilated causal convolution module captures the temporal causal relationships and long-term dependencies between data, allowing for parallel processing of time-series data. This addresses the challenges of real-time detection and highlights the innovation of this study in real-time monitoring.

(3) Implications for future research: The results of this study provide a simple and effective new method for online detection of rail corrugation, promoting the application of car body acceleration in track inspection. Future research directions may explore improvements to the LCRTR-Net model to adapt to different track conditions (such as varying track materials, levels of wear, and environmental factors) to enhance its adaptability and accuracy; and investigate how to integrate data from other sensors (such as temperature, humidity, and track geometry data) to enhance detection capabilities for track defects, thereby providing a more comprehensive understanding of track conditions and improving detection accuracy and reliability, offering comprehensive technical support for railway transportation safety.

The structure of the remaining sections is as follows: Section 2 outlines the design methodology of the LCRTR-Net architecture. Section 3 introduces the measurement methods for rail corrugation and car body acceleration. Chapter 4 provides an in-depth discussion of the superiority, generalization ability, dimensionality reduction capability, and potential applications of LCRTR-Net. Finally, Section 5 summarizes the main conclusions of this study, clarifying future research directions, the limitations of LCRTR-Net, and potential improvement methods.

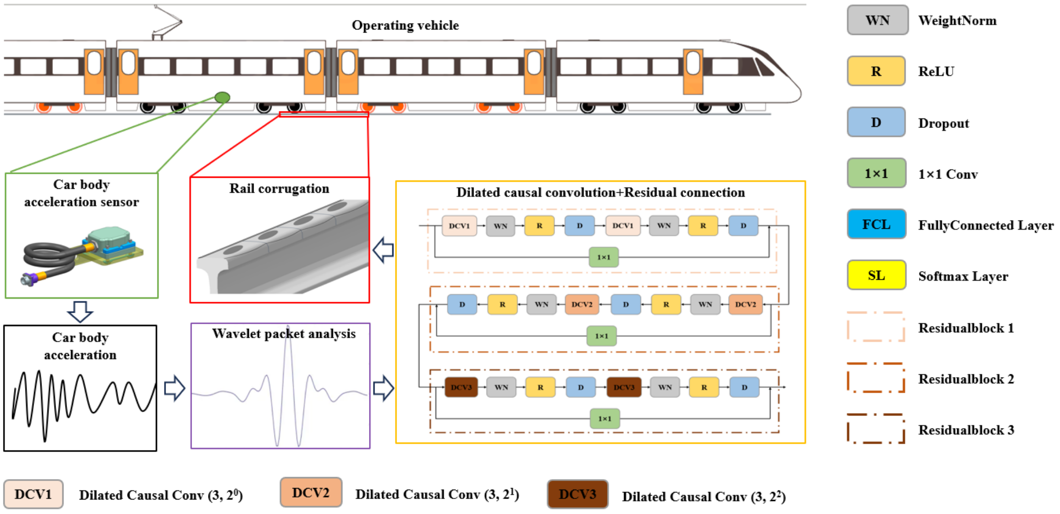


Fig. 3 Overview of the study.

2 Methodology

In this section, the specific methodology used for detecting rail corrugation with the LCRTR-Net is detailed. This approach consists of several key steps. First, wavelet packet decomposition is performed on the car body acceleration signals to extract representative sub-signals. Next, wavelet packet reconstruction techniques are employed to rebuild the signals, ensuring that the features of interest are emphasized. Then, the reconstructed signals are input into the LCRTR-Net, where relevant features are extracted through dilated causal convolution layers. Finally, the classification layer integrates these features to produce identification results related to rail corrugation. The specific implementations and technical details of each step are elaborated upon.

2.1 Wavelet packet decomposition

Wavelet packet decomposition is a multi-resolution analysis method that can decompose a signal into sub-signals of different frequency bands. The process of wavelet packet decomposition can be achieved by recursively applying low-pass and high-pass filters. Let $x(t)$ be the signal to be decomposed, using the wavelet basis function $\psi(t)$ and the scaling function $\phi(t)$ for decomposition. The mathematical expressions for wavelet packet decomposition are as follows:

Definitions of scaling and wavelet functions [42]:

$$\phi(t)=\sum_{n}h[n]\phi(2t-n) \tag{1}$$

$$\psi(t)=\sum_{n}g[n]\phi(2t-n) \tag{2}$$

Here, $h[n]$ and $g[n]$ are the coefficients of the low-pass and high-pass filters, respectively.

Recursive relationship for wavelet packet decomposition:

For the j-th level of decomposition, the expressions for the scaling and wavelet functions are [42]:

$$S_{j,k}=\sum_{n}h[n]S_{j-1,k}(2t-n) \tag{3}$$

$$D_{j,k}=\sum_{n}g[n]S_{j-1,k}(2t-n) \tag{4}$$

Here, $S_{j,k}$is the k-th scaling component at the j-th level, and $D_{j,k}$ is the k-th wavelet component at the j-th level.

Construction of the wavelet packet space:

The wavelet packet space is constructed from all combinations of scaling and wavelet functions, which can be expressed as [42]:

$$W_p=\{S_{j,k},D_{j,k}:j\in\square,k\in\square\} \tag{5}$$

The implementation details and specific parameter choices for wavelet packet decomposition are as follows:

(1) Filter selection: Select Meyer as the wavelet basis function. The Meyer wavelet basis function has the following characteristics:

1) Smoothness: The Meyer wavelet is a smooth wavelet with infinitely many continuous derivatives. This means that its variations are very smooth at all points, making it suitable for processing signals that require high smoothness.

2) Support: The support of the Meyer wavelet is finite, but its main support interval is wider than that of many other wavelets. This characteristic allows it to capture changes in the signal over a larger range.

3) Frequency Domain Characteristics: The Meyer wavelet performs exceptionally well in the frequency domain. It has good localization properties, allowing it to effectively capture both high-frequency and low-frequency features of the signal. The frequency spectrum of the Meyer wavelet is smooth, resulting in more accurate analysis of frequency components.

4) Orthogonality

The Meyer wavelet exhibits orthogonality, meaning that the wavelet functions at different scales and positions are orthogonal to each other. This property allows for the independent representation of different components of the signal in wavelet transforms.

The advantages of the Meyer wavelet in signal processing primarily lie in its excellent denoising capability and feature extraction performance. Due to its support and orthogonality, the Meyer wavelet can effectively separate useful information from noise in vehicle body acceleration signals, achieving efficient signal denoising. Furthermore, its smoothness enhances the accuracy of the features extracted for pattern recognition and classification tasks, thereby improving the overall performance of the model.

Additionally, the Meyer wavelet exhibits strong flexibility and localization analysis capabilities in multiresolution analysis. By analyzing signals at different scales, it can capture subtle changes in non-stationary signals (such as the acceleration signals of vehicles caused by rail wear) and effectively process them across various frequency ranges. This characteristic not only improves computational efficiency but also provides a fine analysis tool for vehicle body acceleration signals, allowing it to excel in various application scenarios.

(2) Decomposition level: Choose a decomposition level of 2. This choice provides sufficient frequency resolution while avoiding excessive computational complexity. With a decomposition level of 2, the signal can be divided into 4 different frequency bands, effectively capturing high-frequency noise while retaining the trend of low-frequency signals. This is suitable for analyzing the features of rail corrugation and ensures the reliability of the analysis results.

(3) Boundary processing: Use mirror processing to handle signal boundaries. This method effectively reduces the impact of boundary effects on the decomposition results, avoiding artifacts at the edges of the signal. By employing mirror processing, the continuity of the signal is maintained, ensuring the accuracy of the resulting frequency bands and enhancing the reliability of subsequent analyses. Mirror processing is particularly suitable for handling periodic or approximately periodic signals, such as the rail corrugation.

2.2 Wavelet Packet Reconstruction

Wavelet packet reconstruction is the process of recombining the various frequency bands obtained from wavelet packet decomposition back into the original signal. The reconstruction process can be achieved through the following formulas [43]:

$$x(t)=\sum_{j=0}^{J}\sum_{k}S_{j,k}(t)+D_{j,k}(t) \tag{6}$$

Here, J is the number of decomposition levels, and$S_{j,k}(t)$ and $D_{j,k}(t)$ are the scaling and wavelet components at the j-th level, respectively.

Reconstruction process:

For the reconstruction at each level, it can be expressed as [43]:

$$S_{j-1,k}(t)=\sum_{n}h[n]S_{j,k}(2t-n)+\sum_{n}g[n]D_{j,k}(2t-n) \tag{7}$$

The implementation details and specific parameter choices for wavelet packet reconstruction are as follows:

(1) Filter selection: Just like during decomposition, it is essential to select an appropriate wavelet basis during reconstruction. Typically, the same wavelet basis used during decomposition should be employed to ensure the preservation and consistency of signal features throughout the reconstruction process. This study continues to use Meyer , as it exhibits excellent time and frequency localization properties, making it suitable for processing subtle features.

(2) Reconstruction level: During the reconstruction process, the reconstruction level should be consistent with the level used during the initial decomposition. Since J = 2 was chosen during decomposition, the same level should be maintained during reconstruction. This ensures that the various frequency bands from the decomposition can be accurately restored during the reconstruction process.

(3) Boundary processing: Boundary processing is equally important during reconstruction. This study continues to use the mirror processing method to reduce the impact of boundary effects on the reconstruction results, ensuring the continuity and accuracy of the signal.

2.3 Dilated causal convolution

Dilated convolution is a special convolution operation that increases the receptive field without increasing the computational cost by introducing gaps (i.e., dilation) between the elements of the convolution kernel. The basic idea is that during convolution, instead of applying all elements of the kernel adjacently to the input, the spacing between the elements in the kernel is controlled by a dilation factor *r*.

Given an input sequence *x* and a convolution kernel *w*, the output *y* of the dilated convolution can be represented as [44]:

$$y[t]=\sum_{i=0}^{k-1} x[t-r\cdot i]\cdot w[i] \tag{8}$$

Where, *t* is the current output position. *k* is the size of the convolution kernel. *r* is the dilation factor, indicating the spacing between adjacent elements in the convolution kernel.

Causal convolution is a special type of convolution that ensures the output only depends on the current and previous inputs. The output of causal convolution can be represented as [45]:

$$y[t]=\sum_{i=0}^{k-1} x[t-i]\cdot w[i] \tag{9}$$

Where, $x[t-i]$ depends only on the current and previous inputs.

Dilated causal convolution combines the features of dilated convolution and causal convolution. The formula is as follows [45]:

$$y[t]=\sum_{i=0}^{k-1} x[t-r\cdot i]\cdot w[i], t\geq r\cdot(k-1) \tag{10}$$

The implementation details and specific parameter choices for dilated causal convolution are as follows:

(1) Kernel size: The kernel size is chosen to be 3. Smaller kernels are effective in capturing local features in signals, especially when dealing with time series data. Small kernels focus on subtle changes in the signal, avoiding feature blurring, and are computationally more efficient, reducing the number of model parameters and thereby lowering the risk of overfitting.

(2) Dilation factor: The dilation factor is chosen to be 2^0, 2^1, 2^2, 2^3, etc. By using exponentially increasing dilation factors, the receptive field can be gradually expanded, allowing the network to capture a broader context. This is particularly important for processing long time series data, as features in the signal may manifest over longer time spans. Increasing the dilation factor helps improve the model's ability to recognize complex patterns.

(3) Padding strategy: For causal convolution, the forward padding size is set to *k* - 1; for dilated causal convolution, the padding size is set to (*k* - 1)×r. This padding strategy ensures the causality of the model, meaning that the output depends only on the current and previous inputs, which is suitable for time series prediction tasks. Additionally, the padding adjustment in dilated causal convolution can prevent distortion at the signal boundaries, maintaining the integrity of the signal and avoiding negative impacts on model performance due to boundary effects.

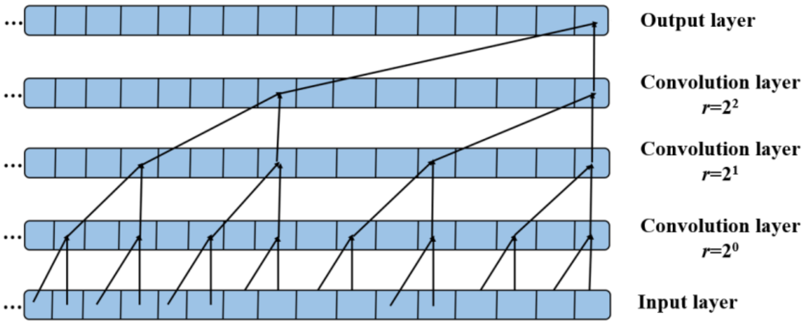


Fig. 4 Dilated causal convolution.

2.4 The architecture of LCRTR-Net

LCRTR-Net (an expanded causal convolutional neural network for track irregularity detection) consists of multiple modules designed to extract features related to track irregularities from the car body acceleration signals. The entire network architecture can be divided into the following main parts, as shown in Fig. 5.

Part I Input Layer:

The input to LCRTR-Net is the car body acceleration signal A, which is decomposed using wavelet packet decomposition to obtain the reconstructed signal $A_{corrugation}$ used for subsequent processing.

Part II Wavelet Packet Decomposition Module:

Wavelet packet decomposition (WPD): This module employs a 2-layer wavelet packet decomposition, breaking down the input car body acceleration signal A into four sub-signals: $S1$,$S2$,$S3$,$S4$. Among these, $S1$ represents acceleration variations caused by the original irregularities of the track, while $S2$,$S3$,$S4$ are used for further analysis. The decomposition formula is [42]:

$$S_j = WPD(A, j) \quad (j = 1, 2, 3, 4) \tag{11}$$

Part III Feature Selection and Reconstruction Module:

Feature selection:Only $S2$,$S3$,$S4$ are selected for reconstruction to eliminate the impact of original irregularities on the car body acceleration.

Wavelet packet reconstruction (WPR): The selected signals are reconstructed to obtain the final input to the network, $A_{corrugation}$. The reconstruction formula is [43]:

$$A_{corrugation} = WPR(S2, S3, S4) \tag{12}$$

Part IV Dilated Causal Convolution Layers:

Dilated causal convolution: This layer is the core of LCRTR-Net, utilizing dilated convolutions to capture features at different scales. Dilated convolutions allow the network to expand its receptive field without increasing computational complexity. Each convolution layer uses a kernel size of 3×1, and each residual block contains 64 convolution kernels.

Dilation factors: The dilation factors used in the network are 2^0, 2^1, 2^2, corresponding to three residual blocks. This enables the network to effectively capture features at different temporal scales.

Part V Residual Connections:

Residual blocks: Each residual block consists of two dilated causal convolution layers with residual connections to alleviate training issues in deep networks. The residual connection helps the model better learn complex functions by directly adding the input to the output.

Part VI Regularization and Optimization:

Weight normalization: Weight normalization is applied in the residual blocks to accelerate training and improve the generalization ability of the model.

Dropout: During training, Dropout layers are used to randomly drop a certain percentage of neurons, typically set to a dropout probability of 0.3, to prevent overfitting. This means that during training, 30% of the neurons in the dropout layer are randomly deactivated in each forward pass, helping to improve the model's generalization ability.

Activation function: Rectified Linear Unit (ReLU) is used as the activation function to increase the non-linearity of the network, prevent gradient vanishing, and make the network sparser.

Part VII Fully Connected Layer:

Feature integration: After passing through the three residual blocks, as shown in Fig. 6. The $A_{corrugation}$ enters a fully connected layer to integrate features and perform final classification.

Part VIII Classification Layer:

Softmax classification: Finally, a Softmax layer is used for classification, outputting the probability of categories related to rail corrugation.

Part IX Training Setup:

Hyperparameter settings: Initial learning rate is set to 0.005, with training epochs totaling 1000. The learning rate decay factor is 0.8, and it decays every 50 epochs. The batch size is 128, and the L2 regularization factor is 0.0005. The optimizer used is

Adam.

Experimental environment: The experiments were conducted on a desktop computer equipped with a 12th generation Intel Core i7-12700KF CPU and an RTX 3070 GPU.

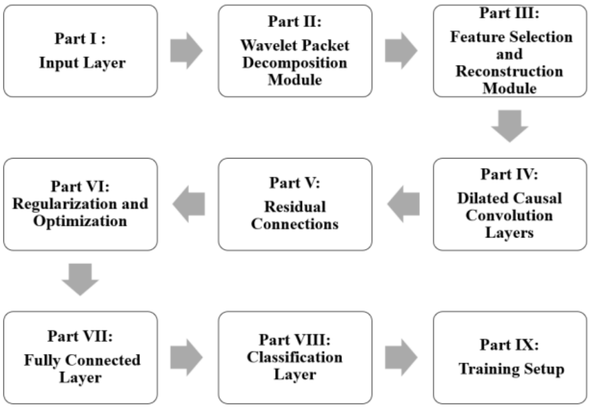


Fig. 5 The architecture of LCRTR-Net.

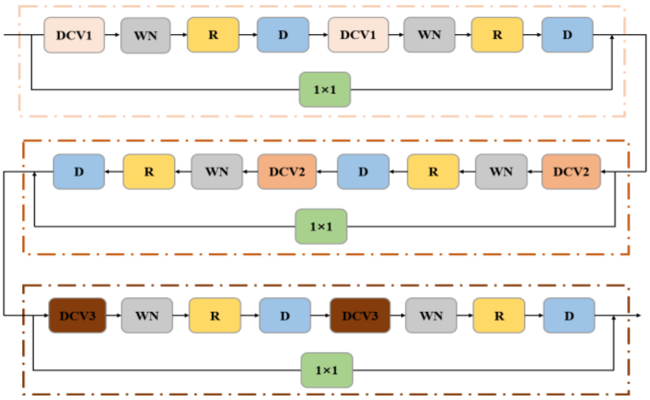


Fig. 6 Residual blocks

2.5 LCRTR-Net Tuning Process

The process of model adjustment and parameter selection mechanisms includes the following aspects:

(1) Kernel Selection:

1) Number and Size of Kernels: A total of 64 kernels, each sized 3×1, are employed in the dilated causal convolution layers. This configuration is optimal based on preliminary experiments, effectively balancing robust feature representation with computational efficiency.

2) Rationale for Kernel Size: The choice of a 3×1 kernel is deliberate, aimed at capturing local features within acceleration signals. This design is essential for discerning the nuances in the data while keeping the computational load manageable. By focusing on local feature extraction, we maintain an appropriate level of model complexity.

(2) Wavelet Packet Decomposition:

1) Implementation of WPD: A 2-layer WPD is conducted, yielding four sub-signals: S1, S2, S3, and S4.

2) Selection of Wavelet Functions: The analysis focuses on wavelet functions S2, S3, and S4. This selection is based on their effectiveness in representing varying frequencies in car body acceleration data. By minimizing noise associated with S1, which mainly reflects original track irregularities, the model can enhance its analytical accuracy.

(3) Dilation Rates:

1) Chosen Dilation Factors: Dilation rates of 2^0, 2^1, and 2^2 are selected to augment the model's capacity to capture features across multiple temporal scales.

2) Receptive Field Expansion: These dilation factors expand the receptive field of the model, integrating long-range dependencies within the data without a significant increase in parameters. This capability is crucial for understanding the temporal dynamics of acceleration signals.

(4) Training Configuration:

1) Learning Rate: An initial learning rate of 0.005 is set, providing a conservative yet effective starting point for stable convergence during training.

2) Training Epochs: Training spans 1000 epochs, allowing ample opportunity for the model to learn while closely monitoring for potential overfitting. Care is taken to prevent excessive training that could compromise model integrity.

3) Learning Rate Decay: A decay factor of 0.8 is applied to the learning rate every 50 epochs. This gradual reduction aids stability and allows the model to fine-tune its parameters as it converges.

4) Batch Size: A batch size of 128 is selected to achieve a balance between computational efficiency and the model's generalization ability. This size provides a representative sample without overwhelming the training process.

5) Regularization: An L2 regularization factor of 0.0005 is implemented to mitigate overfitting, particularly in complex models. This penalty encourages the model to maintain simpler representations by discouraging large weights.

6) Optimizer: The Adam optimizer is chosen for its adaptive learning rate capabilities, which typically accelerate convergence by adjusting the learning rate based on gradient moment estimates. This adaptability is particularly advantageous in scenarios with varying data distributions.

2.6 The process of rail corrugation detection

The steps for LCRTR-Net to identify rail corrugation can be divided into the following steps, as shown in Fig. 7.

Step 1. Data Acquisition: Acceleration signals from the car body are collected during train operation using acceleration sensors. These signals reflect the vibration characteristics caused by track irregularities, providing the foundational data for subsequent analysis.

Step 2. Wavelet Packet Decomposition: The collected acceleration signals undergo 2 levels of wavelet packet decomposition to obtain multiple sub-signals. Sub-signals related to rail corrugation (such as $S2$, $S3$, $S4$) are selected for further analysis to extract more representative features.

Step 3. Feature Reconstruction: The selected sub-signals are reconstructed to obtain the final input signal $A_{reconstructed}$. This process aims to eliminate the impact of original track irregularities on the ca body acceleration, ensuring that the input signal better reflects the characteristics of rail corrugation.

Step 4. Network Architecture: The $A_{corrugation}$ is input into residual blocks, where the network extracts signal features through multiple dilated causal convolution layers. Each convolution layer uses different dilation factors to capture features at various time scales.

Step 5. Feature Integration: After passing through multiple convolution layers, the features are integrated through fully connected layers. This step aims to fuse the extracted feature information to provide effective input for subsequent classification.

Step 6. Classification Layer: A Softmax layer is used to classify the integrated features, outputting the probabilities of various categories related to rail corrugation. This process converts the extracted features into specific identification results for rail corrugation.

Step 7. Model Training: The model is trained using the Adam optimizer, with appropriate learning rates and batch sizes set. Hyperparameters are adjusted to improve the model's accuracy and generalization ability.

Step 8. Recognition: Once training is complete, new acceleration signals are input into the trained LCRTR-Net. The model processes these signals to identify the corresponding categories of rail corrugation based on the features extracted.

Step 9. Output: The model outputs the severity of rail corrugation based on the features of the input signals, providing a basis for subsequent maintenance.

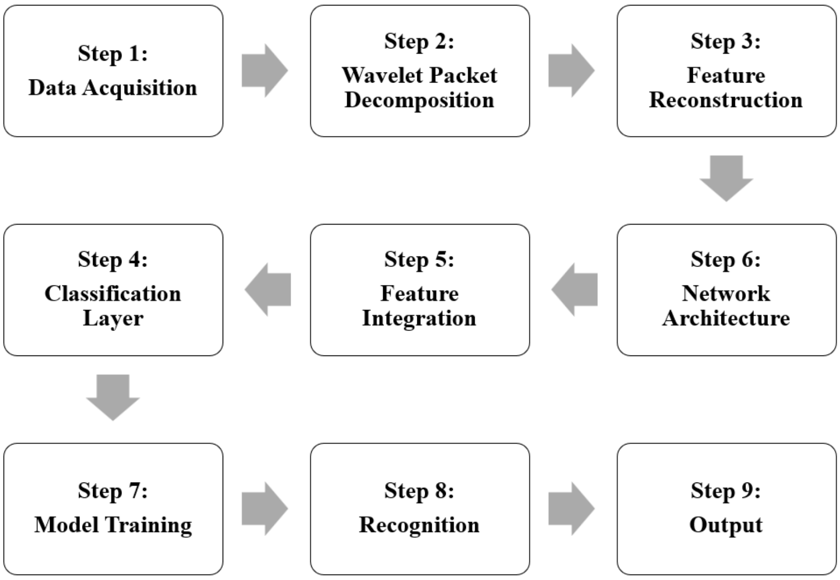


Fig. 7 The process of rail corrugation detection

3 Data acquisition

In this study, the car body acceleration is measured using a acceleration sensor, while the wear of the rail is monitored by Miniprof, as shown in Fig. 9. The car body acceleration sensor is a three-axis sensor based on micro-electromechanical systems (MEMS) technology, capable of measuring acceleration in different directions of the car body. Its measurement rangeis ±2g, with a resolution of up to 16 bits, allowing it to detect minute changes in acceleration. The sensor's sensitivity is generally 1000 mV per g. The operating voltage is usually 5V, accommodating various power configurations, and it has a wide operating temperature range, typically between -40° C and +85° C, making it suitable for various environmental conditions. The sensor's dimensions are generally a few centimeters square, and its weight is typically under 10 grams, facilitating installation in car body. Its protection level is usually IP67, ensuring normal operation in humid or dusty environments. The design includes vibration resistance and complies with relevant electromagnetic compatibility (EMC) standards, allowing stable operation in electromagnetic interference environments.

The main differences between the axle box accelerometer and the body accelerometer are reflected in the following aspects:

(1) Design and Function:

1) The sensor used in this study is specifically designed for monitoring the dynamic response of the vehicle. Its design emphasizes high sensitivity and precision, allowing it to capture subtle changes in performance. It is typically installed on the vehicle's floor, as shown in Fig. 9.

2) In contrast, the axle box sensor is primarily used to monitor the dynamic response of the axle, focusing on its robustness when subjected to high levels of vibration and shock. It is generally installed in the axle box position of the wheel, as illustrated in Fig. 8.

(2) Range and Sensitivity:

1) Range: Axle box accelerometers usually have a larger range (e.g., ±200g or higher), suitable for monitoring significant vibrations and shocks, whereas body accelerometers have a smaller range, typically ±2g, performing better in detecting small dynamic changes.

2) Sensitivity: Axle box accelerometers have a lower sensitivity (e.g., 50 mV/g), making them unsuitable for detecting minor acceleration changes. The sensitivity of body accelerometers is around 1000 mV/g, which is more appropriate for high-resolution and dynamic response applications, such as monitoring the dynamic reactions caused by rail wear.

The reasons for not applying axle box accelerometers in other parts of the vehicle are mainly as follows:

(1) The design of axle box accelerometers is primarily intended for monitoring vibrations and shocks near the wheels. This

positioning is more effective for capturing dynamic behaviors related to the axle, such as vibrations, shocks, and abnormal wear. Therefore, their design and performance characteristics may not be suitable for monitoring needs in other areas.

(2) Axle box accelerometers typically have a larger range, which is suitable for detecting intense vibrations but have lower sensitivity to small acceleration changes. In other areas, especially in applications requiring high sensitivity and detailed monitoring, body accelerometers are more appropriate.

(3) Axle box accelerometers are usually installed in compact spaces within the axle box, where they are more affected by environmental vibrations. In other parts of the vehicle, sensors may face different spatial constraints and environmental conditions, which could impact their performance and the accuracy of the data.

The vibration characteristics and interference signals in other parts of the vehicle may affect the performance of axle box accelerometers. For example, the overall vibrations and noise environment of the body may obscure the subtle changes that the axle box sensor needs to monitor.

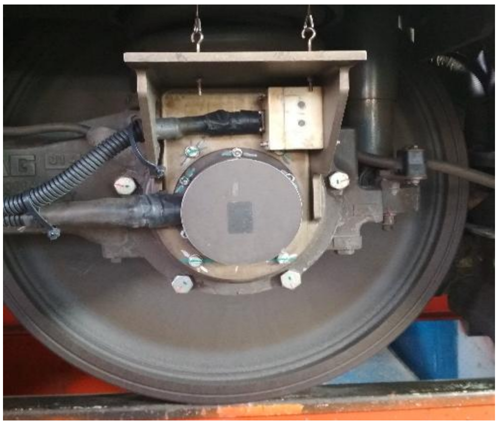


Fig. 8 Axle box accelerometer

To eliminate the influence of vibration on measurement accuracy, this study primarily considers two aspects: sensor installation and data preprocessing.

Sensor installation：

(1) Sensors are installed on the flat and rigid floor of the car body, ensuring they are positioned away from other mechanical components and electrical devices. This choice effectively reduces electromagnetic interference and mechanical vibrations caused by nearby equipment.

(2) High-strength industrial adhesives, such as epoxy resin, are used to secure the sensors to the car body, ensuring that they do not loosen in a vibrating environment. Preliminary experiments are conducted to verify the stability of the adhesive under different temperature and humidity conditions.

(3) To prevent damage to the cables due to vibrations, bolts are used to secure the sensor cables to the car body, ensuring that the cables do not twist or stretch. This measure effectively reduces signal interference caused by cable movement.

Data preprocessing:

(1) Check the dataset for missing values and outliers. Use statistical methods (such as Z-score) to identify outliers [46].

$$Z = \frac{x-\mu}{\sigma} \tag{13}$$

If $|Z|>3$, it is considered an outlier.

(2) Replace the identified outliers with NaN to facilitate further processing. Use linear interpolation to fill in the missing values and outliers, where the formula for linear interpolation is [47]:

$$x_i = x_{i-1} + \frac{(x_{i+1} - x_{i-1})}{2} \tag{14}$$

Where, x_i is the missing or outlier value. $x_{i-1}$and x_{i+1} are the known values before and after the missing or outlier value.

MiniProf is a portable rail profile detection device that features high precision and ease of operation. Its measurement accuracy is better than ±54 μm, with a repeatability of ±20 μm, and it weighs only 0.8kg, making it easy to carry. The device can measure at a speed of up to 10,000 points per second, completing a profile scan in less than 5 seconds, with a processing time of less than 1 second. The generated profile file size is between 20-30 kb. MiniProf is equipped with a magnetic attachment design and is made of high-precision titanium material. It has a built-in data cable that connects via USB, ensuring simple operation. Additionally, the device comes with a shockproof and waterproof transport case, supports railway-specific software, and provides a Chinese interface. It can calculate wear parameters in real-time and compare them with benchmarks, supporting ASCII file formats, and can accommodate up to 500-800 measurement points per profile. This device is particularly suitable for quality control and maintenance of railway tracks.

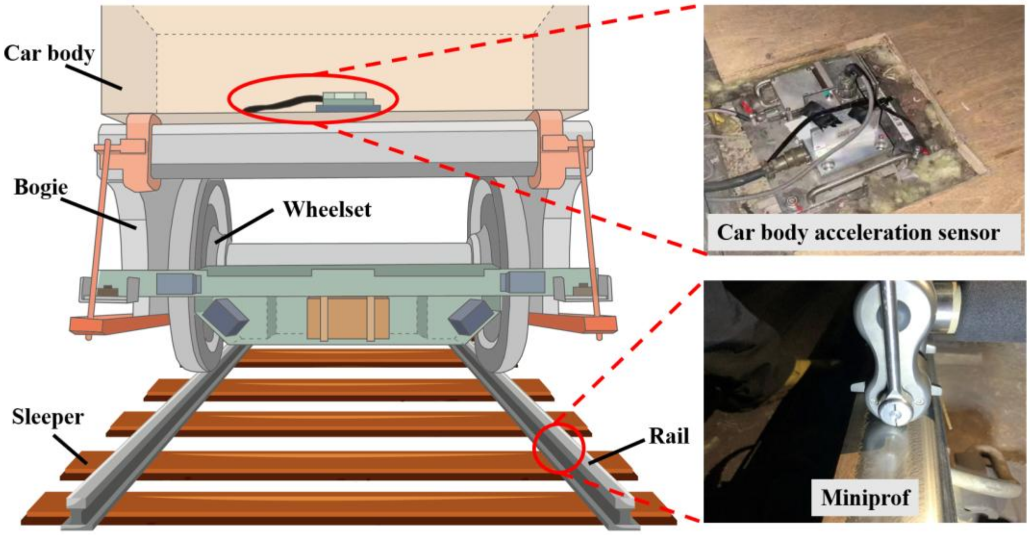
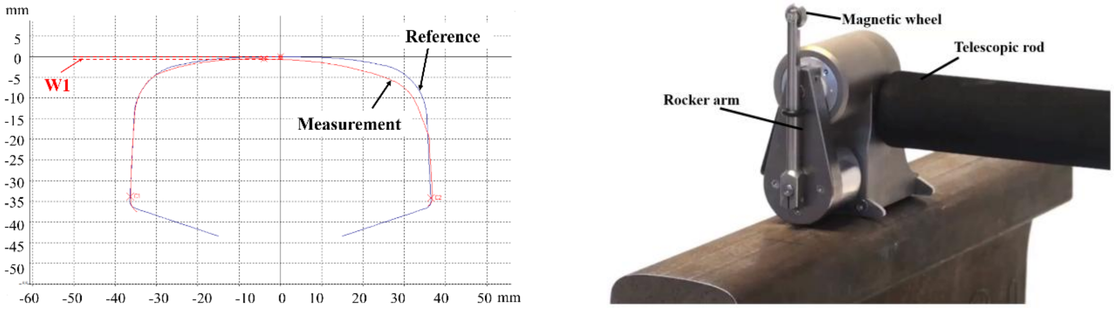


Fig. 9 Data acquisition equipment.

The defects on the rail are primarily distributed on the top surface and the side of the rail, as shown in Fig. 10 (a). The specific dimensions of the rail defects are measured using the MiniProf device, as illustrated in Fig. 10 (b). Vertical wear of the rail refers to the wear value of the top surface along the vertical center line of the rail, denoted as W1. Rail corrugation typically occurs on the rail's top surface, making W1 a suitable indicator in this study. Initially, the top of the MiniProf device is positioned on the rail's edge, ensuring the telescopic rod's tail end aligns perpendicularly to the rail. Subsequently, the rocker arm is rotated, and the magnetic wheel moves slowly along the rail surface until the rail profile measurement is complete.



(a) The measured rail profile is compared with the reference

(b) MiniProf device

Fig. 10 Rail wear measurement using MiniProf.

To determine the amplitude of rail corrugation based on car body acceleration, measurements were taken under different conditions: without corrugation (0 mm amplitude) and with amplitudes of 0.1 mm, 0.2 mm, and 0.3 mm, as shown in Fig. 7. When there is no rail corrugation, the car body acceleration is relatively smooth, showing no obvious burr features; however, when rail corrugation is present, the car body acceleration exhibits small amplitude burr features, which slightly increase with the amplitude of the corrugation.

For each specific amplitude of rail corrugation, each group of samples contains 4,000 sampling points. A 'sampling point' refers to a specific data point at each measurement location along the direction of the track. In this study, the spatial distance between sampling points is 0.04 m. Given that there are 4000 sampling points in each group, and the wavelength of the rail corrugation is 0.1 m, the number of track variations contained in each group is calculated as follows: 4000×0.04÷0.1=1600 times. The dataset consists of a total of 1,152 groups of samples. The distribution of these samples is as follows: 288 groups correspond to 0 mm amplitude (no corrugation), 288 groups to 0.1 mm, 288 groups to 0.2 mm, and 288 groups to 0.3 mm.

To facilitate model training and evaluation, the dataset was divided into a training set and a test set in an 8:2 ratio. For each specific rail corrugation amplitude, the training set contains 230 groups of data, and the test set contains 58 groups of data, as shown in Table 1.

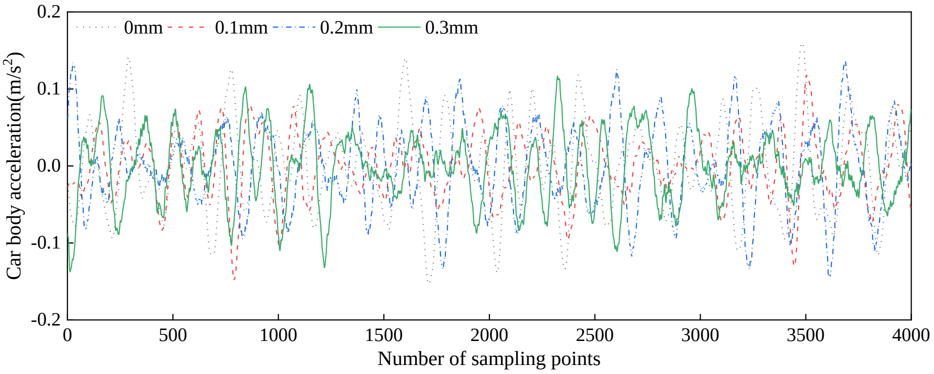


Fig. 11 The measured car body acceleration under different rail corrugation amplitudes.

Table 1 Dataset for rail corrugation recognition.

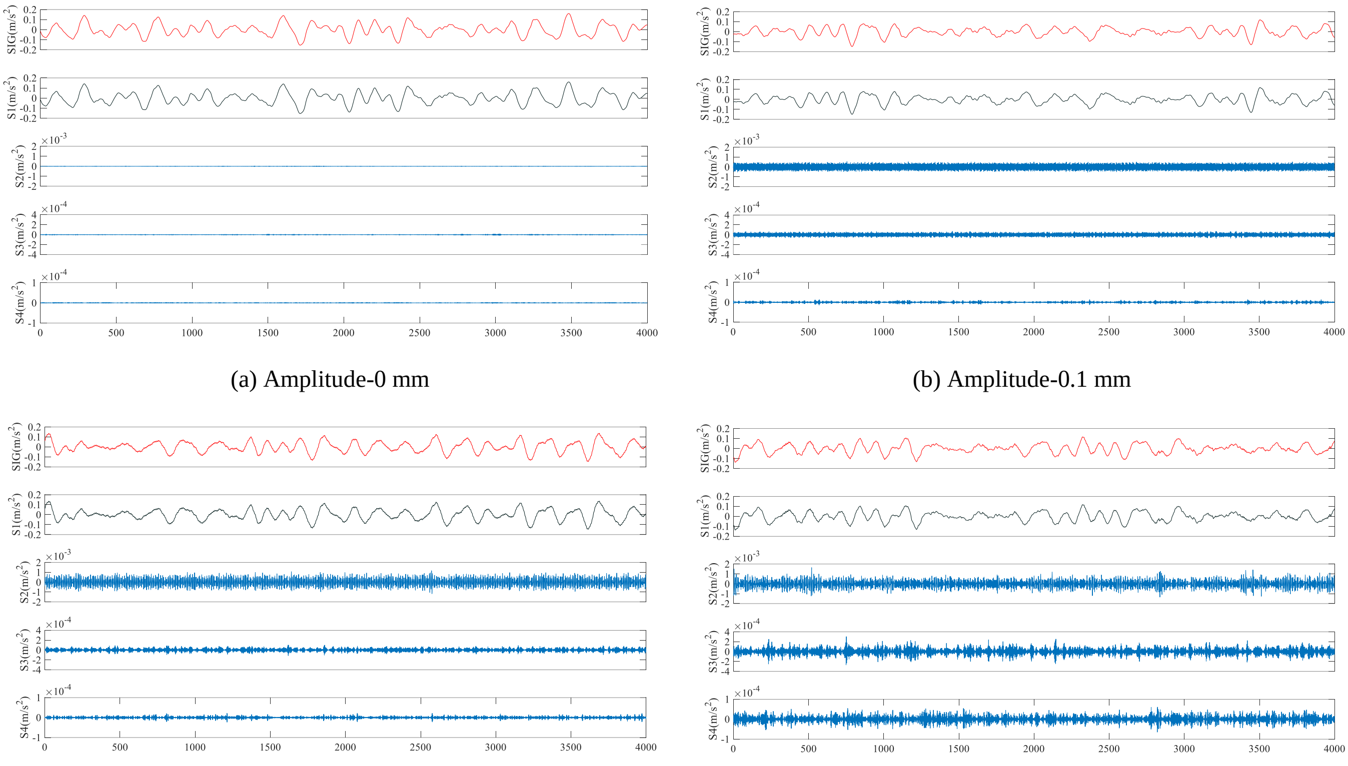
Dataset	Rail corrugation amplitude	Number of groups of car body acceleration
Training set	0mm	230
	0.1mm	230
	0.2 mm	230
	0.3 mm	230
Test set	0mm	58
	0.1mm	58
	0.2 mm	58
	0.3 mm	58

4 Results

4.1 The effectiveness of wavelet packet decomposition and reconstruction

Since the original car body acceleration is primarily influenced by track irregularities, the distinction in car body acceleration across various rail corrugation amplitudes is minimal. Utilizing the original car body acceleration as the input layer of the classifier makes it challenging to discern the amplitudes of rail corrugation. To address this challenge, wavelet packet decomposition and reconstruction methods are employed to preprocess the original car body acceleration.

Taking the car body acceleration at a vehicle speed of 300 km/h as an example, the relationship between the original car body acceleration and the decomposed car body acceleration components at different rail corrugation amplitudes is illustrated in Fig. 12. SIG represents the original car body acceleration, while $S1$ to $S4$ denote the car body acceleration components resulting from wavelet packet decomposition. Specifically, $S1$ corresponds to a low-frequency trend component, primarily reflecting the influence of original track irregularity on car body acceleration. It is noteworthy that $S1$ exhibits minimal variation across different rail corrugation amplitudes. Rail corrugation, categorized as short-wave irregularities, exerts a significant impact on the high-frequency vibrations of car body acceleration, predominantly observed in components *S*2 to *S*4. These components, *S*2 to *S*4, are utilized to reconstruct the car body acceleration.



(a) Amplitude-0 mm

(b) Amplitude-0.1 mm

(c) Amplitude- 0.2 mm

(d) Amplitude-0.3 mm

Fig. 12 Wavelet packet decomposition of car body acceleration at different amplitudes.

The car body acceleration $A_{corrugation}$ caused solely by rail corrugation is illustrated in Fig. 13. After the signal is reconstructed, the acceleration of the car body is amplified by the change of the rail corrugation amplitude. When the difference of rail corrugation amplitude is large, the separability of car body acceleration is enhanced. For example, under the conditions of 0.1 mm and 0.3 mm, the difference of car body acceleration peak is obvious. The LCRTR-Net has a strong adaptive ability, which can capture complex features that are difficult to distinguish from the measured data of large-scale car body acceleration. Therefore, based on the LCRTR-Net, this study eliminates the most disturbing trend in the rail corrugation disturbance signal, and identifies the small fluctuations of the $A_{corrugation}$, so as to achieve the recognition of rail corrugation.

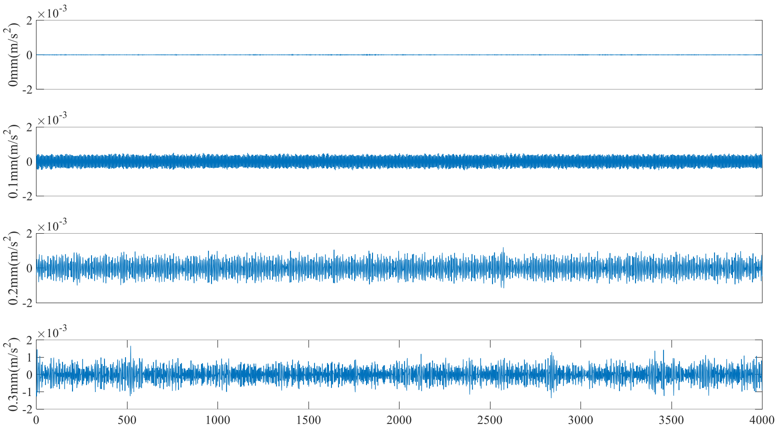
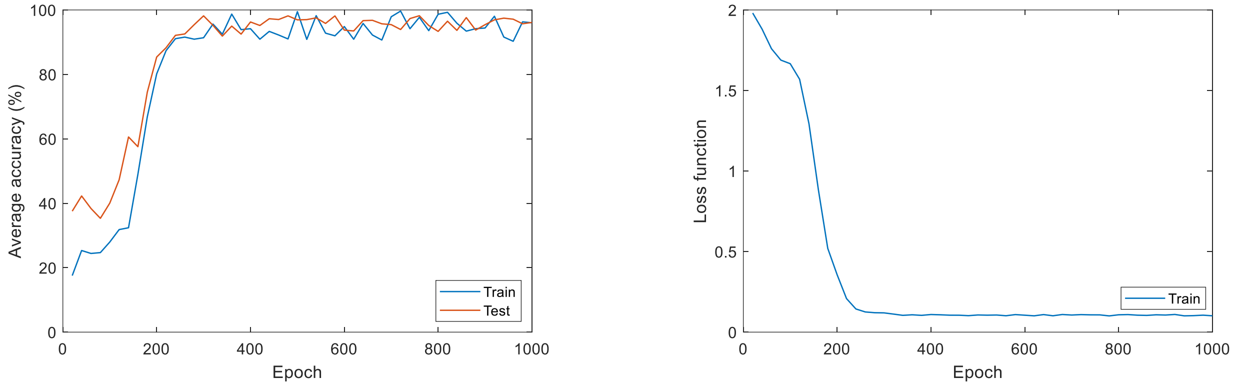


Fig. 13 The car body acceleration $A_{corrugation}$ caused solely by rail corrugation.

4.2 Superiority of LCRTR-Net

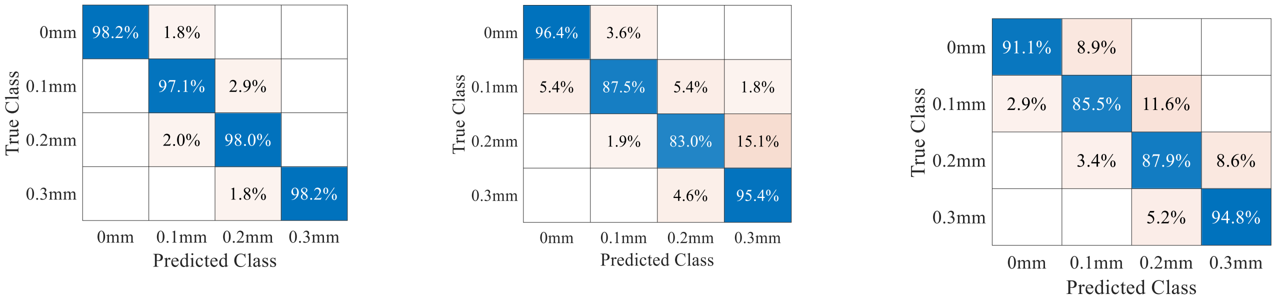
During the training process of the LCRTR-Net, the average accuracy of the training set and test set, as well as the loss of the training set, are shown in Fig. 14. In the initial phase, the loss function is relatively high, around 2.0, and the accuracy is low, only about 20%. At this point, the model has not yet effectively captured the features of the data. As training progresses, the loss function gradually decreases while the accuracy steadily increases, indicating that the LCRTR-Net is gradually converging on both the training and test sets. After 200 epochs, the loss function drops from 2.0 to 0.35, while the accuracy improves from 37% to 85%. Subsequently, the loss function continues to decrease to 0.1, after which no significant changes are observed, and the accuracy fluctuates around 97%.



(a) The average accuracy of train and test sets (b) Loss function of train set

Fig. 14 The average accuracy and loss function of the LCRTR-Net

Among the existing models, Multi-Layer Perceptron (MLP)[56], K-Nearest Neighbor (KNN) [57], Naive Bayesian (NB) [58], and Support Vector Machine (SVM)[35,38] have obvious advantages in training efficiency. In terms of training accuracy, Transformer[52], Long Short Term (LSTM) [53], , Recurrent Neural Network (RNN)[55]， Gate Recurrent Unit (GRU) [54]， Convolutional Neural Network (CNN)[39-40], Naive Bayesian-Convolutional Neural Network (NB-CNN)[36], Empirical Mode Decomposition-Back Propagation (EMD-BP)[34], and Wavelet Packet Decomposition- Support Vector Machine (WPD-SVM[37] perform outstandingly. To illustrate the superiority of LCRTR-Net in rail corrugation recognition, this paper will compare LCRTR-Net with the above existing models from two aspects of training efficiency and training accuracy. This comparison aims to further verify the effectiveness of LCRTR-Net.



	0mm	0.1mm	0.2mm	0.3mm
0mm	98.2%	1.8%		
0.1mm		97.1%	2.9%	
0.2mm		2.0%	98.0%	
0.3mm			1.8%	98.2%

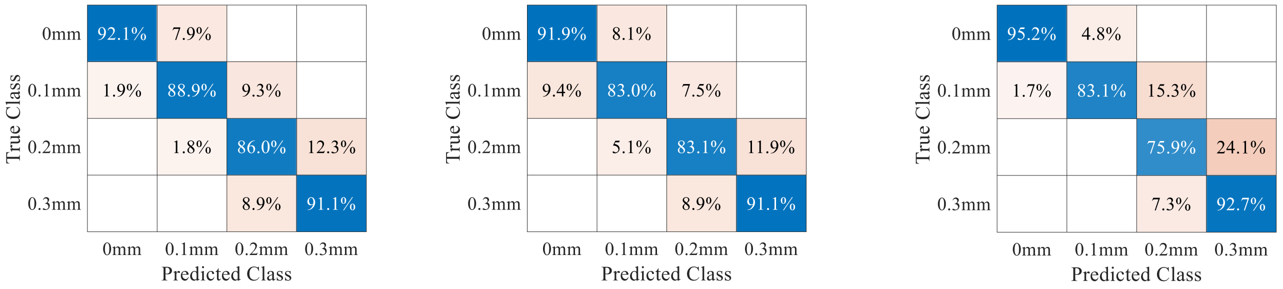
(a) LCRTR-Net

	0mm	0.1mm	0.2mm	0.3mm
0mm	96.4%	3.6%		
0.1mm	5.4%	87.5%	5.4%	1.8%
0.2mm		1.9%	83.0%	15.1%
0.3mm			4.6%	95.4%

(b) Transformer

	0mm	0.1mm	0.2mm	0.3mm
0mm	91.1%	8.9%		
0.1mm	2.9%	85.5%	11.6%	
0.2mm		3.4%	87.9%	8.6%
0.3mm			5.2%	94.8%

(c) NB-CNN



	0mm	0.1mm	0.2mm	0.3mm
0mm	92.1%	7.9%		
0.1mm	1.9%	88.9%	9.3%	
0.2mm		1.8%	86.0%	12.3%
0.3mm			8.9%	91.1%

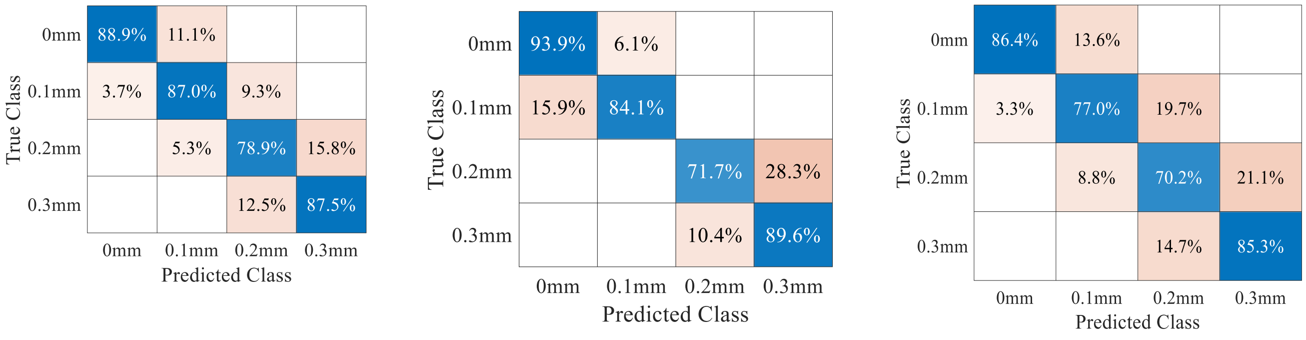
(d) EMD-BP

	0mm	0.1mm	0.2mm	0.3mm
0mm	91.9%	8.1%		
0.1mm	9.4%	83.0%	7.5%	
0.2mm		5.1%	83.1%	11.9%
0.3mm			8.9%	91.1%

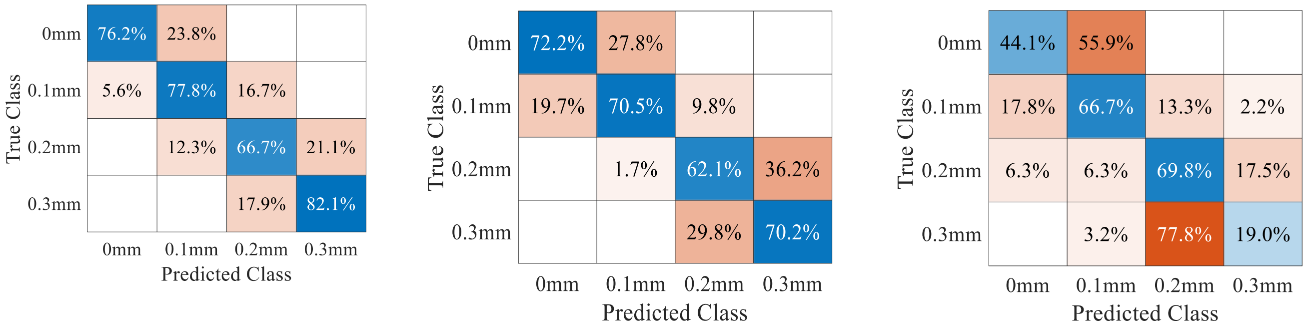
(e) WPD-SVM

	0mm	0.1mm	0.2mm	0.3mm
0mm	95.2%	4.8%		
0.1mm	1.7%	83.1%	15.3%	
0.2mm			75.9%	24.1%
0.3mm			7.3%	92.7%

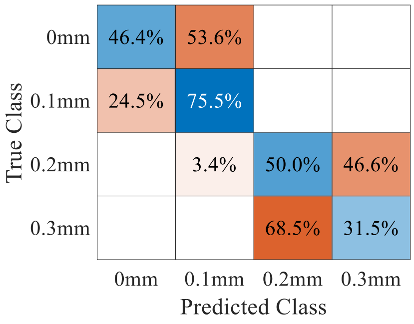
(f) LSTM



(g) CNN (h) GRU (i) RNN



(j)SVM (k)MLP (l)KNN



(m)NB

Fig. 15 Recognition effect of different networks on rail corrugation.

The recognition accuracy of different models at different amplitudes is shown in Fig. 15 and Table 2. LCRTR-Net demonstrates significant advantages across all evaluation metrics, especially in precision, recall, and F1 score, which reached 100%, 98.2%, and 0.991, respectively. The recognition time is only 0.17 ms, significantly lower than that of other models, indicating its potential and advantages in real-time applications.

In contrast, the Transformer model achieved precision and recall rates of 94.7% and 96.4%, respectively. Although its performance is commendable, it falls short of LCRTR-Net in both F1 score and accuracy, with a recognition time of 1.22 ms, indicating weaker real-time processing capability. NB-CNN and EMD-BP perform similarly. Their average accuracy is close to 90%, but it is still less than that of LCRTR-Net, and the recognition time is as long as 3 ms. The average accuracy of WPD-SVM, LSTM and CNN is around 86%, and their recall and F1 scores are significantly lower than those of LCRTR-Net, indicating their limitations in specific tasks.

Further analysis of the GRU, RNN and SVM models reveals relatively weaker performance, particularly in recall and F1 score, indicating their shortcomings in handling complex tasks. The recognition times are also longer, at 2.33 ms and 2.53 ms, respectively. Traditional models such as MLP, KNN, and NB perform significantly worse than deep learning models, with lower precision and recall rates. Although there are differences in recognition times among these traditional models, their real-time capabilities are clearly inferior to those of LCRTR-Net.

To demonstrate the recognition performance of LCRTR-Net when used in conjunction with other algorithms, this study employs the relatively superior Transformer method. The results, shown in Fig. 16 and the last row of Table 2, indicate that while LCRTR-Net-Transformer performs well in terms of precision and recall, with an enhancement percentage of 3.0% over the Transformer, its recognition accuracy is lower than that of LCRTR-Net. The recognition time for LCRTR-Net-Transformer is 339 ms, indicating a significant gap in real-time performance compared to LCRTR-Net.

Overall, LCRTR-Net performs the best across all metrics, particularly in terms of precision, F1 score, and recognition efficiency, demonstrating its significant advantages in practical applications.

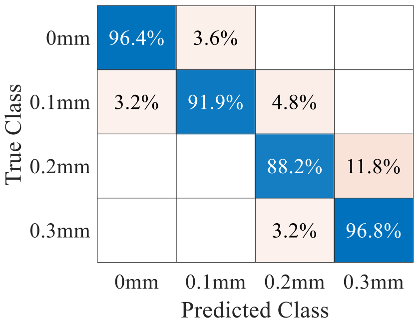


Fig. 16 Recognition effect of LCRTR-Net-Transformer

Table 2 Comparison of the recognition performance of various models under different amplitudes.

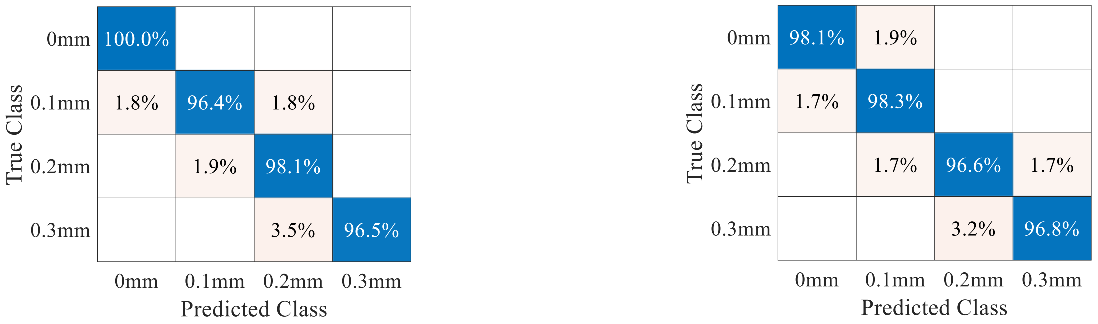
Models	Evaluation metrics				
	Precision (%)	Recall (%)	F1 Score	Average accuracy (%)	Recognition time (ms)
	0mm/0.1mm/ 0.2mm/0.3mm	0mm/0.1mm/ 0.2mm/0.3mm	0mm/0.1mm/ 0.2mm/0.3mm		
LCRTR-Net	100/96.2/95.4/100	98.2/97.1/98/98.2	0.991/0.967/0.967/0.991	97.9	0.17
Transformer	94.7/94.1/89.3/85.0	96.4/87.4/83.0/95.4	0.955/0.906/0.860/0.899	90.6	1.22
NB-CNN	96.9/87.4/84.0/91.7	91.1/85.5/88.0/94.8	0.939/0.865/0.859/0.932	89.8	3.38
EMD-BP	98.0/90.2/82.5/86.6	82.1/87.2/85.9/91.1	0.950/0.887/0.842/0.888	89.5	2.99
WPD-SVM	90.7/86.3/83.5/88.5	91.9/83.1/83.0/81.1	0.913/0.847/0.833/0.898	87.3	2.71
LSTM	98.3/94.5/77.1/79.4	95.2/83.0/75.9/92.7	0.967/0.884/0.765/0.855	86.7	2.27
CNN	96.0/84.1/78.4/84.7	88.9/87.0/78.9/87.5	0.923/0.856/0.786/0.861	85.6	2.08
GRU	85.5/93.2/87.3/76.0	93.9/84.1/71.7/89.6	0.895/0.884/0.788/0.822	84.8	2.33
RNN	96.3/77.5/67.1/80.2	86.4/77.0/70.1/85.3	0.911/0.772/0.686/0.827	79.7	2.53
SVM	93.2/68.3/65.8/79.6	76.2/77.7/66.6/82.1	0.838/0.727/0.662/0.808	75.7	0.98
MLP	78.6/70.5/61.1/66.0	72.2/70.5/62.1/70.2	0.753/0.705/61.6/0.680	68.8	0.44
KNN	64.6/50.5/43.4/49.1	44.1/66.7/70.0/19.0	0.524/0.575/53.5/0.274	49.9	0.21
NB	65.4/57.0/42.2/40.3	46.4/75.5/50.0/31.5	0.543/0.650/45.8/0.354	50.9	0.25
LCRTR-Net-Transformer	96.8/96.2/91.7/89.1	96.4/92.0/88.2/96.8	0.966/0.941/0.899/0.928	93.3	3.39

4.3 Generalization of LCRTR-Net

To verify the generalization capability of LCRTR-Net, the recognition accuracy at different vehicle speeds is analyzed. The confusion matrix for the recognition accuracy of LCRTR-Net at varying speeds is presented in Fig. 17.

The average recognition accuracy of LCRTR-Net for a single amplitude at different speeds is calculated and shown in Fig. 18 (a). The red rectangle represents LCRTR-Net, the dark blue rectangle represents LCRTR-Net-Transformer, and the yellow rectangle represents the Transformer. At an amplitude of 0 mm and vehicle speeds of 200 km/h, 250 km/h, 300 km/h, and 350 km/h, the recognition accuracy of LCRTR-Net is 100%, 96.4%, 98.1%, and 96.5%, respectively. The average recognition accuracy can be calculated as follows: (100%+96.4%+98.1%+96.5%)/4=98.65%. Using the same method, the recognition accuracy of LCRTR-Net for different rail corrugation amplitudes is consistently above 97.0%.

Next, the average recognition accuracy of LCRTR-Net for a single speed under varying rail corrugation amplitudes is calculated. As shown in Fig. 18 (b), when the vehicle speed is 200 km/h and the corrugation amplitudes are 0 mm, 0.1 mm, 0.2 mm, and 0.3 mm, the recognition accuracy is 100%, 96.4%, 98.1%, and 96.5%, respectively. The average recognition accuracy can be calculated as: (100%+96.4%+98.1%+96.5%)/4=97.75%. According to this method, the accuracy of LCRTR-Net in recognizing rail corrugation amplitudes at a vehicle speed of 200 km/h is more than 97.45%.

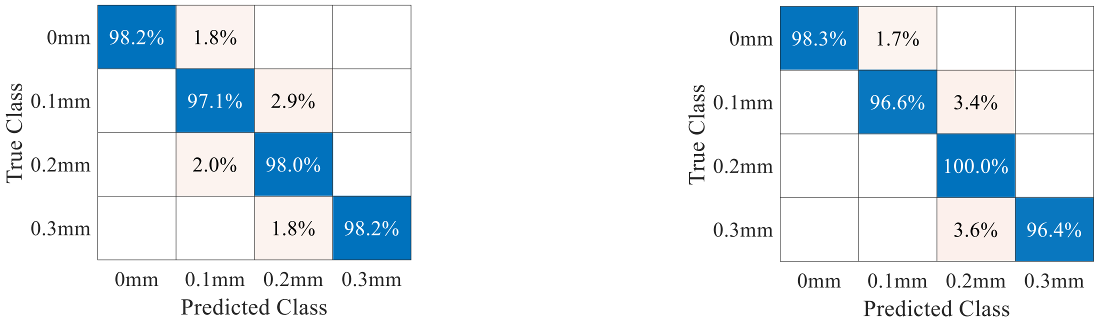


True Class \ Predicted Class	0mm	0.1mm	0.2mm	0.3mm
0mm	100.0%			
0.1mm	1.8%	96.4%	1.8%	
0.2mm		1.9%	98.1%	
0.3mm			3.5%	96.5%

(a) 200 km/h

True Class \ Predicted Class	0mm	0.1mm	0.2mm	0.3mm
0mm	98.1%	1.9%		
0.1mm	1.7%	98.3%		
0.2mm		1.7%	96.6%	1.7%
0.3mm			3.2%	96.8%

(b) 250 km/h



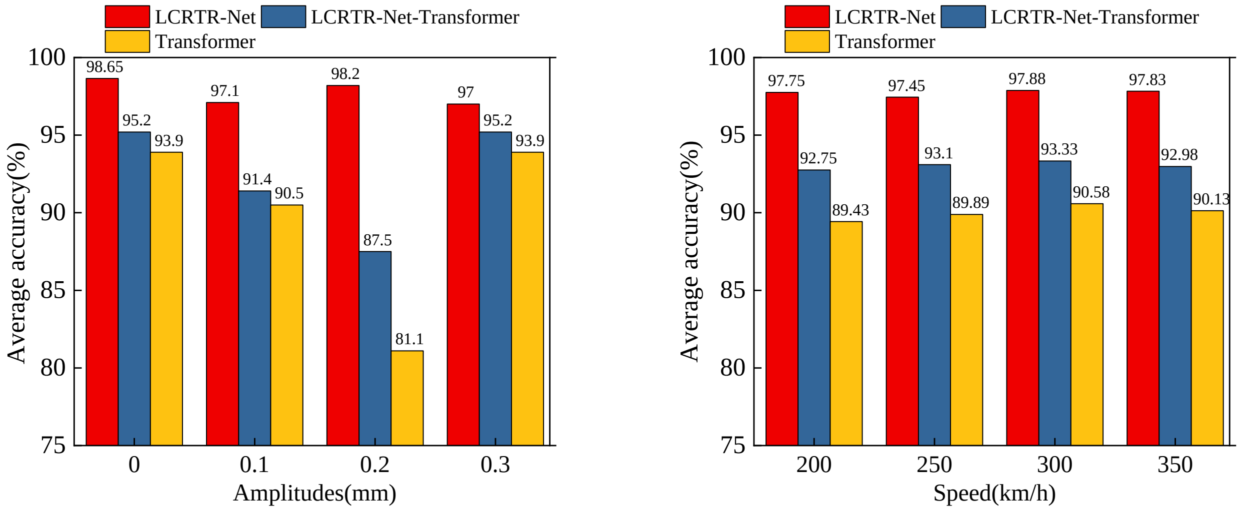
True Class \ Predicted Class	0mm	0.1mm	0.2mm	0.3mm
0mm	98.2%	1.8%		
0.1mm		97.1%	2.9%	
0.2mm		2.0%	98.0%	
0.3mm			1.8%	98.2%

(c) 300 km/h

True Class \ Predicted Class	0mm	0.1mm	0.2mm	0.3mm
0mm	98.3%	1.7%		
0.1mm		96.6%	3.4%	
0.2mm			100.0%	
0.3mm			3.6%	96.4%

(d) 350 km/h

Fig. 17 The confusion matrix of LCRTR-Net



(a) Accuracy of different rail corrugation amplitudes

(b) Accuracy of different vehicle speeds

Fig. 18 Generalization of LCRTR-Net.

When analyzing the recognition performance of different models under varying amplitudes and speed conditions, LCRTR-Net exhibits significant advantages, as shown in Table 3 and Table 4. Specifically:

Under different amplitudes, LCRTR-Net consistently demonstrates superior recognition performance, with accuracy exceeding 97%, which is significantly higher than that of the LCRTR-Net-Transformer and Transformer models. Notably, at amplitudes of 0 mm and 0.3 mm, its precision reaches 100%, and the F1 score exceeds 0.985. These results indicate that LCRTR-Net can maintain a high level of accuracy and recall across various amplitudes, showcasing its effectiveness and robustness in multi-class classification tasks. In contrast, LCRTR-Net-Transformer performs adequately at amplitudes of 0 mm and 0.1 mm, but its precision and F1 score decline significantly at amplitudes of 0.2 mm and 0.3 mm, with the F1 score dropping to 0.918 at 0.3 mm. This suggests that this model lacks recognition capability under more complex amplitude conditions. The Transformer model generally performs poorly, with F1 scores below 0.9 across all amplitudes, particularly at 0.2 mm and 0.3 mm, highlighting its limitations in complex environments.

Under different speeds, LCRTR-Net continues to demonstrate significant advantages, with accuracy, precision, and F1 score all maintained at high levels, specifically with accuracy exceeding 97.45%. Notably, at speeds of 200 km/h and 350 km/h, its F1 score reaches 0.989, indicating the reliability and adaptability of LCRTR-Net under high-speed conditions. In contrast, the performance of LCRTR-Net-Transformer declines significantly at speeds of 300 km/h and 350 km/h, with F1 scores dropping to 0.892 and 0.918, respectively, indicating insufficient adaptability under high-speed conditions. The Transformer model shows relatively consistent performance across all speeds, but overall precision and F1 scores remain lower than those of the other two models, particularly at a speed of 300 km/h, where the F1 score is only 0.848.

In summary, LCRTR-Net demonstrates excellent performance under various amplitudes and speed conditions, proving its strong adaptability and stability in practical applications. In contrast, the recognition performance of LCRTR-Net-Transformer and Transformer is significantly inferior in complex environments.

Table 3 Recognition performance of various models under different amplitudes.

Models	Evaluation metrics		
	Precision (%)	Recall (%)	F1 Score
	0mm/0.1mm/0.2mm/0.3mm	0mm/0.1mm/0.2mm/0.3mm	0mm/0.1mm/0.2mm/0.3mm
LCRTR-Net	100/95.8/95.2/100	97.1/96.8/97.1/97.5	0.985/0.963/0.961/0.987
LCRTR-Net-Transformer	95.5/95.2/90.2/88.3	95.5/91.7/87.3/95.5	0.955/0.934/0.887/0.918
Transformer	93.9/92.7/87.1/83.0	95.3/96.0/82.2/93.7	0.956/0.943/0.846/0.880

Table 4 Recognition performance of various models under different speeds.

Models	Evaluation metrics		
	Precision (%)	Recall (%)	F1 Score
	200(km/h)/250 (km/h) /300 (km/h) /350 (km/h)	200(km/h)/250 (km/h) /300 (km/h) /350 (km/h)	200(km/h)/250 (km/h) /300 (km/h) /350 (km/h)
LCRTR-Net	100/96.0/95.2/100	97.8/97.4/98.3/97.8	0.989/0.967/0.967/0.989
LCRTR-Net-Transformer	96.4/95.8/91.1/88.3	95.5/91.2/87.4/95.5	0.960/0.934/0.892/0.918
Transformer	94.1/92.8/87.1/83.0	95.2/86.0/82.6/93.9	0.947/0.893/0.848/0.935

4.4 Dimensionality reduction ability of LCRTR-Net

The dimensionality reduction capabilities of LCRTR-Net were further analyzed using t-distributed Stochastic Neighbor Embedding (t-SNE), a dimensionality reduction method based on conditional probability. t-SNE is a nonlinear dimensionality reduction technique used for visualizing high-dimensional data. Its goal is to map high-dimensional data into a lower-dimensional space while preserving the local similarities between data points.

Step 1: Compute Similarities in High-Dimensional Space

(1) Similarity Calculation: For each point x_i in the high-dimensional dataset, the similarity with other points x_j is calculated. First, for each point x_i a Gaussian distribution is computed:

$$p_{j|i} = \frac{\exp\left(-\frac{\| x_i - x_j \|^2}{2\sigma_i^2}\right)}{\sum_{k \neq i} \exp\left(-\frac{\| x_i - x_k \|^2}{2\sigma_i^2}\right)} \tag{15}$$

Where, $p_{j\mid i}$ represents the conditional probability of point x_i being a neighbor given point x_j, and σ_i is the parameter controlling the width of the Gaussian distribution.

(2) Symmetrization of Similarity: The conditional probabilities are symmetrized to obtain:

$$p_{ij} = \frac{p_{j|i} + p_{i|j}}{2N} \tag{16}$$

Where, N is the total number of points in the dataset.

Step 2: Construct Low-Dimensional Space

(1) Probability Distribution in Low-Dimensional Space: In the low-dimensional space, similarities are computed using a t-distribution. For the low-dimensional representations y_i and y_j, the similarity is given by:

$$q_{ij} = \frac{(1 + \| y_i - y_j \|^2)^{-1}}{\sum_{k \neq l} (1 + \| y_i - y_k \|^2)^{-1}} \tag{17}$$

(2) Optimization Objective: The low-dimensional embedding is optimized by minimizing the Kullback-Leibler divergence between the probability distributions in the high-dimensional and low-dimensional spaces:

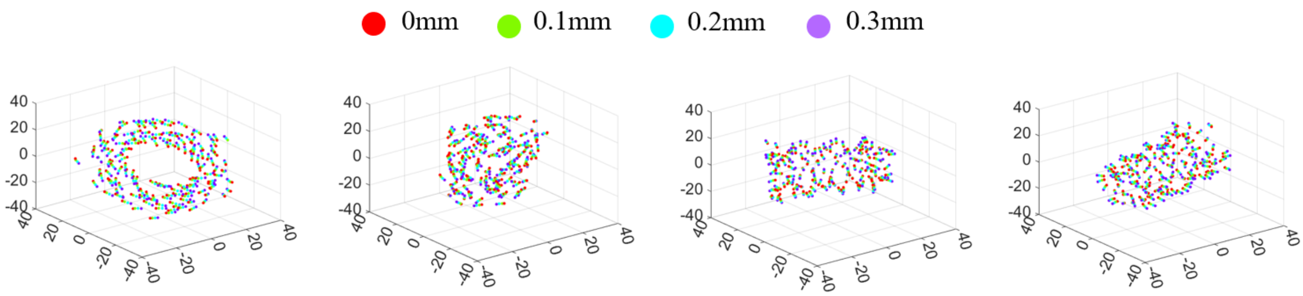
$$C = KL(P \| Q) = \sum_{i \neq j} p_{ij} \log \frac{p_{ij}}{q_{ij}} \tag{18}$$

In this objective, P is the similarity distribution in the high-dimensional space, and Q is the similarity distribution in the low-dimensional space.

t-SNE offers several key advantages for visualizing high-dimensional data. Firstly, it effectively handles nonlinear structures, preserving local neighborhood information, unlike linear methods like PCA, which may overlook local relationships. By finely modeling similarities in high-dimensional space, t-SNE can clearly reveal clusters and patterns within the data, aiding in the identification of potential classifications. Secondly, t-SNE provides excellent visualization results, where similar data points are tightly clustered while points from different categories remain distant, allowing for a more intuitive understanding of the underlying relationships in the data.

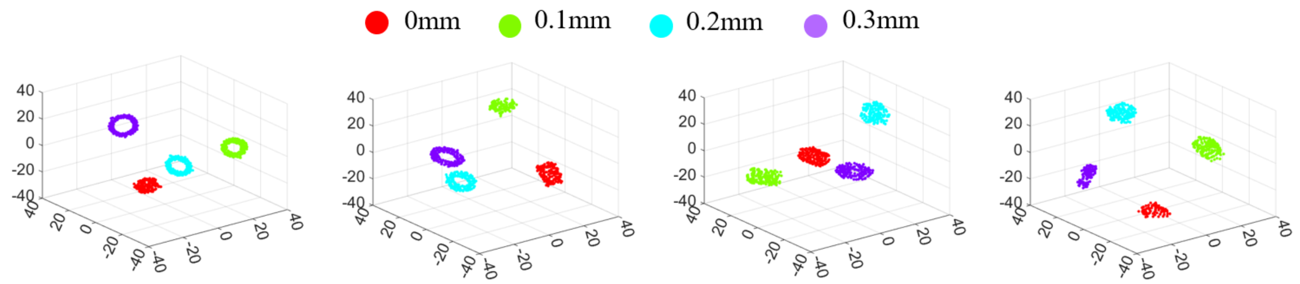
Using the original car body acceleration as the input for the t-SNE model, the three-dimensional visualization results are shown in Fig. 19. It is evident that when the amplitude of rail corrugation ranges from 0 mm to 0.3 mm, there is a strong similarity among the original car body acceleration data corresponding to different rail corrugation amplitudes, making it challenging to distinguish between them in three-dimensional space.

In contrast, using the car body acceleration classified by LCRTR-Net as the input for the t-SNE model, the three-dimensional visualization results are shown in Fig. 20. After treatment with LCRTR-Net, the car body acceleration data corresponding to different rail corrugation amplitude values are more dispersed in three-dimensional space, allowing for clearer differentiation. These findings indicate that LCRTR-Net possesses strong dimensionality reduction capabilities.



(a) 200 km/h (b) 250 km/h (c) 300 km/h (d) 350 km/h

Fig. 19 Dimensionality reduction visualization of the original car body acceleration.



(a) 200 km/h (b) 250 km/h (c) 300 km/h (d) 350 km/h

Fig. 20 LCRTR-Net processed car body acceleration dimensionality reduction visualization.

4.5 The cost analysis of LCRTR-Net

This study defines "low cost" based on several key factors compared to traditional track waviness recognition methods. These factors include:

(1) Hardware Costs: LCRTR-Net utilizes accelerometers installed at the bottom of the train, which are significantly less expensive than complex machine vision systems that require substantial hardware investment. Specifically, commonly used high-precision cameras and image processing equipment on the market typically cost several thousand to tens of thousands of RMB, while the price of body accelerometers generally falls within the range of a few hundred RMB. This presents a clear cost advantage for large-scale deployment.

(2) Installation and Maintenance Costs: LCRTR-Net relies on the installation of accelerometers at the bottom of the train, making the installation process relatively simple and not requiring specialized electromechanical systems or complex calibration procedures. Additionally, since body accelerometers generally require less maintenance, our solution also reduces ongoing maintenance costs. In contrast, systems based on high-precision cameras or axle box accelerometers are not only more susceptible to environmental conditions but also increase the complexity and costs of subsequent maintenance.

(3) Operational Costs: In terms of data processing, LCRTR-Net is capable of real-time processing. By achieving real-time performance (0.17 milliseconds per sample) and high accuracy (over 97%), LCRTR-Net reduces the need for extensive manual inspections and associated labor costs. This efficiency translates into cost savings over time, as the resources required for monitoring track conditions are reduced. Furthermore, due to the enhancement in real-time performance, we are able to respond quickly and process monitoring data, thereby lowering potential safety hazards and saving on possible accident costs in overall operations.

4.6 The potential applications of LCRTR-Net

The potential applications of LCRTR-Net and its impact on real-world railway maintenance and safe operations are primarily reflected in the following five aspects:

(1) Track condition monitoring:

LCRTR-Net utilizes deep learning algorithms to analyze track data collected from sensors, enabling real-time monitoring of track conditions. This data may include irregularities such as track geometry, rail waviness, cracks, and deformations. Through high-frequency data collection and analysis, railway operators can identify potential risks before issues worsen. For example, when LCRTR-Net detects that rail waviness exceeds safety thresholds, it can promptly issue alerts, prompting maintenance teams to conduct on-site inspections and repairs. This real-time monitoring not only enhances the speed of problem response but also improves the overall safety of the railway system.

(2) Data-driven maintenance decisions:

The data generated by LCRTR-Net provides a scientific basis for decision-making in railway maintenance departments. By analyzing historical data and real-time monitoring results, maintenance teams can more accurately assess the health of the tracks, leading to the formulation of reasonable maintenance plans. For instance, maintenance teams can decide whether to carry out routine preventive maintenance or take emergency repair measures based on detected waviness levels and track wear conditions. This data-driven decision-making process not only enhances maintenance efficiency but also optimizes resource allocation, ensuring the rational use of manpower and materials, thereby reducing operational costs.

(3) Extending track lifespan:

The timely detection and handling capabilities of LCRTR-Net effectively extend the lifespan of the tracks. Through regular monitoring and maintenance, railway companies can promptly identify and repair minor issues, preventing them from evolving into larger failures. For example, small cracks detected early, if not addressed in time, could lead to a complete structural failure of the track, resulting in more severe safety hazards. By implementing effective maintenance strategies, railway companies can reduce the frequency of large-scale repairs, thus lowering overall maintenance costs while extending the usable life of the tracks and improving return on investment.

(4) Enhancing passenger safety:

Safety is the top priority in railway transportation. The application of LCRTR-Net can significantly reduce the risk of accidents caused by track damage. Through high-precision track inspections, railway operators can ensure the safety of the tracks before train operations, avoiding accidents caused by uneven or damaged tracks. For instance, LCRTR-Net can monitor abnormal conditions in real-time and promptly notify the train dispatch center for adjustments, ensuring safe train operations. This proactive safety management measure not only enhances the travel experience for passengers but also builds public trust in railway transportation.

(5) Promoting smart railway development:

The application of LCRTR-Net marks an important step towards the intelligent development of the railway industry. By combining advanced monitoring technology with data analysis, railway operators can achieve more efficient operational management. For example, LCRTR-Net can be integrated with other intelligent systems (such as wheel condition monitoring, weather information, etc.) to form a comprehensive monitoring network. This intelligent system can analyze various data in real-time and dynamically adjust operational and maintenance strategies, thereby improving overall transportation efficiency. Additionally, smart railway systems can enable remote monitoring and automated management, further reducing labor costs and enhancing management efficiency.

In summary, the application of LCRTR-Net in track monitoring and maintenance not only improves the accuracy and timeliness of track inspections but also brings higher safety standards and management efficiency to the railway industry. Through data-driven decision-making and intelligent management, railway companies can achieve sustainable development, enhancing passenger safety and satisfaction.

5 Conclusions

In this study, LCRTR-Net is proposed for recognizing rail corrugation using car body acceleration, offering a low-cost and real-time recognition method. The generalization, superiority, and dimensionality reduction capabilities of LCRTR-Net are demonstrated. The conclusions are as follows:

(1) LCRTR-Net comprises three main modules: wavelet packet decomposition and reconstruction, dilated causal convolution, and residual connection. Wavelet packet decomposition and reconstruction effectively eliminate the trend component of car body acceleration caused by random irregularities, allowing for the extraction of the underlying characteristics of car body acceleration related to rail corrugation. The dilated causal convolution integrates the benefits of causal and dilated convolutions, enabling the capture of temporal causality and long-term dependencies within the potential features of car body acceleration. The residual connection enhances the feature extraction capability of LCRTR-Net while ensuring computational efficiency.

(2) The car body acceleration is decomposed using wavelet packet analysis. The $S1$ component primarily represents the impact of track random irregularities on car body acceleration, while the high-frequency vibrations associated with rail corrugation significantly influence car body acceleration, mainly reflected in components $S2$ to $S4$. Consequently, $S2$ to $S4$ are utilized to reconstruct the car body acceleration, which serves as input to the dilated causal convolution neural network.

(3) LCRTR-Net has significant advantages, strong generalization ability, computational efficiency, and dimensionality reduction capability. Compared with existing methods, LCRTR-Net achieves a recognition accuracy of over 97.0% for a single amplitude of rail corrugation, and over 97.45% for all amplitudes of rail corrugation at different speeds. The recognition of each rail corrugation sample takes only 0.17 ms, enabling real-time identification of rail corrugation. LCRTR-Net-Transformer performs well in terms of precision and recall, with an enhancement percentage of 3.0% over the Transformer. However, the accuracy of the enhanced model does not surpass that of LCRTR-Net. After processing with LCRTR-Net, the car body acceleration induced by different rail corrugation amplitudes is more dispersed and does not overlap in three-dimensional space, indicating its strong dimensionality reduction ability.

In addition, this paper discusses future research directions, the limitations of the model, and potential improvements, as follows.

Future research directions:

(1) Model Improvement: We will explore how to adjust and optimize the LCRTR-Net model for different track conditions (such as varying track materials, levels of wear, and environmental factors) to enhance its adaptability and accuracy.

(2) Data Fusion: We will discuss how to integrate data from other sensors (such as temperature, humidity, and track geometry data) to improve the detection capabilities for rail corrugation. This data fusion can provide a more comprehensive understanding of track conditions, thereby enhancing detection accuracy and reliability.

Limitations:

(1) While our study contributes significantly to the field of rail corrugation detection, it is important to acknowledge the reliance of deep learning approaches on large quantities of labeled rail corrugation data. This dependency presents a challenge, as obtaining such data can be resource-intensive and time-consuming.

(2) Before conducting rail corrugation detection, the body acceleration sensors must be installed on vehicles with intact wheels to avoid the influence of wheel conditions on the detection results.

Potential improvements:

(1) Employing more advanced deep learning techniques (such as data augmentation, unsupervised learning, transfer learning, or self-supervised learning) to address the challenges posed by limited labeled data.

(2) By integrating wheel condition monitoring technologies with signal processing algorithms, or by applying signal separation algorithms (such as Independent Component Analysis, ICA) to distinguish between wheel damage signals and rail damage signals, the accuracy of rail corrugation detection can be improved.

Data availability

Data will be made available on request.

CRediT authorship contribution statement

Xueyang Tang: Conceptualization, Methodology, Supervision, Writing-review & editing. Xiaopei Cai: Conceptualization, Project administration, Supervision, Data curation, Funding acquisition. Yuqi Wang: Software, Writing- review & editing, Data

curation. Yue Hou: Supervision, Data curation.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The work was supported by the National Key R&D Program of China (No. 2022YFB2602901), the Project of Science and Technology Research and Development Program of China State Railway Group Co.. Ltd. (No. SY2022T002) and the National Natural Science Foundation of China (No. 52178405).

References

[1] H. Yang, J. Liu, G. Mei, Research on real-time detection method of rail corrugation based on improved ShuffleNet V2, Eng. Appl. Artif. Intell. 126 (2023) 106825, https://doi.org/10.1016/j.engappai.2023.106825.

[2] L. Zhuang, H. Qi, Z. Zhang, The automatic rail surface multi-flaw identification based on a deep learning powered framework, IEEE Trans. Intell. Transp. Syst. 23 (2022) 12133-12143, https://doi.org/10.1109/TITS.2021.3109949.

[3] X. Tang, X. Cai, H. Peng, et al., Experimental and simulation investigation into the cause and treatment of rail corrugation for metro, J. Cent. South Univ. 29 (2023) 3925-3938, https://doi.org/10.1007/s11771-022-5179-2.

[4] X. Cai, X. Tang, S. Pan, et al., Intelligent recognition of defects in high-speed railway slab track with limited dataset, Comput.-Aided Civ. Infrastruct. Eng. (2023) 1-18, https://doi.org/10.1111/mice.13109.

[5] X. Cai, X. Tang, W. Chang, et al., Machine learning-based rail corrugation recognition: a metro vehicle response and noise perspective, Phil. Trans. R. Soc. A 381 (2023) 20220171, https://doi.org/10.1098/rsta.2022.0171.

[6] S. Wang, Y. Xu, Y. Zhou, et al., Detection and evaluation of curve corrugation of urban mass transit, Urban Mass Transit 10 (2011) 56-60, https://doi.org/10.16037/j.1007-869x.2011.10.021.

[7] S. Grassie, Rail corrugation: advances in measurement, understanding and treatment, Wear 258 (2005) 1224-1234, https://doi.org/10.1016/j.wear.2004.03.066.

[8] L. Luo, G. Zhang, W. Wu, et al., Control of track smoothness in wheel-track systems, Beijing: China Railway Press, (2006) 115-135.

[9] L. Chen, Research on key issue of rail corrugation dynamic measurement based on chord measurement method, Hunan Univ. 2019, https://doi.org/10.227135/d.cnki.ghudu.2019.002497.

[10] H. Wei, H. Liu, Z. Ma, et al., A wide-area measurement method of rail corrugation based on the combination-chord system, J. Northwest Univ. (Nat. Sci. Ed.) 48 (2018) 199-208, https://doi.org/10.16152/j.cnki.xdxbzr.2018-02-007.

[11] S. Grassie, Measurement of railhead longitudinal profiles: a comparison of different techniques, Wear 191 (1996) 245-251, https://doi.org/10.1016/0043-1648(95)06732-9.

[12] J. Xu, P. Wang, L. Wang, et al., Research on the distribution characteristics and influence factors of sensitive wavelength of track vertical profile irregularity, J. China Railway Soc. 37 (2015) 72-78, https://doi.org/10.3969/j.issn.1001-8361.2015.07.012.

[13] F. Zhou, G. Zhang, K. Zhu, Laser vision dynamic measuring device and measuring method for rail, CN, CN1776364A (2006).

[14] W. Wang, Q. Liu, H. Wang, Laser measuring equipment for rail corrugation, CN, CN202522210U (2012).

[15] Q. Li, H. Zhang, S. Ren, Detection method for rail corrugation based on rail image feature in frequency domain, China Railway Sci. 37 (2016) 24-30, https://doi.org/10.3969/j.issn.1001-4632.2016.01.04.

[16] Z. Ma, Y. Dong, H. Liu, et al., Rail corrugation dynamic measurement method based on multi-line structured-light vision, J. Sci. Instrum. 39 (2018) 189-197, https://doi.org/10.19650/j.cnki.cjsi.J1803154.

[17] Y. Sunaga, T. Ide, M. Kanao, A Practical use of axle-box acceleration to control the short wave track irregularities on Shinkansen, Railway Tech. Res. Inst. 9 (1995) 35-40.

[18] M. Yoshida, S. Shinowaki, Y. Sunaga, A study on lateral axle-box acceleration for detecting irregularities of rail joints on sharp curves, Jointed Railway Technol. Symp. 11 (2004) 91-54.

[19] R. Potter, G. Kopp, H. Green, Visible speech, New York, NY, USA: Van Nostrand, 1947.

[20] B. Hou, Z. Xu, P. Gong, Prediction of damage to in-service steel rails, J. China Railway Soc. 20 (1998) 127-131, https://doi.org/10.3321/j.issn:1001-8360.1998.03.023.

[21] M. Roppongi, Y. Shibuya, K. Chiyoda, A new detecting method for rail corrugation by using wavelet analysis, WCRR (1999) 919.

[22] F. Coudert, Y. Sunaga, K. Takegami, Use of axle box acceleration to detect track and rail irregularities, WCRR 7 (1999) 1-7.

[23] B. Hopkins, S. Taheri, Broken rail prediction and detection using wavelets and artificial neural networks, Proc. ASME/ASCE/IEEE 2011 Joint Rail Conf., Colorado, USA, March 16-18, 2011, pp. 77-84.

[24] B. Hopkins, S. Taheri, Track health monitoring using wavelets, Proc. ASME 2010 Rail Transp. Div. Fall Tech. Conf., Virginia, USA, October 12-13, 2010, pp. 9-15.

[25] R. Gomes, A. Batista, M. Ortigueira, et al., Railscan: a tool for the detection and quantification of rail corrugation, Theor. Exp. Chem. 314 (2010) 153-157, https://doi.org/10.1007/978-3-642-11628-5_44.

[26] X. Wei, F. Liu, L. Jia, Urban rail track condition monitoring on in-service vehicle acceleration measurements, Measurement 80 (2016) 217-228, https://doi.org/10.1016/j.measurement.2015.11.033.

[27] T. Kojima, H. Tsunashima, A. Matsumoto, Fault detection of railway track by multi-resolution analysis, WIT Trans. Built Environ. 88 (2006) 955-964, https://doi.org/10.2495/CR060931.

[28] Q. Xie, G. Tao, M. Liu, et al., Application of mathematical morphology filter in recognition of rail corrugation wavelength, J. Cent. South Univ. (Sci. Technol.) 52 (2021) 1724-1732, https://doi.org/10.11817/j.issn.1672-7207.2021.05.033.

[29] C. Zhou, J. Gao, Dynamic effect and identification of rail corrugation irregularity based on the three-dimensional wheel-rail transient dynamic model, J. China Railway Soc. 17 (2020) 841-848, https://doi.org/CNKI:SUN:CSTD.0.2020-04-007.

[30] J. Li, Analysis of wheel-rail vibration signals for rail transit based on frequency-sliced wavelet transforms, Nanjing Univ. Sci. Technol. 2014.

[31] H. Li, J. Chen, Y. Ren, The research of rail wavy abrasion detection system, Southwest Jiaotong Univ. 2014, https://doi.org/10.3969/j.issn.1672-6944.2014.05.065.

[32] M. Tankaka, Q. Cai, The on-vehicle supervision method for rail corrugation, Foreign Rolling Stock 54 (2017) 31-36, https://doi.org/10.3969/j.issn.1002-7610.2017.01.005.

[33] L. Wang, The fast testing method of high speed railway rail corrugation on the basis of vibration response, Chin. Railway 7 (2017) 44-49, https://doi.org/10.19549/j.issn.1001-683x.2017.07.044.

[34] H. Jiang, C. Shang, R. Gao, Wheel/rail fault noise diagnosis method based on EMD and neural network, J. Vib. Shock 33 (2014) 34-38, https://doi.org/CNKI:SUN:ZDCJ.0.2014-17-008.

[35] Z. Zhou, N. Hu, Y. Huang, et al., Research on orbital wave wear detection method based on support vector machine, Proc. 13th Nat. Conf. Vib. Theory Appl. (2019) 193-197.

[36] L. Zhao, Research on detection and recognition of rail corrugation based on train vibration information, Beijing Jiaotong Univ. 2021, https://doi.org/10.26944/d.cnki.gbfju.2021.000840.

[37] Z. Zhang, Rail corrugation based on time-frequency analysis and data mining, North China Electr. Power Univ. (Beijing) 2021, https://doi.org/10.27140/d.cnki.ghbbu.2021.001546.

[38] B. Xiao, J. Liu, L. Niu, et al., Diagnosis method of rail corrugation for heavy haul railway based on WPD-ASTFT and SVM, J. China Railway Soc. 59 (2021) 31-35+48.

[39] Q. Xie, G. Tao, Z. Wen, Detection method of metro rail corrugation based on 1-dimensional convolutional neural network, J.Cent. South Univ. 52 (2021) 1371-1379, https://doi.org/10.11817/j.issn.1672-7207.2021.04.033.

[40] Q. Xie, G. Tao, Bin He, et al., Rail corrugation detection using one-dimensional convolution neural network and data-driven method, Measurement. 200(2022)111624. https://doi.org/10.1016/j.measurement.2022.111624.

[41] Z. Chen, Q. Wang, K.Yang, Deep learning for the detection and recognition of rail defects in ultrasound B-scan images, Transp. Res. Rec. 23(2022)12133-12143. 2675(2021) 888-901. https://doi.org/10.1177/03611981211021547

[42] X. Wu,Y. Li, Z. Ma, et al. Capacity configuration method of hybrid energy storage system based on improved wavelet packet decomposition, Acta Energiae Solaris Sin. 44(2023):23-29, https://doi.org/10.19912/j.0254-0096.tynxb.2022-0578.

[43] N .Cong,Y. Ren, J. Chen, et al., A simulation method for vehicle frequency domain non-stationary vibration spectrum based on wavelet

packet reconstruction and fuzzy C-mean clustering, J. Vib. Shock. 37(2018):92-99, https://doi.org/10.13465/j.cnki.jvs.2018.21.014.

[44] W. Huang,G. Yin, K. Geng, et al., Object detection in complex driving scene based ondilated convolution feature adaptive fusion, J. Southeast Univ. 51(2021):1076-1083. https://doi.org/10.3969/i.issn.1001 -0505.2021.06.021.

[45] G. Jiang, J. Zhou, X. Wu, et al., Wind turbine anomaly detection based on dilated causai convolution network, Acta Energiae Solaris Sin. 44(2023):368-375. https://doi.org/10.19912/j.0254-0096.tynxb.2021-1581.

[46] X. Liu, Z. Yang, H. Wan, et al., An improved maximum consistency geometricprimitives fitting algorithm for point cloud, J. Tongji Univ.,43(2015):1246-1253. https://doi.org/10.11908/j.issn.0253-374x.2015.08.019

[47] S. Yue, N. Li, W. Wang, et al., Linear interpolation based adaptive feature-preserving filtering method onmesh models, J. Computer-Aided Des. Comput. Graphics. 32(2020):1377-1388. https://doi.org/ 10.3724/SP.J.1089.2020.18104

[48] K. Ferkous, M. Guermoui, S. Menakh, et al., A novel learning approach for short-term photovoltaic power forecasting - A review and case studies, Eng. Appl. Artif. Intell. 133 (2024) 108502, https://doi.org/10.1016/j.engappai.2024.108502.

[49] Y. Lyu, Y. Xu, C. Yang, et al., WSRGAN: A wavelet-based GAN for super-resolution of plane-wave ultrasound images without sampling loss, Eng. Appl. Artif. Intell. 127 (2024) 107834, https://doi.org/10.1016/j.engappai.2023.107384.

[50] W. Yang, K. Xia, Z. Wang, et al., Self-Attention Causal Dilated Convolutional Neural Network for Multivariate Time Series Classification and Its Application, Eng. Appl. Artif. Intell. 122 (2023) 106151, https://doi.org/10.1016/j.engappai.2023.106151.

[51] T. M. Khan, S. S. Naqvi, E. Meijering, ESDMR-Net: A lightweight network with expand-squeeze and dual multiscale residual connections for medical image segmentation, Eng. Appl. Artif. Intell. 133 (2024) 107995, https://doi.org/10.1016/j.engappai.2024.107995.

[52] P. Li, T. Si, C. Ye, et al., Semantic-aware transformer with feature integration for remote sensing change detection, Eng. Appl. Artif. Intell. 135(2024) 108774, https://doi.org/10.1016/j.engappai.2024.108774.

[53] E. Koşar, B. Barshan, A new CNN-LSTM architecture for activity recognition employing wearable motion sensor data: Enabling diverse feature extraction, Eng. Appl. Artif. Intell. 124 (2023) 106529, https://doi.org/10.1016/j.engappai.2023.106529.

[54] H. Guo, Y. Li, Y. Li, et al., A hybrid attention-guided ConvNeXt-GRU network for action recognition, Eng. Appl. Artif. Intell. 49 (2024) 108243, https://doi.org/10.1016/j.engappai.2024.108243.

[55] N. S. Chauhan, N. Kumar, Confined attention mechanism enabled Recurrent Neural Network framework to improve traffic flow prediction, Eng. Appl. Artif. Intell. 136(2024) 108791, https://doi.org/10.1016/j.engappai.2024.108791.

[56] A. Mahmoud , T. Hu, X. Zeng, et al., Hydraulic informed multi-layer perceptron for estimating discharge coefficient of labyrinth weir, Eng. Appl. Artif. Intell. 123(2023) 106435, https://doi.org/10.1016/j.engappai.2023.106435.

[57] P. Sahu, B. K. Singh, N. Nirala, Optimized k-nearest neighbors for classification of prosthetic hand movements using electromyography signal, Eng. Appl. Artif. Intell. 133 (2024) 108390, https://doi.org/10.1016/j.engappai.2024.108390.

[58] L. Li, J. Xie, X. Zhao, et al., Screening out potentially defective products in micro-transformer production by intelligently integrating mechanical and electronic signals, Eng. Appl. Artif. Intell. 126 (2023) 107186, https://doi.org/10.1016/j.engappai.2023.107186.