International Journal for Housing Science and Its Applications
Publish August 6, 2025. Volume 46, Issue 3 Pages 6235-6247

https://doi.org/10.70517/ijhsa463535

Research on the application of image-based feature extraction algorithm in the detection of railroad track defects

Yuefei Liu[1,*]
[1] School of Geodesy and Geomatics, Wuhan University, Wuhan, Hubei, 430079, China
Corresponding authors: (e-mail: yuefeiliu@whu.edu.cn).

Abstract Railroad track safety is critical to transportation operations, and track defects may lead to serious accidents. This paper proposes a railroad track defect detection method based on image feature extraction algorithm and BP neural network. Firstly, X-ray imaging is used to obtain the rail defect image, and then the image quality is enhanced by wavelet transform noise reduction processing and contrast and entropy. Then, the improved Otsu threshold method is used to segment the image and extract the geometric features of the defective image; finally, a BP neural network with 5-15-6 structure is constructed to classify and recognize the defects. The experimental results show that the running time of the proposed method in the threshold segmentation stage is only 0.1526 seconds, which is reduced by 62.3% compared with the maximum Shannon entropy multi-threshold segmentation method. For the six defects of wave abrasion, optical band anomaly, serration, abrasion, corrosion and pitting, five geometric features of frequency, amplitude, area, long axis and short axis are selected for the BP neural network training, and the network reaches an error of 0.01 after 112 trainings, which realizes the accurate classification of defects. The method improves the automation degree and accuracy of track defect detection through the combination of image feature extraction and machine learning, which provides an important guarantee for the safe operation of railroads.

Index Terms Railroad track defects, Image feature extraction, Otsu threshold segmentation, Geometric features, BP neural network, Defect classification

I. Introduction

Railroad is an important mode of transportation in China, occupying a backbone position in the transportation system [1]. However, due to the underdeveloped technology of China's railroad system and the backward structure of the network, it is difficult to adapt to the growing demand for transportation, and the speed upgrading of railroads as well as the continuous construction and development of high-speed railroads have become the inevitable trend of social development [2], [3]. China's railroad system has had six large speed upgrading project, the introduction and development of high-speed railroad speed of 350 km/h or more, the railroad track safety inspection technology has become the cornerstone of the maintenance of high-speed railroad operation safety [4]. High-speed trains have extremely strict requirements for railroad tracks, and track injuries are a greater threat to the safety of trains [5]. Heavy-duty railroad transportation is highly valued because of its large transport volume and high economic efficiency, and China's resource pattern determines that heavy-duty railroad is an effective way to solve the uneven geographical distribution of resources, while the damage to the track by heavy-duty railroad is very serious [6]-[8]. With the development of metal smelting technology and nondestructive flaw detection technology, the probability of the original injury inside the track is getting smaller and smaller, but the risk of the internal injury of the track triggered by the surface defects of the railroad track is still larger [9]. The main types of track surface defects are spalling, scratches, cracks, scars, etc. These surface defects can cause the formation of internal injuries on the railroad track, which in turn leads to fracture and derailment of the track [10]-[12]. Timely detection of track surface defects is the best measure to prevent the development of railroad track injuries and ensure the safety of railroad traffic, so the detection of track surface defects is an important aspect of the daily inspection and maintenance of the railroad system [13]-[15].

Machine vision technology came into being along with the development of laser technology and CCD technology, and has been widely promoted in the inspection field [16]. For example, Min, Y et al. proposed a method for real-time detection of railroad track surface defects using machine vision, aiming to improve the efficiency and adaptability of railroad maintenance by replacing manual inspection with an automated system [17]. Shi, Y et al. proposed a method for detecting welding defects in railroads based on machine vision technology, which has high practical value with an accuracy of more than 90% compared with traditional methods [18]. Wu, Y and Lu, Y

combined machine vision with the support vector machine (SVM) algorithm in machine learning to develop a method for detecting surface defects on railroad tracks, which achieved more than 98% detection accuracy [19]. Machine vision technology is based on the input image, which is analyzed by an algorithm to obtain information about the depth in the image, such as positional information, texture information, etc [20]. The machine vision system can be divided into three components, image acquisition, image feature extraction, and output of data [21]. Kumar, A and Harsha, S. P provided a comprehensive overview of the research related to image processing technology based defect detection methods for railroads in recent years, stating that the combination of computer vision algorithms, machine learning models, and deep learning techniques can increase the detection accuracy and improve the safety and reliability of railroads [22]. Dubey, A. K and Jaffery, Z. A. proposed a visual inspection technique based on maximum stable extreme region marking to detect railroad track defects focusing on three surface defects with good results [23].

Images of railroad tracks are captured in an online manner during the acquisition process, so they need to be captured by high-quality camera lenses [24]. Aldao, E et al. compared different methods used for autonomous inspection of railroad bolts and fixtures by means of a prototype system equipped with LIDAR and camera sensors, evaluating the performance of these methods applied to the detection of railroad track defects [25]. The processing of the acquired rail images is the core part of the machine vision system, the content of the acquired images is influenced by the background and the measured area in the image is not a single image [26]. In this regard, Zhuang, L et al. developed a deep learning-driven two-phase approach designed to reduce the influence of image content by background when used for automated detection of major components of railroad tracks, achieving high accuracy of 95.2% mAP [27]. In the processing of track surface defects, the shape and bit depth of the defects cannot be determined in advance, leading to ambiguity in the detection of track defects [28]. To solve the above blurring problem, Cheng, M. Y et al. proposed a novel Hybrid Generative Adversarial Network (HybridGAN) to enhance the quality of dynamically acquired railroad inspection images to improve resolution and reduce blurring, and demonstrated its effectiveness in improving automated railroad safety maintenance [29]. The above study lays the foundation for our research on railroad track defect detection, however, due to the fact that the image acquisition equipment is installed on high-speed running locomotives, the track is prone to reflect the light, and the noise of locomotives and vehicles also affects the image, which results in a reduction of the accuracy of the defect detection, and further research is still needed.

As an important national transportation infrastructure, the safe and stable operation of railroad is directly related to the national economic development and the safety of people's lives and properties. Railroad track as the carrier of train operation, its integrity has a decisive impact on the safe operation of the train. In the long-term operation of the railroad, due to the vehicle load, the natural environment and the aging of materials and other factors, the surface of the track is prone to wave abrasion, light band abnormalities, jagged, abrasion, corrosion and pitting and other types of defects. These defects, if not found and repaired in time, will lead to the deterioration of track geometry and even cause derailment and other major accidents. Therefore, the development of efficient and accurate track defect detection technology is of great significance to the safe operation of railroads. Traditional track defect detection mainly relies on manual visual inspection, low efficiency and low precision; ultrasonic detection can find internal defects, but limited ability to identify surface defects; electromagnetic induction detection is subject to external interference and lack of stability. In recent years, the application of computer vision technology in the industrial field has become more and more extensive, and the defect detection method based on image processing has become a research hotspot for track defect detection due to its advantages of non-contact, high accuracy and high efficiency. However, the existing image processing methods still face the problems of large noise interference, insufficient feature extraction, and low classification accuracy when dealing with track defect images in complex backgrounds. This study proposes a railroad track defect detection method based on the combination of image feature extraction algorithm and BP neural network. High-quality images of rail defects are obtained using X-ray imaging technology, and the image noise reduction process is carried out by wavelet transform and image enhancement is realized by using local gain function based on contrast and entropy; then, the Otsu threshold segmentation algorithm is improved to realize the accurate segmentation of defective regions; then, the geometric features of defects are extracted, including key parameters such as frequency, amplitude, area, long and short axes, etc. Finally, a BP neural network with 5-15-6 structure is constructed to detect rail defects. 5-15-6 structure of BP neural network model to classify and recognize six typical track defects. The effectiveness of the proposed method is verified by comparing and analyzing the performance of different threshold segmentation methods; the accuracy of the neural network classification is evaluated by a large number of samples training and testing.

II. Railway track defect detection method research

II. A. Acquisition of defect images at the defect

In order to obtain more accurate images, this paper adopts X-ray imaging technology in the acquisition of railroad track defect images, i.e., the use of digital detectors to obtain railroad track rays, and then output digital images. In this process, the screen converts the X-rays into beams visible to the naked eye, after which the beams are converted into video signals, thus realizing the effect of real-time generation and output of railroad track defect images [30]. When the railroad track is penetrated by X-rays, because there are differences in the absorption effect of X-rays between the defective and non-defective places, clear and accurate images of railroad track defects can be captured by the above method. The expression using X-ray imaging is shown in equation (1):

$$Q_k = Q_s e^{-\delta h} \tag{1}$$

where Q_k denotes the intensity of X-rays that penetrate the railroad track defect. h denotes the length of the rail defect. Q_k denotes the initial intensity of the X-rays. δ denotes the value of the attenuation coefficient of the X-rays. When the intensity of X-ray projection at the track defect is greater than that in the region near the defect, the X-ray film obtained shows white X-rays and black defects. Conversely, the X-rays are black and the defects are white. Because X-ray imaging is a source of electromagnetic radiation, the X-ray film needs to be digitized again to obtain a digital image of the track defect.

II. B. Preprocessing defect images at defects

(1) Image noise reduction

In the acquisition of track defect images, due to the influence of the external environment and other factors, there is a large amount of noise in the image, and the details of the defects are not obvious enough to seriously affect the accuracy of the subsequent defect detection. Therefore, before further defect detection, it is necessary to perform noise reduction on the collected track defect images. In this paper, the wavelet transform will be utilized to carry out noise reduction processing of defect images. Image noise reduction by wavelet transform is to use the mapping relationship of wavelet signal space to obtain the original noise-free signal [31]. Because the quantum noise data in the defect image formed by X-ray imaging has different frequency domain components in space, the wavelet transform coefficients are also somewhat different in the case of different defect sizes in the image. In this paper, wavelet noise reduction based on this principle can realize the separation of effective data and noise data of defect images under the premise of ensuring the integrity of image data. According to the definition of wavelet transform, the wavelet mother function is processed in a certain way to obtain the wavelet sequence expression as shown in equation (2):

$$M_{s,p}(t) = \frac{1}{\sqrt{|s|}} M_{s,p}\left(\frac{t-p}{s}\right) \tag{2}$$

In Eq. (2), $M_{s,p}(t)$ denotes the wavelet mother function that depends on s and p. t denotes the time series of defective images. s denotes the scaling factor data used to process the wavelet mother function. p denotes the translation factor data used to process the wavelet mother function. Because s and p are two changing data in wavelet transform, the wavelet mother function can be processed by these two data, and then the wavelet sequence can be obtained. Through this formula to obtain the track defect image signal in the space of the local mode maximum value point, according to the signal of different frequency domain components, can be separated from the noise signal, in order to achieve the purpose of image noise reduction.

(2) Image enhancement

After the noise reduction process, the track defect image has the disadvantage of low contrast, which generally shows the phenomenon that the overall defect image is brighter or darker. Therefore, this paper needs to improve the image quality and restore the image contrast through image enhancement processing. Because in the detection of rail defects, the defects in the image have a diversity of characteristics, resulting in each type of defects and the contrast of the background are not the same, so if you use the general algorithm to enhance the image, the computational process is complex and the adaptive effect is poor, which is not conducive to the subsequent detection of defects. Based on the actual situation of track defects in a section of the line, this paper selects contrast and entropy as the parameters of image optimization, and processes the grayscale characteristics of local areas in the image. The expressions of contrast and entropy are shown in equation (3) and equation (4):

$$D = \frac{1}{2}\left(\frac{h_1 - \bar{h}}{h_1 + \bar{h}} + \frac{h_2 - \bar{h}}{h_2 + \bar{h}}\right) \tag{3}$$

$$Z = -\sum_{n=h_{\min}}^{h_{\max}} g_n \lg g_n \tag{4}$$

where D denotes the contrast parameter of the defective image. Z denotes the entropy parameter of the defective image. h_1, h_2 denote the grayscale data of pixels in two different regions of the image and the range of values of the grayscale of pixels in these two regions is $[h_{\min}, h_{\max}]$. The $\bar{h}$ denotes the average gray scale value. g_n denotes the probabilistic data for region pixel gray value of n. Then, the formula for obtaining the local gain of the defective image using contrast and entropy is shown in equation (5):

$$Y(a,b) = \begin{cases} (\mu D + (1-\mu)Z) \times \dfrac{\bar{h}}{(h_{a,b} - \bar{h})^2}, \mu \geq \dfrac{1}{2} \\ ((1-\mu)D + \mu Z) \times \dfrac{\bar{h}}{(h_{a,b} - \bar{h})^2}, \mu < \dfrac{1}{2} \end{cases} \tag{5}$$

In Eq. (5), $Y(a,b)$ denotes the local gain function of the region with pixel point coordinates (a,b) in the defect image. μ denotes the proportionality constant. $h_{a,b}$ denotes the grayscale data of the region with pixel point coordinates of (a,b) in the defect image. Gaining the local pixels of the track defect image according to this formula can obtain the image enhancement effect, which in turn makes the features of the defect image more prominent.

II. C. Threshold Segmentation

Orbit defect segmentation can be divided into two categories: adaptive thresholding and global thresholding. Adaptive thresholding is to divide the image into multiple subgraphs and segment the defects with corresponding thresholds respectively. Global thresholding is to find a fixed threshold to segment defects according to different algorithms. In this paper, global thresholding is used to segment out the track defects, and the widely used global thresholding methods are: iterative thresholding, maximum entropy and maximum inter-class variance (Olsu).

Olsu thresholding method is better compared to iterative thresholding method and maximum entropy method [32], and Otsu threshold segmentation is calculated as:

$$\sigma^2(t) = \omega_1(t)(\mu_1(t) - \mu_T)^2 + \omega_2(t)(\mu_2(t) - \mu_T)^2 \tag{6}$$

where, t is the assumed set threshold. μ_1, μ_2 and μ_T are denoted as: the defective part gray mean, the background part gray mean and the global gray mean, respectively. $\omega_i(t)$ and $\omega_2(t)$ are the defective part and background part probabilities. The optimal threshold is calculated as:

$$t^* = \arg\max_{i \in [0,255]} (\sigma^2(t)) \tag{7}$$

Due to the large value of the defect proportion $\omega_1(t)$ and the small value of the background proportion $\omega_2(t)$, which makes the ideal segmentation threshold t high, it is necessary to increase the proportion of background weights, at which time the optimal threshold is calculated as Eq:

$$t^* = \arg\max_{i \in [0,255]} (\omega_2(t)^{\varepsilon} \sigma^2(t)) \tag{8}$$

where the weight parameter ε controls the extent to which $\omega_2(t)$ plays a role in the background share of the image.

II. D. Extraction of defective image features

After the pre-processing of the railroad track defect image, it is necessary to further extract the features of the image in order to subsequently realize defect detection based on the quantified image features. In this paper, the primary feature driver in the detection of rail defects is the geometric parameters of the image, the image geometric features mainly include the perimeter of the defects, the area of the defects and the defects of the degree of roundness. Before extracting the geometric features of the track defect image, it is necessary to define the defect size in the image, i.e., it is necessary to realize the conversion between the geometric feature parameters and the actual track defect size, which is realized by the wire transmissometer in this paper. Then, the feature extraction of the track defect image is carried out, and the geometric feature parameters of the defects in the image are calculated as shown in equation (9):

$$\begin{cases} J_s = N_s \times \dfrac{l}{N_l} \\ J_z = N_z \times \dfrac{l}{N_l} \\ J_y = \dfrac{4\pi \times J_s}{J_z^2} \end{cases} \tag{9}$$

In Eq. (9), J_s, J_z, J_y denote the defect area feature parameter, defect perimeter feature parameter, and defect roundness feature parameter in the geometrical features extracted from the railroad track defect image, respectively. l denotes the length between 2 wires in the wire transmissometer. N_l denotes the number of pixel points present between the 2 wires in the wire transmissometer. N_s denotes the number of pixel points contained in the defective region in the track defect image. N_z denotes the number of pixel points contained at the boundary of the defective region in the track defect image. The geometric features of track defect image obtained by this formula can provide a strong basis for track defect recognition and classification.

II. E. Experimental results and analysis

II. E. 1) Threshold Segmentation Effect Analysis

The proposed Otsu thresholding method is validated through a large number of image segmentation experiments, and the image segmentation results are given and compared with the maximum Shannon entropy multi-threshold segmentation method (Method 1), the two-dimensional Shannon entropy segmentation method (Method 2), the two-dimensional Tsallis gray-scale entropy segmentation method (Method 3), the inverse gray-scale entropy multi-threshold segmentation method (Method 4) were performed to compare the image segmentation performance.

Three defective images are selected for threshold segmentation, and the optimal thresholds for segmentation as well as the running time of the five methods are compared as shown in Table 1. Since Shannon entropy involves logarithmic operations, the maximum Shannon entropy multi-threshold segmentation method and the two-dimensional Shannon entropy segmentation method take longer time. The inverse gray scale entropy multi-threshold segmentation method takes slightly less time than the first two methods, but the algorithm still takes longer to run due to the use of adaptive two-particle swarm for optimal threshold search. The two-dimensional Tsallis gray entropy segmentation method uses recursion to calculate the intermediate variables, which reduces the complexity of the algorithm and the running time is relatively short. The experiments utilize the Olsu thresholding method with the shortest time consumption, which can more accurately divide the target and background.

Table 1: Optimal threshold and run time comparison

Segmentation method	Threshold value	Time/s	Threshold value	Time/s	Threshold value	Time/s
Method 1	(83,164)	0.4046	(103,145)	0.4045	(88,173)	0.3821
Method 2	(83,78)	0.3676	(73,75)	0.3667	(78,76)	0.3491
Method 3	(75,79)	0.1984	(67,67)	0.1928	(84,86)	0.1647
Method 4	(46,113)	0.3251	(42,116)	0.3245	(51,116)	0.3142
Otsu	(81,156)	0.1526	(67,143)	0.1516	(76,147)	0.1485

II. E. 2) Characterization of rail surface defects

(1) Representation and description of defect geometric features

Geometric features are used to describe the target appearance and morphology of the characteristics of the bright, is a common feature of the rail surface defects, the main choice of geometric features in this paper are frequency f, amplitude A, perimeter d, area S, the long axis of the L_{MER} and the short axis of the W_{MER}.

1) Frequency f

Waveform frequency is the number of times a waveform repeats itself per unit line, and is a quantity used to measure the frequency of periodic waveforms, calculated as follows:

$$f = \frac{1}{r} \tag{10}$$

where T is the period of waveform repetition.

2) Amplitude A

Amplitude is the gray value of the line and the maximum possible value, indicating the line of the maximum number of white dots.

3) Perimeter d

Perimeter can be used to measure the variation of rail defects in the direction along the rail and its perpendicular direction. There are two calculation methods, one is to use the Euclidean distance calculation, through the defect edge neighbor perimeter and get perimeter, the method has the advantages of the obtained perimeter and the real perimeter of the difference is smaller, but also so the calculation efficiency is low. The other is to use the 8-neighborhood distance and directly take the number of pixel points of the defect edge as the perimeter, which has high computational efficiency, but the difference between the obtained perimeter and the real perimeter is larger. After comprehensive consideration, this paper chooses to use

Euclidean distance to calculate the perimeter.

4) Area A

Area is a quantity used to describe the size of a defect and is one of the defect characteristics. Area is the number of pixel points in the defect, the size of the area is determined by the defect boundary, and is not related to the gray level of the pixel points within the defect.

5) The long axis L_{MER} and the short axis W_{MER}

When the defect boundary is determined, the defect geometry can be characterized by the minimum outer rectangle MER, and the length and width of the minimum outer rectangle are the long and short axes, respectively.

(2) Selection of defect features

The six classified defects of wave abrasion, light band abnormality, serration, abrasion, corrosion and pitting are compared using geometric features. The corresponding grayscale waveforms of the six defects are shown in Fig. 1, and (a)~(f) represent the grayscale waveforms of wave abrasion, optical band anomaly, serration, abrasion, corrosion, and pitting, respectively.

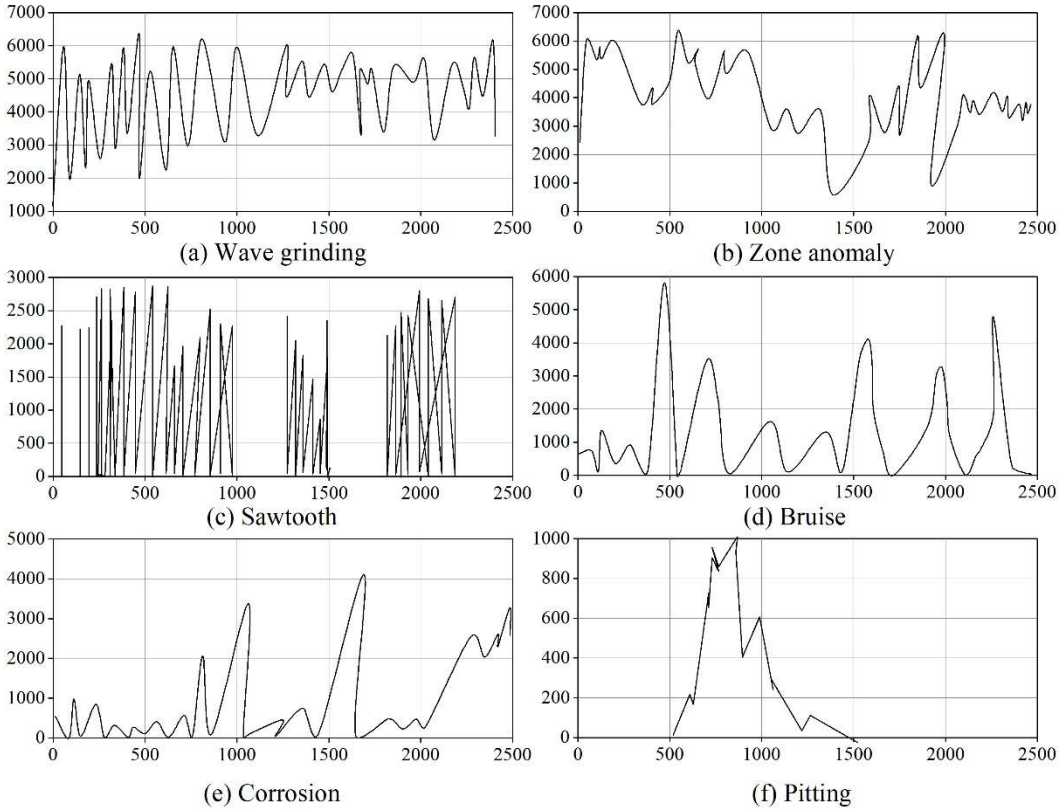

Figure 1: Six kinds of defects corresponding to the gray form

The results of geometric feature comparison are shown in Table 2. Wave abrasion defects have the largest area, perimeter, long-axis and short-axis among the six types of defects, pitting defects have the smallest area, amplitude and perimeter, sawtooth frequency is the highest, optical band anomaly defects are in the second place of the six features, abrasion defects have the third place in the long-axis, area and perimeter and abrasion waveforms are

relatively uniform, and corrosion of the short-axis is the shortest, and the area and perimeter are ranked in the fourth place. Since the area and perimeter are highly correlated, this paper selects five geometric features, namely, frequency, amplitude, area, long-axis and short-axis, to form a feature vector of dimension 5 to describe the six rail surface defects.

Table 2: Comparison of geometric characteristics

	Wave grinding	Zone anomaly	Sawtooth	Bruise	Corrosion	Pitting
Frequency	0.01864	0.00316	0.03024	0.00236	0.00414	0.00063
Amplitude	6133	6853	2504	5894	4868	986
Perimeter	5896	3059	570	800	126	52
Area	84624	48967	7256	10366	2565	302
Long axis	1896	957	16	118	30	13
Short axis	42	8	5	7	4	8

III. Research on the classification method of railroad track defects

III. A. Defect classification method based on BP neural network

Artificial neural network applied to pattern classification, the most commonly used model is BP neural network, is a kind of multilayer feedforward network containing hidden layers and trained according to the error backpropagation algorithm, which has the ability to recognize nonlinear patterns [33].

Because the morphology of the rail surface defects that need to be detected in this paper is more complex, in order to accurately classify the six types of defects, this paper chooses the BP neural network in the artificial neural network to recognize the defects.

III. A. 1) Determination of network structure

BP neural network is generally divided into 3 layers, i.e., input layer, hidden layer and output layer. The input layer of the BP neural network applied to this detection system is 5 neurons, which corresponds to the defect's frequency f, amplitude A, perimeter d, area S, long axis L_{MER} and short axis W_{MER}. These are the 5 characteristic parameters. Based on experience, the hidden layer was determined to be 15 neurons. The output layer is 6 neurons, corresponding to the six defect categories of wave abrasion, light band abnormality, serration, abrasion, corrosion and pitting. The structure of the BP neural network applied to this inspection system is shown in Figure 2.

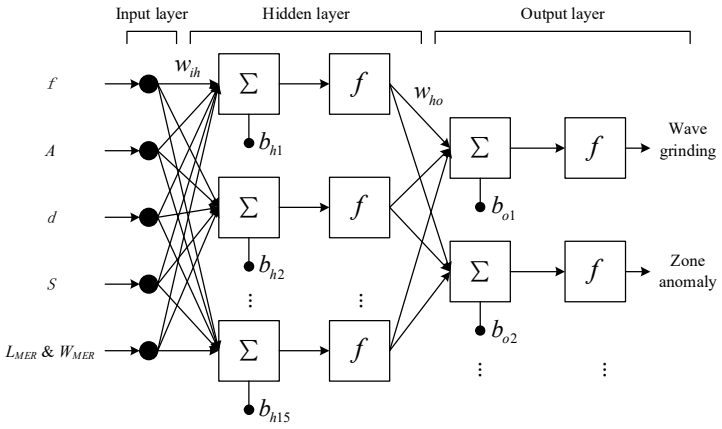

Figure 2: Neural network structure of track defect recognition

where the input vector is $x=(x_1,x_2,x_3,x_4,x_5)$ =(frequency, amplitude, perimeter, area, long and short axes), and the implicit layer input vector is $hi=(hi_1,hi_2,...,hi_{15})$, and the hidden layer output vector is $ho=(ho_1,ho_2,...,ho_{15})$, the input vector of the output layer is $yi=(yi_1,yi_2)$, the output vector of the output layer is $yo=(yo_1,yo_2)$ =(wave abrasion, optical band anomaly, jaggedness, abrasion, corrosion, pitting), the threshold of each neuron of the hidden layer is b_h, and the threshold of each neuron in the output layer is b_o. The weights between the input layer and the hidden layer are w_{ih}, and the weights between the hidden layer and the output layer are w_{ho}, where i takes the

value of (1, 2, ···, 5), h takes the value of (1, 2, ···, 15), and o takes the value of (1, 2, ···, 5). The transfer function is $f()$, where the transfer functions between the input layer and the implicit layer, and between the implicit layer and the output layer are the tansig function and the logsig function, respectively, which are derivable everywhere and restrict the output to between [-1, +1].

III. A. 2) BP Neural Network Training

The BP neural network is trained on the weights w_{ih} between the input layer and the hidden layer, the weights w_{ho} between the hidden layer and the output layer, the thresholds b_h of each neuron in the hidden layer, and the thresholds b_o of each neuron in the output layer, of which there are 120 in w_{ih}, 75 in w_{io} has 75, b_h has 15, and b_o has 5, so the total number of trained parameters is 215. The network training objective is to minimize the error function, when the weights w_{ih}, w_{ho}, b_h and b_o are optimal. The BP neural network is trained using gradient descent method with the following steps:

In the first step, the network is initialized. Random numbers are taken in the interval [-1, 1] and assigned as initial values to the connection weights w_{nh} and w_{ho}, respectively. Based on the experience, the initial values of threshold b_h for each neuron in the hidden layer and threshold b_o for each neuron in the output layer are given. Setting the error function e, the error after the k th training is:

$$e = \frac{1}{2}\sum_{o=1}^{2}(d_o(k) - yo_o(k))^2 \tag{11}$$

Set the output training accuracy ε of the BP neural network as 10-10 and the maximum number of learning times M as 3500. Based on the experience, this paper chooses the parameter tuning method as gradient descent method.

In the second step, 20 pairs of normalized learning pattern pairs for each of the 6 types of defects are provided to the network, i.e., a total of 120 sets of defective samples with feature vectors $x(l) = (x_1(l), x_2(l), x_3(l), x_4(l), x_5(l))(l = 1, 2, \ldots, 120)$ and its corresponding expected category vector $d_o(l) = (d_1(l), d_2(l))(1 = 1, 2, ..., 120)$, the category vector d_o is $[10]$ if the defect is a scar and d_o is $[01]$ if the defect is a corrugated abrasion. Input the learned pattern pairs into the BP neural network. Let this be the Nth training, and randomly select the k th learning pattern pair to input into the neural network, including the input samples $x(k) = (x_1(k), x_2(k), ..., x_5(k))$ and its corresponding expected output $d_o(k) = (d_1(k), d_2(k))$.

In the third step, the input and output of each neuron in the hidden and output layers are computed as follows:

$$hi_h(k) = \sum_{i=1}^{5} w_{ih}x_i(k) - b_h \quad h = 1, 2, ..., 10 \tag{12}$$

$$ho_h(k) = f_1(hi_h(k)) \quad h = 1, 2, ..., 10 \tag{13}$$

$$yi_o(k) = \sum_{h=1}^{10} w_{ho}ho_h(k) - b_o \quad o = 1, 2 \tag{14}$$

$$yo_o(k) = f_2(yi_o(k)) \quad o = 1, 2 \tag{15}$$

where $f_1()$ and $f_2()$ are the transfer functions from the input layer to the hidden layer and from the hidden layer to the output layer, respectively.

In the fourth step, the actual category vectors of the defect type outputs and samples identified by the neural network are utilized and Eq. (16) is applied to compute the partial derivatives $\delta_o(k)$, i.e., the partial derivatives of the error function e with respect to neuron yi_o of the output layer:

$$\begin{aligned}\delta_o(k) = \frac{\partial e}{\partial yi_o} &= \frac{\partial(\frac{1}{2}\sum_{o=1}^{2}(d_o(k) - yo_o(k)))^2}{\partial yi_o} = -(d_o(k) - yo_o(k))yo_o'(k) \\ &= -(d_o(k) - yo_o(k)f'(yi_o(k))\end{aligned} \tag{16}$$

In the fifth step, the partial derivatives $\delta_h(k)$ of the error function e with respect to the hidden layer neurons hi_h are computed:

$$\begin{aligned}\delta_h(k) &= \frac{\partial e}{\partial hi_h(k)} = \frac{\partial(\frac{1}{2}\sum_{o=1}^{2}(d_o(k)-yo_o(k))^2)}{\partial ho_h(k)}\frac{\partial ho_h(k)}{\partial hi_h(k)} \\ &= -\sum_{o=1}^{2}(d_o(k)-yo_o(k))f'(yi_o(k))w_{ho}\frac{\partial ho_h(k)}{\partial hi_h(k)} \\ &= -(\sum_{o=1}^{q}\delta_o(k)w_{ho})f'(hi_h(k))\end{aligned} \tag{17}$$

In the sixth step, the partial derivatives $\delta_o(k)$ and the outputs of the neurons in the hidden layer ho_h are used to correct the weights of the connections between the hidden layer and the output layer $w_{ho}(k)$ and the thresholds of each neuron in the output layer b_o:

$$\Delta w_{ho}(k) = \alpha\delta_o(k)ho_h(k) \tag{18}$$

$$w_{ho}^{N+1} = w_{ho}^{N} + \alpha\delta_o(k)ho_h(k) \tag{19}$$

$$\Delta b_o(k) = \alpha\delta_o(k) \tag{20}$$

$$b_o^{N+1} = b_o^{N} + \alpha\delta_o(k) \tag{21}$$

In the seventh step, the partial derivatives $\delta_h(k)$ and the input layer neuron inputs x_i are used to correct the weights of the connection between the input layer and the hidden layer $w_{ih}(k)$ and the thresholds of each neuron in the hidden layer b_h:

$$\Delta w_{ih}(k) = \beta\delta_h(k)x_i(k) \tag{22}$$

$$w_{ih}^{N+1} = w_{ih}^{N} + \beta\delta_h(k)x_i(k) \tag{23}$$

$$\Delta b_h(k) = \beta\delta_h(k)x_i(k) \tag{24}$$

$$b_h^{N+1} = b_h^{N} + \beta\delta_h(k) \tag{25}$$

The α in the formula in the sixth step and β in the formula in the seventh step are both learning coefficients.

In the eighth step, determine whether all 120 pairs of pattern pairs have been trained in this round of training session, and if all of them have been trained, add 1 to the number of learning times and go to the ninth step. If there are still unused learning pattern pairs, update the learning pattern pairs, i.e., select the next learning pattern pair that has not been used yet, and go to step two.

In the ninth step, calculate the global error:

$$E = \frac{1}{2\times 40}\sum_{k=1}^{40}\sum_{o=1}^{2}(d_o(k)-y_o(k))^2 \tag{26}$$

In the tenth step, determine whether the global error or the number of learning times meets the neural network requirements, and end the algorithm if the global error is less than or equal to the pre-set accuracy or the number of network training reaches the pre-set maximum value. Otherwise, reuse the previous 120 pairs of learning pattern pairs, go to step 2, and enter the next round of learning.The flowchart of the training algorithm of BP neural network is shown in Fig. 3.

III. B. Network training results

Selected in the BP neural network input training samples, 5 types of features, 6 types of defects, the number of samples 120, to 1 ~ 6, respectively, that the defective types of wave abrasion, light band anomalies, serration, abrasion, corrosion, pitting, the training error is set to 0.01, and some of the training samples feature values of the information shown in Table 3. By inputting training samples to the BP neural network for network training, the network has reached the training purpose at this time.

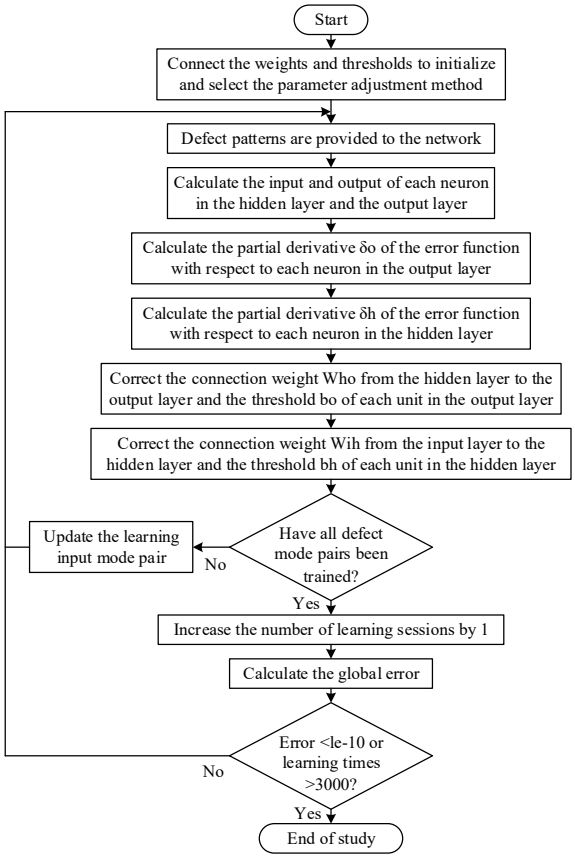

Figure 3: Flowchart of BP Neural Network Training

Table 3: Training sample

Serial number	f	A	d	S	L_{MER}	W_{MER}	Expected output
1	0.0174	1073	1751	59418	610	9	6
2	0.0181	2701	2769	47145	1431	8	5
3	0.0152	5527	5640	46508	468	14	4
4	0.0187	2495	4244	38964	345	25	3
5	0.0133	1347	2269	48419	176	22	3
6	0.0147	2835	2170	12477	1235	12	5
7	0.0132	969	2998	76779	149	41	1
8	0.0098	4939	4665	80572	1711	25	3
9	0.0106	1354	588	60516	884	22	3
10	0.0125	4892	3045	23291	839	4	5
11	0.0136	3435	3571	76474	1841	18	5
12	0.0094	3902	1773	75354	1558	29	2
13	0.0132	1873	943	50884	949	8	5
14	0.0145	3754	5097	79244	843	12	4
15	0.0099	3889	4653	81358	439	30	1
16	0.0181	3018	5332	64204	1458	40	1
17	0.0117	3952	5627	47220	747	30	1
18	0.0098	5116	2945	70905	66	41	1
19	0.013	2354	3470	23647	523	11	5
20	0.0094	3303	1624	83464	923	2	6

The BP neural network training process is shown in Figure 4. Where the training number reaches 112 times to satisfy the error 0.01.

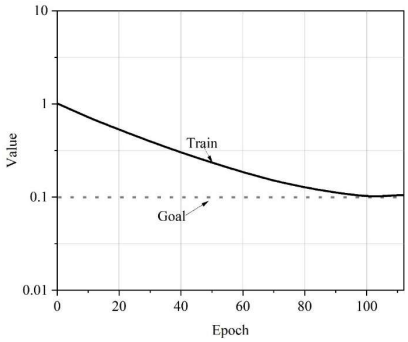

Figure 4: BP neural network training process

Input classification samples to the trained network, classification samples are classified and the results are shown in Table 4. Through BP neural network classification it can be observed that the actual results are the same as the defect results and the defects are classified reliably. This BP neural network has five defective eigenvalues and the input layer has three inputs. The hidden layer can be derived according to the connection characteristics of the neural network, with 15 nodes, and the output is six classes of defects with six nodes. This BP neural network satisfies the classification requirements.

Table 4: Classification of classified samples

Serial number	f	A	d	S	L_{MER}	W_{MER}	Classification result	Actual result
1	0.016	1064	1741	59409	599	15	5	5
2	0.0176	2690	2760	47136	1420	16	5	5
3	0.0019	5517	5630	46498	457	38	1	1
4	0.0386	2484	4233	38953	335	17	5	5
5	0.0096	1336	2260	48412	166	41	1	1
6	0.0093	2824	2160	12466	1227	13	5	5
7	0.0162	958	2989	76769	138	20	4	4
8	0.0047	4929	4654	80563	1700	12	5	5
9	0.0099	1342	578	60506	874	3	6	6
10	0.0109	4882	3035	23281	828	4	6	6
11	0.0063	3424	3562	76462	1831	33	3	3
12	0.0199	3891	1766	75344	1549	40	1	1
13	0.0131	1863	931	50873	940	10	5	5
14	0.0098	3743	5088	79232	836	30	3	3
15	0.0054	3877	4642	81348	429	13	5	5
16	0.0149	3009	5321	64194	1448	41	1	1
17	0.0182	3941	5617	47209	737	3	6	6
18	0.0231	5107	2935	70895	56	28	3	3
19	0.0157	2343	3461	23638	513	8	6	6
20	0.0108	3294	1615	83454	911	41	1	1

IV. Conclusion

The following conclusions are obtained from the study of the application of image feature extraction algorithm and BP neural network in the detection of railroad track defects:

The improved Otsu threshold segmentation method has a significant advantage in rail defect image processing, with a running time of 0.1526 seconds, saving about 0.252 seconds compared with the maximum Shannon entropy multi-threshold segmentation method, and the segmentation effect is more accurate. The threshold segmentation lays the foundation for subsequent feature extraction.

For the six types of orbital defects, namely, wave abrasion, optical band anomaly, serration, abrasion, corrosion and pitting, the five geometric features, namely, frequency, amplitude, area, long-axis and short-axis, have obvious differentiation. The wave abrasion defect has the largest area (84624) and the longest perimeter (5896) among the

six types of defects; the pitting defect has the smallest area (302) and the lowest amplitude (986); and the serration defect has the highest frequency (0.03024). These differences provide an effective basis for BP neural network classification.

The established 5-15-6 structural BP neural network is trained by the gradient descent method, and the error reaches the preset target of 0.01 after 112 trainings, and the classification results of 20 sets of test samples are completely consistent with the actual ones, which verifies the reliability of the network model.

The track defect detection method proposed in this study integrates image processing and artificial intelligence technology, realizes automatic identification and classification of defects, and avoids the subjectivity and instability of traditional manual detection. The method provides a new idea for railroad track condition monitoring and fault early warning, and has important practical value for improving the safe operation of railroads. In the future, the sample library can be further expanded, the neural network structure can be optimized, and the real-time and adaptability of the system can be improved.

References

[1] Zhao, P., & Yuan, D. (2023). China's transport system. In Population Growth and Sustainable Transport in China (pp. 31-58). Singapore: Springer Nature Singapore.

[2] Lin, Y. (2017). Travel costs and urban specialization patterns: Evidence from China's high speed railway system. Journal of Urban Economics, 98, 98-123.

[3] Ferreira, P. A., & López-Pita, A. (2015). Numerical modelling of high speed train/track system for the reduction of vibration levels and maintenance needs of railway tracks. Construction and Building Materials, 79, 14-21.

[4] Gao, Y., Huang, H., Ho, C. L., & Hyslip, J. P. (2017). High speed railway track dynamic behavior near critical speed. Soil Dynamics and Earthquake Engineering, 101, 285-294.

[5] Hodas, S. (2014). Design of railway track for speed and high-speed railways. Procedia Engineering, 91, 256-261.

[6] Li, Z. Q., Huang, W. W., Li, Z., Feng, J. W., Zhang, H., & Li, Q. (2023). Influence of Base Rock Deterioration on Tunnel Bottom Structure Damage under Train Excitation. Advances in Civil Engineering, 2023(1), 4632678.

[7] Bai, W., Zhou, L., Wang, P., Hu, Y., Wang, W., Ding, H., ... & Zhu, M. (2022). Damage behavior of heavy-haul rail steels used from the mild conditions to harsh conditions. Wear, 496, 204290.

[8] Wang, W. J., Guo, H. M., Du, X., Guo, J., Liu, Q. Y., & Zhu, M. H. (2013). Investigation on the damage mechanism and prevention of heavy-haul railway rail. Engineering failure analysis, 35, 206-218.

[9] Yang, H., Wang, Y., Hu, J., He, J., Yao, Z., & Bi, Q. (2021). Segmentation of track surface defects based on machine vision and neural networks. IEEE Sensors Journal, 22(2), 1571-1582.

[10] Dang, D. Z., Lai, C. C., Ni, Y. Q., Zhao, Q., Su, B., & Zhou, Q. F. (2022). Image classification-based defect detection of railway tracks using fiber Bragg grating ultrasonic sensors. Applied Sciences, 13(1), 384.

[11] Alvarenga, T. A., Carvalho, A. L., Honorio, L. M., Cerqueira, A. S., Filho, L. M., & Nobrega, R. A. (2021). Detection and classification system for rail surface defects based on Eddy current. Sensors, 21(23), 7937.

[12] Tastimur, C., Yetis, H., Karaköse, M., & Akin, E. (2016). Rail defect detection and classification with real time image processing technique. International Journal of Computer Science and Software Engineering, 5(12), 283.

[13] Sresakoolchai, J., & Kaewunruen, S. (2022). Railway defect detection based on track geometry using supervised and unsupervised machine learning. Structural health monitoring, 21(4), 1757-1767.

[14] Xiong, Z., Li, Q., Mao, Q., & Zou, Q. (2017). A 3D laser profiling system for rail surface defect detection. Sensors, 17(8), 1791.

[15] Wang, M., Li, K., Zhu, X., & Zhao, Y. (2022). Detection of surface defects on railway tracks based on deep learning. IEEE Access, 10, 126451-126465.

[16] Kumar, A., & Harsha, S. P. (2024). A systematic literature review of defect detection in railways using machine vision-based inspection methods. International Journal of Transportation Science and Technology.

[17] Min, Y., Xiao, B., Dang, J., Yue, B., & Cheng, T. (2018). Real time detection system for rail surface defects based on machine vision. EURASIP Journal on Image and Video Processing, 2018(1), 1-11.

[18] Shi, Y., Zhu, Y. Y., & Wang, J. Q. (2023). Surface Defect Detection Method for Welding Robot Workpiece Based on Machine Vision Technology. Manufacturing Technology, 23(5), 691-699.

[19] Wu, Y., & Lu, Y. (2019). An intelligent machine vision system for detecting surface defects on packing boxes based on support vector machine. Measurement and Control, 52(7-8), 1102-1110.

[20] Yun, Z., Zhi-gang, L. I. N. G., & Yu-qiang, Z. H. A. N. G. (2020). Research progress and prospect of machine vision technology. Journal of graphics, 41(6), 871.

[21] Ren, Z., Fang, F., Yan, N., & Wu, Y. (2022). State of the art in defect detection based on machine vision. International Journal of Precision Engineering and Manufacturing-Green Technology, 9(2), 661-691.

[22] Kumar, A., & Harsha, S. P. (2024). A systematic literature review of defect detection in railways using machine vision-based inspection methods. International Journal of Transportation Science and Technology.

[23] Dubey, A. K., & Jaffery, Z. A. (2016). Maximally stable extremal region marking-based railway track surface defect sensing. IEEE Sensors Journal, 16(24), 9047-9052.

[24] Chang, C. F., Wu, J. L., Chen, K. J., & Hsu, M. C. (2017). A hybrid defect detection method for compact camera lens. Advances in Mechanical Engineering, 9(8), 1687814017722949.

[25] Aldao, E., Fernández-Pardo, L., González-deSantos, L. M., & González-Jorge, H. (2023). Comparison of deep learning and analytic image processing methods for autonomous inspection of railway bolts and clips. Construction and Building Materials, 384, 131472.

[26] Liu, S., Wang, Q., & Luo, Y. (2019). A review of applications of visual inspection technology based on image processing in the railway industry. Transportation Safety and Environment, 1(3), 185-204.

[27] Zhuang, L., Qi, H., Wang, T., & Zhang, Z. (2022). A deep-learning-powered near-real-time detection of railway track major components: A two-stage computer-vision-based method. IEEE Internet of Things Journal, 9(19), 18806-18816.
[28] Kliuiev, P., Latychevskaia, T., Zamborlini, G., Jugovac, M., Metzger, C., Grimm, M., ... & Castiglioni, L. (2018). Algorithms and image formation in orbital tomography. Physical Review B, 98(8), 085426.
[29] Cheng, M. Y., Khasani, R. R., & Setiono, K. (2023). Image quality enhancement using HybridGAN for automated railway track defect recognition. Automation in Construction, 146, 104669.
[30] Li Liangliang,He Chuchao,Wang Peng,Ren Jia,Lü Zhigang,Sha Baolin... & Di Ruohai. (2023). An automatic exposure imaging and enhanced display method of line scan camera for X-ray defect image of solid rocket engine shell. Measurement,217,
[31] Anitha Mary Chinnaiyan & Boyed Wesley Alfred Sylam. (2024). Deep demosaicking convolution neural network and quantum wavelet transform-based image denoising. Network (Bristol, England),21-25.
[32] Shuang Kang,Yinchao He,Wenwen Li & Sen Liu. (2024). Research on Defect Detection of Wind Turbine Blades Based on Morphology and Improved Otsu Algorithm Using Infrared Images. Computers, Materials & Continua,81(1),933-949.
[33] Wang Xing,Wu Mingming,Zhou Jinxia & Wang Jianguo. (2020). Research on Classification Method of Vehicle Body Surface Defects Based on BP Neural Network. IOP Conference Series: Materials Science and Engineering,740,012036-012036.