INTERNATIONAL JOURNAL OF AUTOMOTIVE AND MECHANICAL ENGINEERING
ISSN: 2229-8649 e-ISSN: 2180-1606
VOLUME 22, ISSUE 3, 2025, 12654 - 12664
DOI: https://doi.org/10.15282/ijame.22.3.2025.9.0966

RESEARCH ARTICLE

Defect Detection in Disc Brake Pads Using an Improved AlexNet Convolutional Neural Network

Yu Wei[1,2]*, Jinkai Yin[1,2]

[1]School of Applied Engineering, Henan University of Science and Technology, Sanmenxia, 472000, China
[2]Automotive Academy, Sanmenxia Polytechnic, Sanmenxia, 472000, China

ABSTRACT – In the last few years, with the rapid advancement of the automotive industry, disc brake pads have been increasingly widely used in automotive braking systems. However, due to long-term friction and high temperature effects, they are prone to defects such as cracks, wear, and dents. To raise the accuracy and efficiency of disc brake pad defect detection, the study first optimized image preprocessing and, based on the AlexNet network, introduced a multi-scale convolution module, batch normalization, and global average pooling for structural improvement. Finally, a new type of disc brake pad defect detection model was proposed. The experiment findings denoted that the highest defect detection accuracy of the model was 97.12%, and the mean detection time was 10.53 milliseconds. Compared to other advanced methods, this new model had the lowest missed detection rate of only 1.78%, the lowest resource consumption rate of 36.45%, and the lowest model complexity of 53.23%. This denotes that the new model can efficiently and accurately detect various kinds of defects in disc brake pads, which is suitable for practical industrial inspection scenarios and provides a reliable technical support for the field of disc brake pad defect detection.

ARTICLE HISTORY
Received : 12th Dec. 2024
Revised : 27th May 2025
Accepted : 17th July 2025
Published : 07th Sept. 2025

KEYWORDS
Disc brake pads
Defects
Detection
AlexNet
Preprocessing

1. INTRODUCTION

As a crucial component of modern automotive braking systems, disc brake pads are directly linked to the safety and performance of automobiles. With the continuous development and popularization of automotive technology, the application of disc brake pads is gradually increasing. They utilize friction to achieve braking, enabling vehicles to slow down or stop quickly at high speeds [1]. However, with prolonged use, disc brake pads are affected by various factors, such as friction, temperature, and pressure during operation, which gradually lead to physical defects, including wear, deformation, cracking, and peeling [2]. These defects not only affect the normal service life of brake pads but may also lead to a decrease in braking performance, potentially causing serious safety accidents. Therefore, the detection and identification of defects in disc brake pads has become one of the research hotspots in the field of automotive safety. Feng et al. believed that traditional defect detection methods had poor robustness due to their reliance on manual feature extraction. Therefore, researchers proposed a new method for identifying surface defects on brake discs by combining a Gaussian difference algorithm and a Hough transform algorithm. The experiment findings indicated that the accuracy of this method exceeded 93%, with a false alarm rate of less than 1.5%, significantly outperforming other advanced methods [3]. Gnanasekaran et al. believed that the braking system, as an essential module, requires close attention for continuous monitoring. The research team first extracted defect signals, employed a nested binary method for family classification, and subsequently proposed a new detection model. The experiment findings denoted that the model achieved a maximum accuracy of 93.45% and its performance was relatively stable [4]. Mandziy et al. found that when the brake is applied, the brake linings come into regular contact with the outer cylindrical surface of the metal discs, operate under extreme conditions and are subject to heavy abrasion, which poses certain safety risks. To this end, researchers proposed a method for assessing in-service friction surface degradation [5]. The experimental findings indicated that the detection effectiveness of this method was higher than that of traditional methods, and the assessment of degradation detection severity was more realistic. To enhance the precision and expediency of surface wear identification, Zhao H et al. devised a machine vision-based approach for discerning surface deterioration in automotive metal-ceramic composite brake pads. This methodology was enhanced through the integration of an advanced Retinex algorithm. The outcomes demonstrated that this technique could accurately discern the surface wear patterns of the brake pads with a detection time of less than 500 milliseconds [6].

As computer vision and deep learning technology have advanced, image processing and deep learning-based defect detection methods have gradually emerged as essential research directions for disc brake pad defect detection. Among them, AlexNet was designed by Krizhevsky et al. in 2012. With its deep network structure and powerful image feature extraction ability, it has achieved revolutionary results in image classification tasks [7]. Ullah et al. believed that automatic detection of road cracks may seem simple, but the strength and complexity of the background make it a challenging task. To this end, researchers proposed a road crack category recognition model by combining AlexNet. The experiment's findings indicated that the model's classification accuracy was the highest at 85.20%, and the effect was extremely stable

*CORRESPONDING AUTHOR | Yu Wei | ✉ yuwei202411@126.com

[8]. Jiang W et al. found that traditional neural network models were susceptible to noise interference in rolling bearing fault diagnosis. Therefore, researchers proposed a new diagnostic method by combining AlexNet and variational mode decomposition. The results of the experiment demonstrated the efficacy of the method used in extracting fault features from vibration signals, indicating that it is an effective diagnostic tool with excellent computational efficiency [9]. Sai SA et al. found that existing methods for detecting defects in automotive suspension and brake pads were time-consuming and expensive. Therefore, researchers proposed a new defect fault detection model by combining AlexNet and residual networks. The experiment's findings indicated that this model achieved the highest accuracy in identifying and detecting faults, making it a suitable deep learning model for defect discovery [10]. The vibration detection of car brake pads by Hassan I U and others was easily affected by extreme operating conditions, resulting in poor defect detection results. To this end, researchers proposed a brake pad defect detection method based on sound signals by combining AlexNet. The experiment findings indicated that the defect detection effectiveness under this method was higher than that of traditional methods, and the implementation process was simpler [11].

In summary, previous research has made significant progress in the field of disc brake pad defect detection, but the ability to identify subtle defects is insufficient, the computational complexity is high, and the performance in multiscale feature extraction and real-time performance needs to be improved. To raise the accuracy and efficiency of defect detection in disc brake pads, image preprocessing is studied, including edge extraction, circle fitting, and background removal, to optimize image quality. At the same time, an improved AlexNet network structure is utilized for feature extraction, and the detection ability for subtle defects is enhanced through multiscale convolution. Batch normalization (BN) is used to accelerate convergence and improve the stability of the model. Global average pooling (GAP) is used to minimize the number of parameters and reduce the risk of overfitting. The innovation of the research lies in the targeted improvement of the AlexNet structure, making it more suitable for multi-defect detection tasks of disc brake pads. The contribution of this study lies in the proposed model's enhanced ability to extract subtle defect features. At the same time, by optimizing network parameter configuration, the model achieves efficient multiscale feature fusion and resource utilization, enabling its broader application in automated industrial inspection fields. Additionally, the proposed detection model is suitable for multi-classification tasks, capable of distinguishing between various defect types such as cracks, wear, dents, and hot spots.

The structure of this paper is as follows: Section 1 is the introduction, which reviews the relevant research background and research motivation; Section 2 introduces the image preprocessing process and improved model structure; Section 3 conducts performance experiments and comparative analysis; Section 4 summarizes the research conclusions and looks forward to future work.

2. METHODS AND MATERIALS

The research methodology and technical approach are shown in Figure 1.

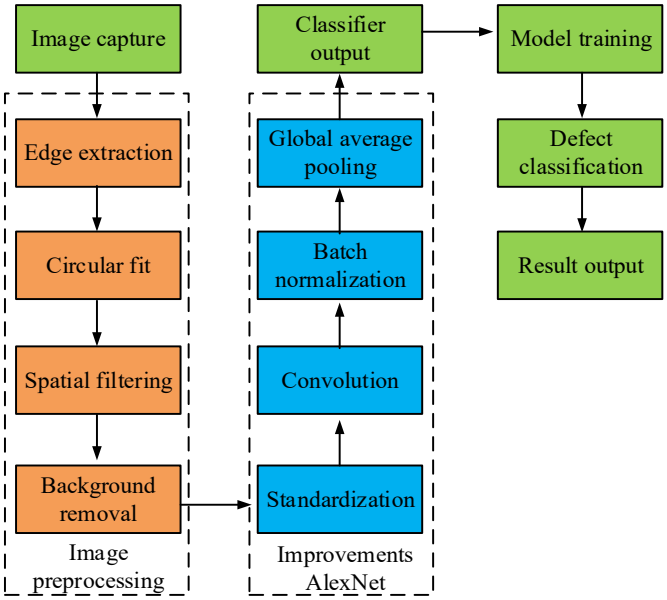
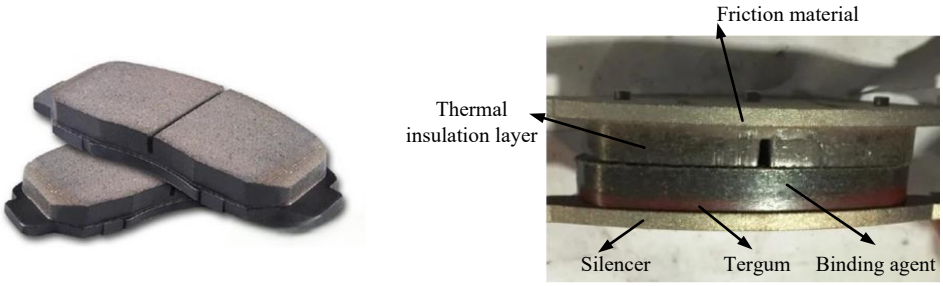


Figure 1. Research methods and technical routes

2.1 Machine Vision Image Preprocessing of Disc Brake Pads

The existing brake pads are divided into drum and disc types, with disc brake pads being the most commonly used. Disc brake pads are the core friction components in modern automotive braking systems, which convert the vehicle's kinetic energy into thermal energy through friction with the brake disc, achieving the functions of deceleration and parking [12-14]. The example and typical structure of disc brake pads are shown in Figure 2.



(a) Disk brake pads in kind (b) Disk brake pad structure

Figure 2. Disc brake pads and their structure display

Figures 2 (a) and (b) show the physical and structural display of the disc brake pad, respectively. In Figure 2, the overall shape of the disc brake pad has a flat and regular appearance, which facilitates close contact with the brake disc. It is mainly composed of friction materials, backing plates, adhesives, sound-absorbing sheets, and insulation layers. Among them, the friction material is the working surface of the brake pad, directly in contact with the brake disc, providing braking force and withstanding high temperature and friction. The back plate is located behind the friction material and is made of high-strength steel to support and transmit braking force, ensuring that the brake pads remain stable during braking [15-17]. However, due to the complex surface texture of disc brake pads and their susceptibility to factors such as lighting changes and background interference, traditional image preprocessing processes have certain limitations in processing such images. For example, conventional grayscale, binarization, and edge detection methods are easily affected by noise, leading to inaccurate feature extraction [18-20]. Therefore, a more suitable process for image preprocessing of disc brake pads has been proposed, as shown in Figure 3.

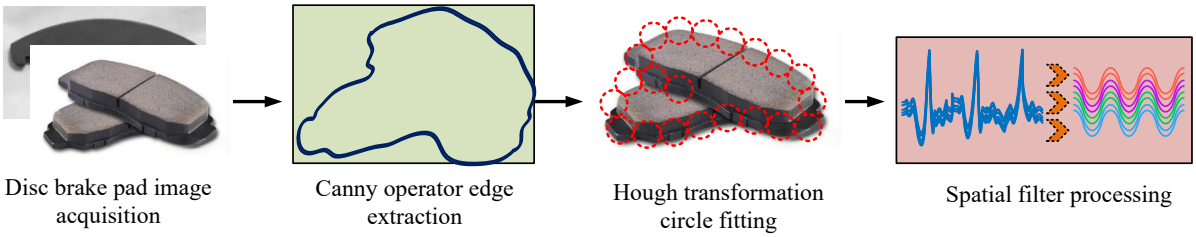


Figure 3. New image preprocessing process for disc brake pads

In Figure 3, the entire process is divided into four modules, namely edge extraction, circle fitting, spatial filtering, and background region removal. Firstly, the Canny operator is used to extract edges from the image and obtain clear contour information. Next, the Hough transform is used for circle fitting to accurately identify and segment the contour of the brake pad. Then, spatial filtering methods such as median filtering are utilized to remove noise and texture interference from the image. Finally, the background removal algorithm is used to remove complex background areas, retaining only the brake pad areas related to defect detection. In the edge extraction stage, firstly, Gaussian filtering is applied to the input image to smooth it out and reduce the impact of noise. The calculation is denoted in Eq. (1).

$$G(x,y)=\frac{1}{2\pi\sigma^2}\exp(-\frac{x^2+y^2}{2\sigma^2}) \tag{1}$$

In Eq. (1), $G(x,y)$ means the Gaussian kernel function; σ stands for standard deviation; x and y respectively indicate the horizontal and vertical coordinates of the image. The calculation method for the gradient amplitude and gradient direction of the filtered image is shown in Eq. (2).

$$\begin{cases} M(x,y)=\sqrt{(I_x^{'})^2+(I_y^{'})^2} \\ \theta(x,y)=\arctan(\frac{I_x^{'}}{I_y^{'}}) \end{cases} \tag{2}$$

In Eq. (2), $M(x,y)$ indicates the gradient amplitude; $M(x,y)$ means the gradient direction; $I_x^{'}$ and $I_y^{'}$ represents the first derivative of the image in the x and y directions, respectively. In addition, the edges of disc brake pads usually have regular circular or approximately circular structures, so fitting the detected edges with circles can accurately locate the contour range of the brake pads. This time, the Hough transform is used, and the calculation expression is indicated in Eq. (3).

$$(x-a)^2+(y-b)^2=r^2 \tag{3}$$

In Eq. (3), (a,b) represents the center coordinates of the fitted circle; r represents the radius of the circle. By using the voting method to search for the center and radius in the parameter space, the circle corresponding to the maximum number of votes is found, which is the fitted edge of the brake pad [21-23]. In addition, to remove possible high-frequency noise and texture interference in the image, spatial filtering was studied, as shown in Figure 4.

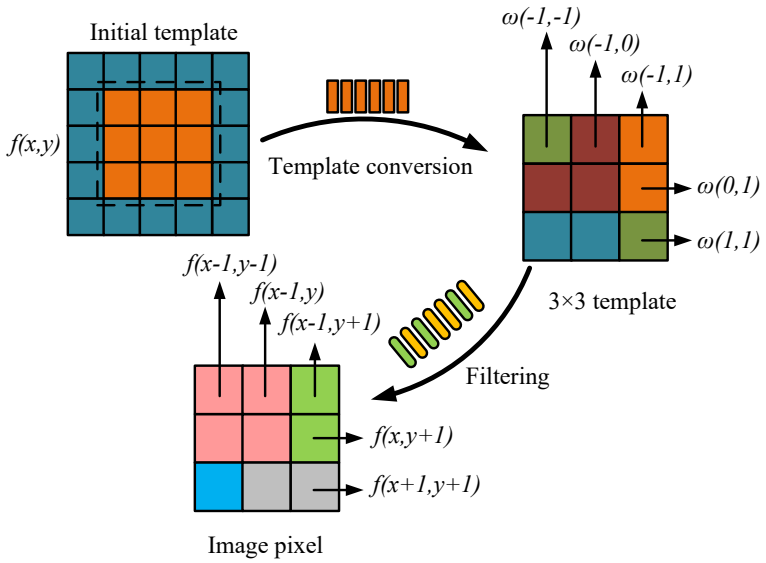


Figure 4. Spatial filtering diagram

In Figure 4, the basic process of spatial filtering is to use a 3 × 3 filtering template to perform local operations on each pixel in the image. The filtering template gradually moves on the image, combining the pixel values within the coverage area with the weights in the template to calculate new pixel values. Specifically, when each pixel $f(x,y)$ in the image is covered by the template, it will be weighted and summed with the corresponding weight $\omega(i,j)$ in the template to obtain the filtered pixel value. The pixel values after filtering are calculated as denoted in Eq. (4).

$$g(x,y)=\sum_{i=-1}^{1}\sum_{j=-1}^{1}\omega(i,j)\cdot f(x+i,y+i) \tag{4}$$

In Eq. (4), $g(x,y)$ indicates the pixel value of the filtered image; $f(x+i,y+i)$ means the pixel value of the image at the $(x+i,y+i)$ position; $\omega(i,j)$ means the weights in a 3x3 filtering template. Finally, to remove complex background interference, an adaptive method of local binarization is adopted, and the calculation expression for local binarization is shown in Eq. (5).

$$T(x,y)=\frac{1}{|W|}\sum_{(i,j)\in W}I^{''}(i,j)-C \tag{5}$$

In Eq. (5), $T(x,y)$ represents the threshold for local binarization; $|W|$ represents the local window size; C represents a constant offset; $I^{''}(i,j)$ means the pixel value of the image at (i,j) after multiple processing. Based on the $T(x,y)$ threshold, the background area is removed and only the foreground target, namely the disc brake pad, is retained. The calculation expression is shown in Eq. (6).

$$B(x,y)=\begin{cases}1 & if I^{''}(x,y)\geq T(x,y)\\ 0 & if I^{''}(x,y)<T(x,y)\end{cases} \tag{6}$$

In Eq. (6), $B(x,y)$ represents the binarized image; $I^{''}(x,y)$ represents the image after three processing steps. A value of 1 represents the foreground area, which is the brake pad, and a value of 0 represents the background area.

2.2 Construction of A Disc Brake Pad Defect Detection Model Based on Improved AlexNet

After completing the machine vision image preprocessing of disc brake pads, to further raise the accuracy and real-time effectiveness of defect detection, AlexNet is introduced to address the complexity and diversity of disc brake pad defects. Compared to other methods, such as support vector machines or decision tree models based on manual feature extraction, AlexNet has stronger automated feature extraction capabilities and can extract low-level to high-level multi-level features from preprocessed images [24-25]. The network structure of AlexNet is shown in Figure 5.

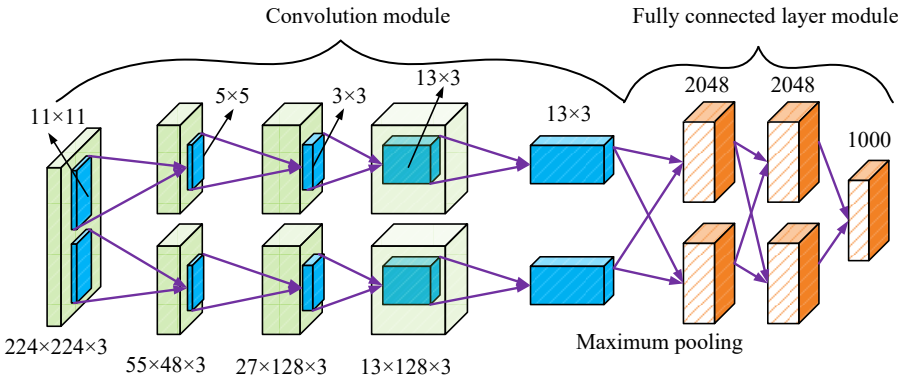


Figure 5. AlexNet network structure

As shown in Figure 5, the network structure of AlexNet consists of 8 layers of backbone, including five convolutional layers and three fully connected layers. The input image size is 224 × 224 × 3. Firstly, a convolution operation is performed using an 11 × 11 convolution kernel with a step size of 4 to extract low-level features of the image, which are then downsampled through a max pooling layer. Next, the second convolutional layer uses a 5 × 5 convolution kernel and combines it with a max pooling operation to further extract intermediate features. In the 3rd to 5th convolutional layers, AlexNet uses a 3x3 convolution kernel to extract deep features from the image. Subsequently, the output of the convolutional layer is unfolded and passed to the fully connected layer, where the first and second fully connected layers each contain 2048 neurons for high-level feature fusion and learning. The final third fully connected layer is the Softmax classification layer, which outputs 1000 categories [26-28]. The convolution calculation method of AlexNet is denoted in Eq. (7).

$$O(i,j,k)=\sum_{m=0}^{M-1}\sum_{n=0}^{N-1}W(m,n,k)\cdot I(i+m,j+n)+b(k) \tag{7}$$

In Eq. (7), $O(i,j,k)$ means the pixel values of the outputting feature map of the convolutional layer at positions (i,j) and channel k; $W(m,n,k)$ represents the weight of the convolution kernel at position (m,n) and channel k; $I(i+m,j+n)$ means the pixel value of the input image at position $(i+m,j+n)$; $b(k)$ represents the bias of the k th convolution kernel; M and N respectively indicate the height and width of the convolution kernel. In addition, AlexNet uses Local Response Normalization (LRN) to normalize the response of local features and suppress irrelevant features [29-31]. The calculation method for LRN is denoted in Eq. (8).

$$\hat{\alpha}_{i,j}=\frac{\alpha_{i,j}}{(\psi+\alpha\sum_{m=\max(0,i-n/2)}^{\min(N-1,i+n/2)}\alpha_{m,j}^{2})^{\beta}} \tag{8}$$

In Eq. (8), $\hat{\alpha}_{i,j}$ represents the normalized activation value; $\alpha_{i,j}$ represents the input activation value; ψ, α, and β all represent hyperparameters used to control the strength of normalization. By suppressing the activation values within the local neighborhood, the features with high activation values are made more prominent. The calculation method for fully connected layers is shown in Eq. (9).

$$y_i=\sum_{j=1}^{N}\omega_{ij}\cdot x_j+b_i \tag{9}$$

In Eq. (9), y_i represents the output value of the i th neuron in the fully connected layer; x_j represents the j th element in the input feature vector; ω_{ij} represents the weight connecting the j th input to the i th neuron; b_i represents the bias term of the i th neuron. Although AlexNet performs well in image recognition, there are issues with insufficient ability to recognize subtle defects and high computational complexity in detecting defects in disc brake pads. For this purpose, an improved AlexNet model is proposed, whose structure is shown in Figure 6.

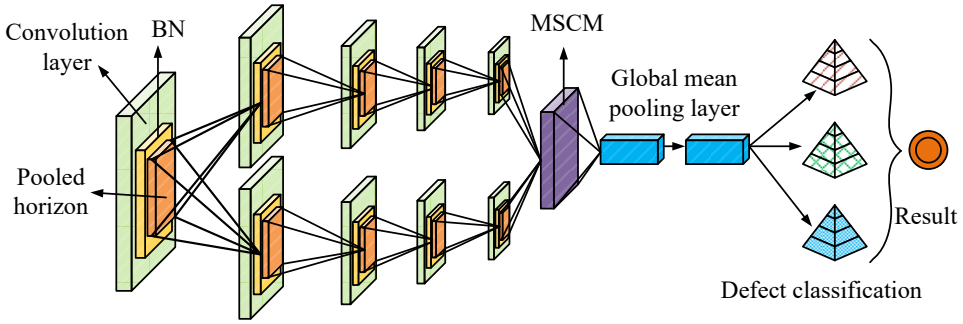


Figure 6. Improved AlexNet model structure

In Figure 6, compared to the original AlexNet, the improved model has undergone three structural optimizations. Firstly, a BN layer is added and inserted into each convolutional and activation layer, which improves the convergence speed and stability of the model. Secondly, two large fully connected layers are removed and replaced with GAP layers, significantly reducing the number of parameters and lowering the risk of overfitting. Finally, a multiscale convolution module (MSCM) is introduced, combined with convolution kernels of different sizes, to strengthen the model's ability to recognize subtle defects. The calculation method for the MSCM is shown in Eq. (10).

$$F_{multi}(i,j)=\sum_{s=1}^{S}(W_s * I)(i,j) \tag{10}$$

In Eq. (10), $F_{multi}(i,j)$ represents the output feature value of the MSCM at position (i,j); W_s represents the s th scale convolution kernel; S represents the number of convolution kernels at different scales. BN is used instead of LRN. The reason for choosing BN is that, compared to LRN, it can reduce the impact of model parameter changes through batch processing, ensuring the convergence and capacity of the model. The BN calculation method is denoted in Eq. (11).

$$\hat{x}=\frac{x-\mu}{\sqrt{\vartheta^2+\varepsilon}}, y=\gamma\hat{x}+\zeta \tag{11}$$

In Eq. (11), $\hat{x}$ represents the standardized activation value; μ and ϑ^2 represent the mean and variance of the previous batch; ε stands for the minimum constant that prevents the denominator from being zero; γ and ζ represent learnable scaling and translation parameters. The improved AlexNet model is applied to the defect detection of disc brake pads, and a new detection model is constructed. The process is shown in Figure 7.

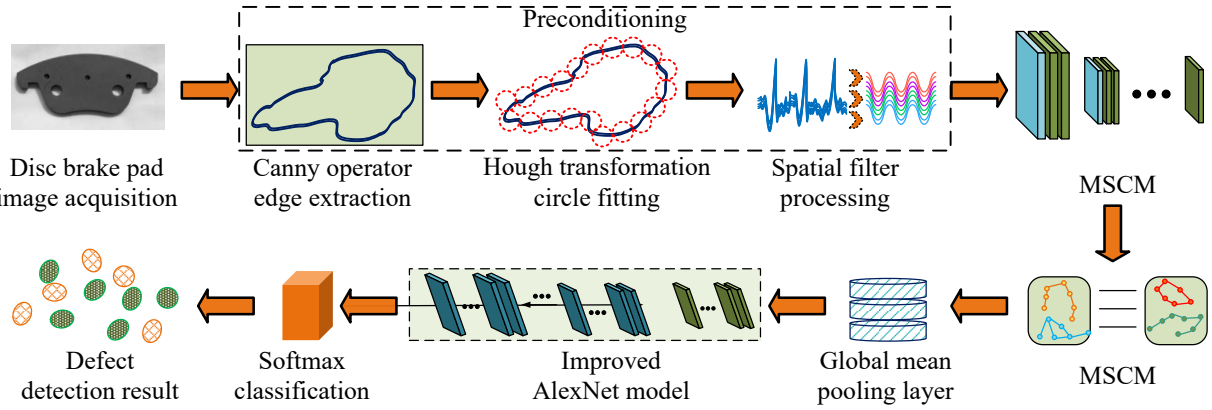


Figure 7. New disc brake pad defect detection model flow

In Figure 7, firstly, the image preprocessing stage extracts the image contour through Canny edge detection, accurately locates the brake pad area using the Hough circle transform, removes noise using median filtering and spatial filtering, and finally uses a local binarization algorithm to remove complex background and preserve the key target area. Next, an improved AlexNet model is used in the feature extraction stage to enhance the extraction of defect features at different scales through MSCMs. BN improves the convergence speed and stability of the model, while GAP effectively reduces the number of parameters and enhances the generalization ability. Finally, in the defect classification stage, the Softmax classification layer is used to classify features, which can automatically identify defect types such as normal, crack, wear, and deformation.

3. RESULTS

3.1 Performance Testing of A New Type of Disc Brake Pad Defect Detection Model

A suitable experimental environment was set up for the study, with Intel Core i7-9700K CPU, NVIDIA GeForce RTX 2080 Ti GPU, 32GB DDR4 memory, Ubuntu 18.04 system environment, and TensorFlow 2.4 deep learning framework. The Mechanical Components Surface Defects Dataset (MSD) and the Structural Defect NETwork Dataset (SDNET) are used as the test data sources. The MSD dataset focuses on the detection of surface defects in mechanical components, including common defects such as scratches, cracks, and pits. The dataset contains approximately 3,200 high-resolution defect images (640×480), all captured using industrial cameras. The annotation information was completed by professional engineers, covering four typical defect categories, with image sources spanning various components and shooting angles. The SDNET dataset includes defect images on structural material surfaces, with defect types including cracks, wear, and corrosion. Although primarily focused on defect detection in structural materials, the crack and wear images exhibit texture features similar to those of disc brake pads, making them valuable for transfer learning. The dataset includes approximately 2,600 images with a resolution of 512×512. Image annotations combine segmentation masks with category labels, enhancing the ability to model defect boundaries with precision during training. The study conducted

value selection tests on the two types of hyperparameters that have the greatest impact on model performance, namely the scaling parameter γ and translation parameter ζ. The test results are shown in Figure 8.

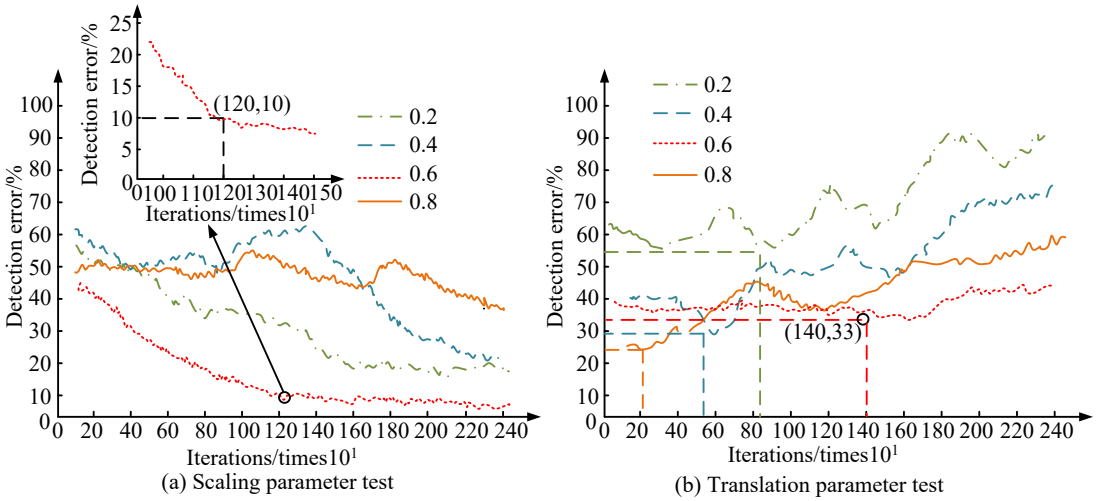


Figure 8. Hyperparameter selection test

Figure 8 (a) showcases the selection test findings of the scaling parameter γ, and Figure 8 (b) showcases the selection test findings of the translation parameter ζ. According to Figure 8 (a), when the number of iterations reached about 1200, the detection error rate was the lowest at a scaling parameter of 0.6, which was about 10%, and it was relatively stable in subsequent iterations. The detection error rates of scaling parameters 0.2, 0.4, and 0.8 were all higher than 0.6, indicating that 0.6 was the optimal choice for scaling parameters. In Figure 8 (b), there were also differences in detection error rates among different translation parameters. For example, when the translation parameters were 0.4 and 0.6, the detection error rate was low in the early stage, but after more than 1000 iterations, the detection error rate of parameter 0.4 began to fluctuate and increase, while parameter 0.6 remained at a low level and stabilized at about 33% after around 1400 iterations. Therefore, a translation parameter of 0.6 could be used as a better choice for hyperparameter configuration in the model. The study used scaling and translation parameters of 0.6 for subsequent testing, and the ablation test findings of the model are indicated in Figure 9.

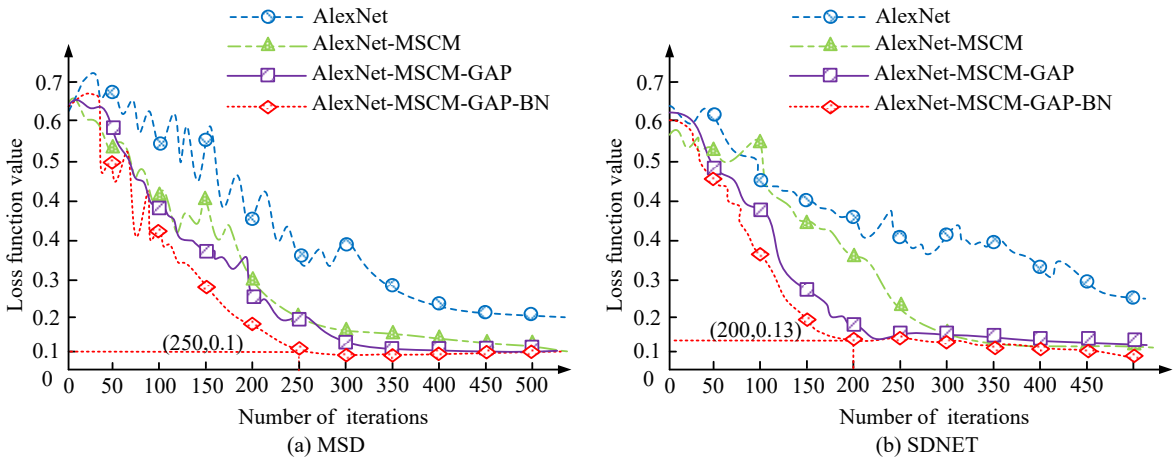


Figure 9. Ablation test results

Figures 9 (a) and (b) show the ablation test results of the model in the MSD and SDNET datasets, respectively. In Figure 9 (a), the loss function value of the basic AlexNet model was relatively high, and the convergence speed slowed down with the increase in iteration times. After adding MSCM, the convergence speed and loss value of the model were improved, but the effect had not yet reached the optimal level. After adding GAP, the loss value of the model further decreased and stabilized after about 250 iterations. After combining BN, the improved model had the fastest convergence speed, the lowest loss value, and finally stabilized around 0.1, showing the best performance. In Figure 9 (b), the efficacy of the model in the SDNET dataset was similar to that in the MSD dataset. When MSCM, GAP, and BN were all combined, the model converged to the lowest loss value of around 0.13 after 200 iterations, showing the best performance. The above findings denoted that the combination of MSCM, GAP, and BN could greatly raise the convergence speed and detection accuracy of the model. It compared advanced defect detection models of the same type, such as Efficient, You Only Look Once version 5 (YOLOv5), and Faster Region-based Convolutional Neural Network (Faster R-CNN). The test results are shown in Table 1, using precision (P), recall (R), and average detection time as indicators.

Table 1. Multi-index test results of different models

Data set	Model	P/%	R/%	Mean detection time/ms
MSD	Efficient	98.41	72.58	24.32
	YOLOv5	88.36	74.75	33.74
	Faster R-CNN	90.06	89.89	11.03
	Our model	97.12	84.75	10.53
	IGD-IHT + PIQEDS-IBPSO-NESN	91.8	78.94	18.41
	AlexNet-Residual	89.45	76.12	21.66
SDNET	Efficient	83.06	93.01	20.03
	YOLOv5	92.25	93.24	19.82
	Faster R-CNN	83.97	87.06	29.87
	Our model	95.36	92.82	13.71
	IGD-IHT + PIQEDS-IBPSO-NESN	87.73	88.66	22.35
	AlexNet-Residual	90.28	85.53	25.48

According to Table 1, in the MSD dataset, the Efficient model achieves the highest accuracy of 98.41%, but it has a longer detection time; YOLOv5 is faster but has lower accuracy; Faster R-CNN has a better recall rate, at 89.89%. The IGD-IHT joint model proposed by Feng et al. is slightly more accurate than YOLOv5, while Sai et al.'s AlexNet-Residual has a slightly lower recall rate and longer detection time. In contrast, the model proposed in this study achieves a precision rate of 97.12% and a recall rate of 84.75% in the MSD dataset, with an average detection time of only 10.53ms, achieving a good balance between accuracy and speed. On the SDNET dataset, our model continues to demonstrate strong stability and practicality, achieving a precision rate of 95.36%, a recall rate of 92.82%, and the shortest detection time of 13.71ms, outperforming most comparison models and demonstrating high overall performance and engineering deployment value. To further verify the stability of the model in complex environments, the study conducted five independent experiments on the MSD and SDNET datasets, recording the average F1 value and standard deviation of each model to assess performance fluctuations.

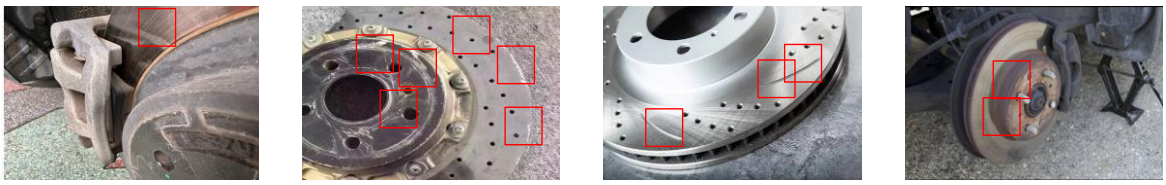
Table 2. Uncertainty analysis table

Dataset	Model	Mean F1 Score (%)	Standard Deviation (%)
MSD	YOLOv5	81.23	1.28
	Faster R-CNN	87.71	1.42
	Proposed Method	91.55	0.61
SDNET	YOLOv5	85.46	1.33
	Faster R-CNN	89.92	1.26
	Proposed Method	92.08	0.57

As shown in Table 2, the proposed method exhibits the smallest standard deviation of F1 scores (±0.61% and ±0.57%) on both datasets, significantly outperforming YOLOv5 (±1.28% and ±1.33%) and Faster R-CNN (±1.42% and ±1.26%). This indicates that the proposed model not only has an advantage in terms of accuracy but also maintains highly consistent detection performance across multiple independent tests, demonstrating excellent robustness and application stability. Especially on the SDNET dataset, the standard deviation of the F1 score for the proposed method is kept within ±0.6%, indicating its strong adaptability to different image sources and background interference, making it suitable for deployment in real-world industrial inspection scenarios for stable operation.

3.2 Simulation Detection of Defect Detection Model for New Disc Brake Pads

To verify the practical application effect of the new disc brake pad defect detection model, four common types of disc brake pad defect images were randomly selected from the MSD dataset for detection and comparison, such as cracks, wear, dents, and hot spots. The images of four types of brake pad defects are shown in Figure 10.



(a) Crack defect (b) Wear defect (c) Sunken defect (d) Heat spot defect

Figure 10. Four types of typical disc brake pad defects

Figures 10 (a), (b), (c), and (d) show the crack, wear, depression, and hot spot defects of the disc brake pad, respectively. Based on the above data, the F1 scores of four types of models were further tested for different defect detections, and the results are shown in Figure 11.

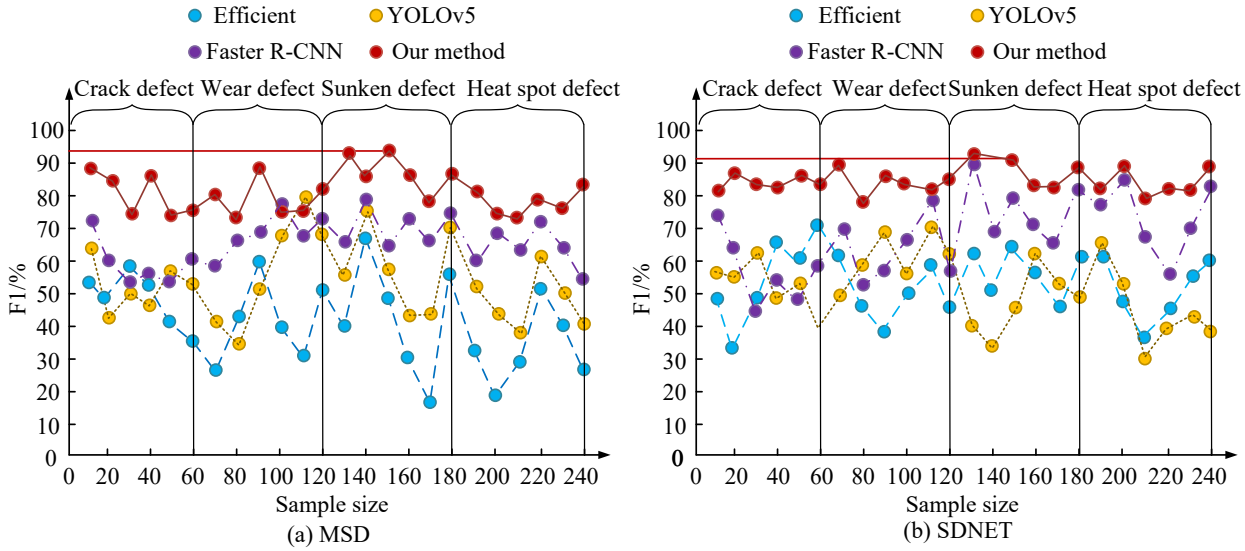


Figure 11. Test results of different algorithmic models for intersection-parallel ratio test

Figures 11 (a) and (b) show the detection outcomes of four types of models on different defects in the MSD and SDNET datasets. In Figure 11(a), the proposed model performed best in detecting various types of defects, particularly in crack and depression defect detection, with F1 scores reaching up to 93.4%. The overall performance of Faster R-CNN was second, with good performance in crack defect detection, achieving an F1 score of approximately 85.6%. However, it was relatively weak in pit defect detection, with F1 scores fluctuating between 70% and 75%. YOLOv5 and Efficient had significant fluctuations in the detection of wear and pit defects, with F1 scores ranging from 60% to 80%, and their stability was not as good as the previous two. From Figure 11(b), the detection efficiency of the proposed model for the four types of defects was relatively low. However, the effect of detecting concave defects was still the best, with an F1 score of 90.3%, indicating stable detection performance. The detection performance of the other three models was significantly lower than that of the proposed method. Further testing was conducted using metrics such as missed detection rate, resource consumption rate, and model complexity; the findings are presented in Table 3.

Table 3. Detection results of real defect detection indexes of different models

Data set	Model	Missed detection rate /%	Resource consumption rate /%	Model complexity /%
Crack defect	Efficient	5.94	61.21	61.93
	YOLOv5	7.44	69.96	69.79
	Faster R-CNN	6.42	66.69	51.32
	Our model	5.93	42.24	53.23
Wear defect	Efficient	4.81	78.72	72.85
	YOLOv5	6.81	67.95	73.19
	Faster R-CNN	4.94	47.69	96.06
	Our model	9.03	36.83	67.73
Sunken defect	Efficient	9.67	47.16	55.17
	YOLOv5	4.45	58.42	60.59
	Faster R-CNN	8.13	68.66	78.83
	Our model	5.76	36.68	34.22
Heat spot defect	Efficient	6.11	51.31	76.67
	YOLOv5	9.33	44.88	76.94
	Faster R-CNN	1.64	55.87	44.73
	Our model	1.78	36.45	39.02

According to Table 3, the proposed model and Faster R-CNN performed well in terms of missed detection rate, particularly in hotspot defect detection. The missed detection rate of the proposed model was 1.78%, while that of Faster R-CNN was 2.54%. In comparison, YOLOv5 Efficient reached 5.33% and 6.72%, respectively. In terms of resource

consumption, YOLOv5 demonstrated superior efficiency, with a resource consumption of only 25.67% in crack defect detection, making it well-suited for detection scenarios that require high real-time performance. In terms of model complexity, the proposed model had a relatively low complexity, achieving 35.89% in hotspot defect detection, whereas Faster R-CNN had a higher complexity, reaching 82.34%. Overall, the model proposed by the research achieved a good balance between missed detection rate, resource consumption rate, and model complexity, making it especially suitable for multi-defect detection tasks involving disc brake pads, and demonstrating strong comprehensive performance and adaptability.

4. CONCLUSION

The various defects caused by high temperature and friction in disc brake pads during long-term use can easily lead to a decrease in braking performance, posing a threat to vehicle driving safety. To address the issues of insufficient accuracy and poor robustness in traditional methods for detecting defects in disc brake pads, a novel defect detection model was developed by combining image preprocessing and an improved AlexNet network. The experiment findings denoted that the detection error rate of the new model reached a minimum of 10% when both the scaling and translation parameters were set to 0.6. At the same time, the ablation test verified the performance promoting effect of each module of the new model on the final model, which could achieve a minimum loss function value of 0.1. Compared to other advanced models, this new model had the highest P-value of 97.12%, the highest R-value of 92.82%, and the shortest average processing time of 10.53ms, demonstrating excellent defect detection efficiency and accuracy. In the defect detection of four types of disc brake pads, the proposed model performed the best, especially in detecting crack and depression defects, with F1 scores reaching up to 93.4%, significantly outperforming the Faster R-CNN model's 85.6%.

Meanwhile, the new model achieved the lowest missed detection rate of 1.78%, while Faster R-CNN had a rate of 2.54%. The new model also exhibited the lowest resource consumption rate of 36.45% and the lowest model complexity of 53.23%. However, for precision mechanical parts such as brake pads, this study did not discuss the impact of lighting conditions on defect detection, particularly the extraction of defect features. Future research could incorporate considerations of lighting factors to optimize the model's feature extraction capabilities further, thereby enhancing its robustness and broad adaptability. Additionally, future work could integrate attention mechanisms or Transformer architectures to enhance the model's ability to identify small defects in complex environments. Simultaneously, lightweight model designs and edge deployment strategies could be explored to enable more efficient, low-power, real-time detection in industrial terminal environments.

ACKNOWLEDGEMENT

The author would like to thank the School of Applied Engineering, Henan University of Science and Technology, Sanmenxia, China, for providing laboratory facilities and its support.

CONFLICT OF INTEREST

The authors declare no competing interests.

AUTHORS CONTRIBUTION

Yu Wei: Data collection, data organization, table preparation, manuscript writing, review, and editing.

Jinkai Yin: Data analysis, manuscript writing, review, and editing.

REFERENCES

[1] T. Dózsa, P. Őri, M. Szabari, E. Simony, A. Soumelidis, and I. Lakatos, “Brake disc deformation detection using intuitive feature extraction and machine learning,” *Machines*, vol. 12, no. 4, pp. 214-217, 2024.

[2] J. Cao, J. Bao, Y. Yin, W. Yao, T. Liu, and T. Cao, “Intelligent prediction of wear life of automobile brake pad based on braking conditions,” *Ind. Lubr. Tribol*., vol. 75, no. 2, pp. 157-165, 2023.

[3] L. Feng, Z. Yu, J. Gao, and Q. An, “Surface defect detection of brake discs based on the IGD-IHT algorithm and the PIQEDS-IBPSO-NESN Algorithm,” *Instrumentation*, vol. 11, no. 3, pp. 62-73, 2024.

[4] S. Gnanasekaran, L. P. Jakkamputi, J. Rakkiyannan, M. Thangamuthu, and Y. Bhalerao, “A comprehensive approach for detecting brake pad defects using histogram and wavelet features with nested dichotomy family classifiers,” *Sensors*, vol. 23, no. 22, pp. 9093-9095, 2023.

[5] T. Mandziy, I. Ivasenko, O. Berehulyak, R. Vorobel, M. Bembenek, and S. Kryshtopa, et al., “Evaluation of the degree of degradation of brake pad friction surfaces using image processing,” *Lubricants*, vol. 12, no. 5, pp. 172-181, 2024.

[6] H. Zhao, and T. Wang, “Surface wear detection of automotive cermet composite brake pads based on machine vision,” *International Journal of Microstructure and Materials Properties*, vol. 17, no. 2, pp. 151-170, 2024.

[7] K. Kılıç, and U. Özcan, “AlexNet architecture optimized for wood defect detection,” *Bozok Journal of Engineering and Architecture*, vol. 2, no. 2, pp. 20-28, 2023.

[8] A. Ullah, H. Elahi, Z. Sun, and A. Khatoon, "Comparative analysis of AlexNet, ResNet18 and SqueezeNet with diverse modification and arduous implementation," *Arabian Journal for Science and Engineering*, vol. 47, no. 2, pp. 2397-2417, 2021.

[9] W. Jiang, Z. Qi, A. Jiang, S. Chang, and X. Xia, "Lightweight network bearing intelligent fault diagnosis based on VMD-FK-ShuffleNetV2," *Machines*, vol. 12, no. 9, pp. 608-609, 2024.

[10] S. A. Sai, S. N. Venkatesh, S. Dhanasekaran, P. A. Balaji, V. Sugumaran, and N. Lakshmaiya, et al., "Transfer learning based fault detection for suspension system using vibrational analysis and radar plots," *Machines*, vol. 11, no. 8, pp. 778-779, 2023.

[11] U. Hassan, K. Panduru, and J. Walsh, "Review of data processing methods used in predictive maintenance for next generation heavy machinery," *Data*, vol. 9, no. 5, pp. 69-72, 2024.

[12] H. R. Hong, and C. W. Ha, "Optimizing friction block location on brake pads for high-speed railway vehicles using artificial neural networks," *Applied Sciences*, vol. 13, no. 17, pp. 9634-9637, 2023.

[13] Z. Kang, M. Zhang, and W. Cheng, "Health status monitoring of high-speed train brake pads considering noise under variable working conditions," *Structural Health Monitoring*, vol. 23, no. 3, pp. 1449-1467, 2023.

[14] S. Wang, Z. Yu, J. Wang, and S. Chen, "Research on CNN-LSTM brake pad wear condition monitoring based on GTO multi-objective optimization," *Actuators*, vol. 12, no. 7, pp. 301-307, 2023.

[15] A. De Martin, G. Jacazio, V. Parisi, and M. Sorli, "Prognosis of wear progression in electrical brakes for aeronautical applications," *PHM Society European Conference*, vol. 7, no. 1, pp. 329-337, 2022.

[16] H. Pan, W. Jiao, T. Yan, A. Ur Rehman, A. Wan, and S. Yang, "Combining kernel principal component analysis and spatial group-wise enhance convolutional neural network for fault recognition of rolling element bearings," *Measurement Science and Technology*, vol. 34, no. 12, pp. 125003-125007, 2023.

[17] X. Yin, S. Zhang, Y. Zhang, Z. Pang, and B. Zhang, "Friction performance prediction of automotive pads under operating conditions using attention-based CNN-BiLSTM deep learning framework," *Journal of Mechanical Science and Technology*, vol. 38, no. 8, pp. 4135-4144, 2024.

[18] X. Hao, Y. Ding, B. Liu, X. Yang, C. Gao, and K. Tian, et al., "Study on vibration and sound signals for fault diagnosis of hoist disc brake system," *Australian Journal of Mechanical Engineering*, vol. 21, no. 3, pp. 979-990, 2021.

[19] S. Wang, L. Zhong, Y. Niu, S. Liu, S. Wang, and K. Li, et al., "Prediction of frictional braking noise based on brake dynamometer test and artificial intelligent algorithms," *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*, vol. 236, no. 12, pp. 2681-2695, 2021.

[20] S. Pal, A. Roy, P. Shivakumara, and U. Pal, "Adapting a swin transformer for license plate number and text detection in drone images," *Artificial Intelligence and Applications*, vol. 1, no. 3, pp. 145-154, 2023.

[21] M. I. M. Ameerdin, M. H. Jamaluddin, A. Z. Shukor, and S. Mohamad, "A review of deep learning-based defect detection and panel localization for photovoltaic panel surveillance system," *International Journal of Robotics and Control Systems*, vol. 4, no. 4, pp. 1746-1771, 2024.

[22] B. Benbouya, H. Cheghib, M. Behim, M. M. Mahmoud, M. F. Elnaggar, N. F. Ibrahim, and N. Anwer, "Dynamic assessment and control of a dual star induction machine state dedicated to an electric vehicle under short-circuit defect," *International Journal of Robotics and Control Systems*, vol. 4, no. 4, pp. 1731-1745, 2024.

[23] F. Furizal, A. Ma'arif, A. A. Firdaus, and W. Rahmaniar, "Future potential of E-nose technology: A review," *International Journal of Robotics and Control Systems*, vol. 3, no. 3, pp. 449-469, 2023.

[24] Y. Cui, Z. Zhang, Z. Zhong, J. Hou, Z. Chen, Z. Cai, and J. Kim, "Bearing fault diagnosis based on multiscale lightweight convolutional neural network," *Processes*, vol. 13, no. 4, pp. 1239-1243, 2025.

[25] A. Liao, D. Hu, R. Liu, W. Shi, and Y. Huang, "Fault diagnosis of wheel tread based on deep transfer convolution neural network," *Journal of Failure Analysis and Prevention*, vol. 25, no. 1, pp. 314-329, 2025.

[26] F. Akbalik, A. Yildiz, and Ö. F. Ertuğrul, "Enhancing vehicle fault diagnosis through multi-view sound analysis: integrating scalograms and spectrograms in a deep learning framework," *Signal, Image and Video Processing*, vol. 19, no. 2, pp. 182-189, 2025.

[27] S. Dou, F. Dai, Y. Du, F. Liu, T. Li, D. Yao, and H. Bai, "Deep learning-based fault diagnosis of gearbox using CBAM-DSC-ResNeXt50 structure," *Journal of Vibration Engineering & Technologies*, vol. 13, no. 6, pp. 418-421, 2025.

[28] Z. Wu, S. Cui, S. Tan, R. Hao, and Q. Zhang, "Dual time-delay Rényi entropy: A new wide-area information fusion and quantification method applied to intelligent fault diagnosis of rotating machinery," *Insight-Non-Destructive Testing and Condition Monitoring*, vol. 67, no. 4, pp. 232-241, 2025.

[29] L. Wang, X. Fan, X. Yu, Z. Huang, and K. Zhao, "A review of road surface recognition and tyre-road friction coefficient estimation methods," *International Journal of Heavy Vehicle Systems*, vol. 32, no. 3, pp. 327-367, 2025.

[30] D. Wang, X. Pan, J. Ma, Y. Dai, C. Sun, S. Li, and X. Yin, "Research on wind turbine blade fault detection based on DenseNet-TL combined with ELM," *Journal of Applied Science and Engineering*, vol. 28, no. 12, pp. 2451-2460, 2025.

[31] H. Fan, B. Chen, X. Cao, Q. Li, H. Xu, and T. Zhang, "An intelligent multi-element fault diagnosis method of rolling bearings considering damage degrees and sensor abnormity under small samples," *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, vol. 239, no. 7, pp. 2334-2356, 2025.