Research Article

Rongcai Wang, Hao Yan, Enzhi Dong, Zhonghua Cheng, Yuan Li*, and Xisheng Jia*

Infrared thermography based fault diagnosis of diesel engines using convolutional neural network and image enhancement

https://doi.org/10.1515/phys-2024-0110
received July 07, 2024; accepted December 13, 2024

Abstract: Diesel engines find extensive application in various production sectors, including industry and agriculture. Strengthening the condition monitoring and fault diagnosis of diesel engines is of paramount importance in ensuring the smooth operation of production systems. Timely detection and elimination of defects play a crucial role in maintaining the normal functioning of these systems. Significant temperature fluctuations during the operation of diesel engines are often associated with malfunctions, including ignition failure, abnormal intake, and exhaust processes. Hence, the application of infrared thermography (IRT) for collecting infrared images of diesel engines and conducting quantitative analysis of the temperature distribution in these images has proven to be a faster and more efficient method for recognizing the health status of diesel engines, compared to other fault diagnosis methods. In recent years, there has been a growing interest in deep learning (DL) for fault diagnosis in various industries. This emerging trend has attracted significant attention from researchers. Convolutional neural network (CNN) has garnered significant attention owing to the exceptional capability in extracting image features. Therefore, the article presents a new fault diagnosis method for diesel engines using IRT and CNN. The proposed method involves conducting adaptive histogram equalization for image enhancement, followed by employing Softmax regression for pattern recognition. Finally, two sets of self-made experimental data are used to investigate the impact of temperature variations on fault diagnosis performance and to validate the efficacy of the proposed method in comparison with three DL methods. The findings indicate that this method exhibits superior performance in the realm of diesel engine fault diagnosis.

Keywords: infrared thermography, convolutional neural network, image enhancement, diesel engine, fault diagnosis

1 Introduction

Diesel engines, as a common type of reciprocating machine, serve as the primary power source for various equipment and find extensive application in agricultural production, petroleum equipment, defense technology, military industry, ship transportation, and other domains. Therefore, enhancing the dependability of diesel engines and guaranteeing the seamless functionality hold immense importance [1]. Diesel engines are susceptible to a range of faults during high-speed and heavy-duty operation due to the intricate and challenging working conditions, as well as certain inherent manufacturing defects. These faults have a significant impact on production efficiency and can even lead to major safety accidents [2]. Through the implementation of real-time condition monitoring and fault diagnosis techniques, the defects can be promptly identified and rectified. Additionally, appropriate maintenance measures can be implemented to effectively mitigate the risk of catastrophic fault and significant economic losses.

Currently, in the field of mechanical equipment fault diagnosis, sensor technologies commonly employed include vibration signals [3–5], sound signals [6], temperature signal analysis [7], as well as oil [8] and infrared image analysis [9–11]. The sensitivity and maintenance costs of the aforementioned methods [12] in the context of mechanical equipment fault diagnosis are illustrated in Figure 1.

Based on the findings presented in Figure 1, it is evident that the higher sensitivity of the sensor technology allows for the detection of abnormalities in mechanical

* **Corresponding author: Yuan Li**, Shijiazhuang Campus of Army Engineering University of PLA, Shijiazhuang, 050003, China, e-mail: hhliyuan123@163.com
* **Corresponding author: Xisheng Jia**, Shijiazhuang Campus of Army Engineering University of PLA, Shijiazhuang, 050003, China, e-mail: asd3v36@163.com
Rongcai Wang, **Hao Yan**, **Enzhi Dong**, **Zhonghua Cheng:** Shijiazhuang Campus of Army Engineering University of PLA, Shijiazhuang, 050003, China

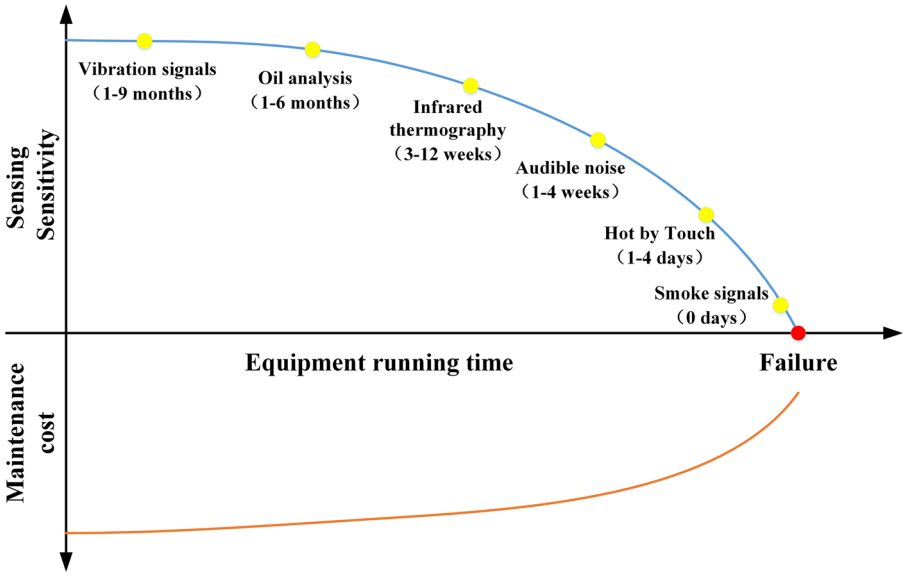

Figure 1: Sensitivity and maintenance costs of various sensor technologies.

equipment through signal analysis. However, accurately pinpointing and resolving faults is not currently feasible, leading to lower accuracy in fault diagnosis and subsequently higher maintenance costs. Conversely, as the sensitivity of the sensor technology decreases, the identification and characterization of faults become clearer, resulting in reduced maintenance costs. However, the decrease in sensitivity also increases the likelihood of equipment failure.

Among the fault diagnosis methods mentioned above, vibration signals are widely applied in fault diagnosis of diesel engines [13–15], primarily due to the ease of measurement and ability to capture crucial dynamic information of mechanical equipment, including the reciprocating motion of piston connecting rod components, crankshaft rotation, and valve opening and closing [16,17]. However, the analysis of vibration signals is subject to certain inherent limitations. The limitations include the need for contact measurement, the presence of complex noise pollution, and the requirement for local detection. Furthermore, the collection process of vibration signals typically takes place in challenging working environments, such as those with high temperatures or oil and gas-filled equipment surfaces. These conditions can easily result in signal measurement distortion and even damage to the sensors [18]. Additionally, the positioning of vibration signal sensors is typically conducted during equipment shutdown, resulting in certain losses associated with the shutdown process.

The structure and composition of diesel engines exhibit a high degree of complexity, while the working environment is characterized by harsh conditions. When faults arise, it is common to observe abnormal occurrences such as heightened friction, unusual vibrations, and fluctuations in temperature [19]. In recent years, temperature signals have garnered significant attention from scholars both domestically and internationally, owing to the abundant information regarding the condition of mechanical equipment [20–22]. In recent years, infrared thermography (IRT) has emerged as a promising non-contact and non-destructive testing technique for acquiring temperature data of devices [23]. The infrared thermal camera serves as the central element of the IRT system, enabling the remote monitoring of equipment surface temperatures. It is capable of converting temperature distributions into infrared images, which can be visually displayed. Compared to contact temperature measurement, IRT offers a simpler and more intuitive approach. By analyzing the abundant fault information present in infrared images, IRT enables a comprehensive assessment of the equipment's fault status. Currently, IRT has been extensively and effectively applied in the domains of equipment condition monitoring and fault diagnosis, as depicted in Figure 2. Furthermore, a comprehensive examination and synthesis of the fundamental principles, current application status, and procedural framework of IRT can be found in the literature [24].

This article aims to leverage the distinctive benefits of IRT to propose a novel approach for fault diagnosis of diesel engines. The proposed method is designed to achieve precise and efficient fault diagnosis. In this study, infrared images of diesel engines exhibiting various fault states were acquired using an infrared thermal camera. Subsequently, CNN was employed to automatically extract the distinguishing features from the infrared images. Finally, the

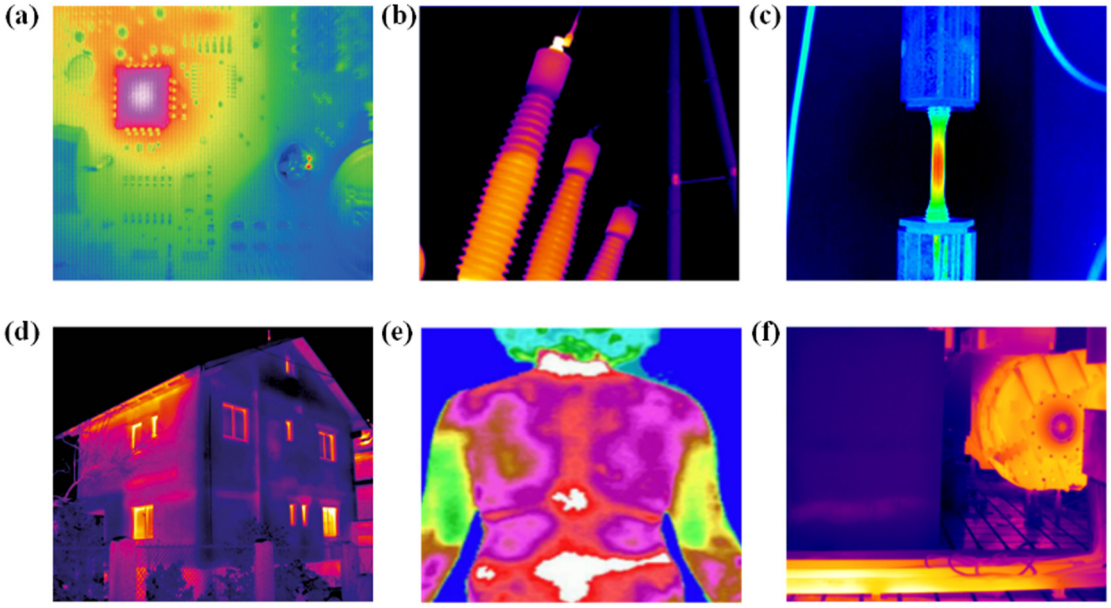

Figure 2: Related IRT application scenarios: (a) Defect detection of chips, (b) power facility troubleshooting, (c) evaluation of material extrusion fatigue, (d) inspection of building structure and internal facilities, (e) lumbar intervertebral disc examination in medical field, and (f) condition monitoring of the wind turbine gearbox.

utilization of a softmax regression (SR) classifier is deployed to effectively classify different fault types of the diesel engine. The efficacy of this method is validated through the analysis of experimental data obtained from a high-pressure common-rail diesel engine.

The subsequent sections of the article are structured as follows: Section 2 provides a comprehensive review of relevant literature, while also analyzing the current research state in the field. In Section 3, the article provides a concise introduction to the related background knowledge of the fault diagnosis method proposed, as well as an overview of the fault diagnosis method framework. Section 4 offers a comprehensive overview of the fundamental elements of the diesel engine system and the data acquisition system. Additionally, it examines the procedural approach to fault diagnosis. In Section 5, the article introduces two self-made datasets, as well as the parameter settings of the network models that are relevant to this study. Section 6 provides a discussion and comparison of the experimental results, thereby verifying the efficacy of the proposed method. Finally, conclusions and contributions of the article are summarized in Section 7.

2 Related literature review and their limitations

In recent years, experts and scholars from various countries have conducted numerous meaningful studies on the significant application value of IRT in the field of mechanical equipment fault diagnosis. This is due to the unique advantages of IRT in temperature measurement. Extracting image histogram features is considered a straightforward and efficient method for representing features. Younus *et al.* [25] introduced a novel approach for fault diagnosis of rotating machinery by using infrared image histogram features. Principal component analysis (PCA) and independent component analysis (ICA) were deployed during the feature extraction phase. Finally, support vector machine (SVM) was applied to classify the various fault states of rotating machinery. In a separate investigation conducted by Younus and Yang [26], the application of the two-dimensional discrete wavelet transform (DCT) was employed to decompose infrared images of rotating machinery. By extracting the histogram feature parameters of each layer following decomposition, the Mahalanobis distance and relief algorithm were applied for feature extraction. Subsequently, fault pattern recognition was performed using SVM and linear discriminant analysis (LDA). Widodo *et al.* [27] employed a self-organizing map (SOM) artificial neural network (ANN) to detect the faults in the rotor system and rolling bearings. The authors extracted histogram features from infrared images, such as mean, variance, standard deviation, central moment, entropy, kurtosis, maximum, and minimum, to identify misalignment, imbalance, looseness, and defects in the outer ring, inner ring, rolling body, and cage of rolling bearings. Bagavathiappan *et al.* [28] used IRT as a remote sensing technique to capture the

atypical temperature patterns on the surface of a machine, successfully implementing the method for real-time monitoring of the blower impeller end bearing, shaft, and motor in the exhaust system of a nuclear power plant, effectively preventing the occurrence of major malfunctions. Janssens *et al.* [29] introduced a novel automated system for fault diagnosis of bearings using IRT, with a specific focus on rotating machinery bearings as the subject of investigation. By utilizing various feature parameters, including temperature standard deviation, Gini coefficient, and optical moment, SVM was employed to classify bearing fault types that are challenging to diagnose with conventional techniques, such as the absence of lubricating oil. In a separate study [19], random forest (RF) was employed to conduct early fault diagnosis of rotating machinery by extracting the Gini coefficient from infrared images. Glowacz and Glowacz [11] proposed a novel method for extracting features from infrared images, which involves selecting the difference region of the image as the fault feature and classifying the fault patterns using K-means, K-nearest neighbor (KNN), and BP network. Li *et al.* [30] introduced a novel method for fault diagnosis of variable-speed rotating machinery using IRT. The feature parameters of infrared images were extracted based on the bag of visual word (BoVW), and the identification of various fault types of rotating machinery was achieved through the implementation of SVM.

Currently, within the realm of diesel engine fault diagnosis, the prevailing approaches primarily encompass the construction of dynamic models [31] and the implementation of condition monitoring techniques [32]. The construction of dynamic models primarily relies on the principles of mechanical equipment dynamics. By developing a dynamic model that accurately represents the mechanical properties of diesel engines, this study aims to investigate the correlation between variations in dynamic parameters and system malfunctions. Consequently, the ultimate goal is to achieve effective fault diagnosis for diesel engines. The limitations of this particular method are rooted in the challenges associated with constructing a model and its limited ability to generalize, thereby complicating the practical application process. To date, the diagnosis method based on condition monitoring has emerged as the most extensively employed approach in the particular field, which achieves real-time fault diagnosis of diesel engines through online monitoring and collection of related signals that can effectively characterize the fault state of diesel engines. The method involves several steps, including signal preprocessing, feature extraction, and pattern recognition, to accurately identify and diagnose faults in real-time. The commonly employed methods for condition monitoring encompass vibration analysis [33], temperature monitoring [34], noise analysis [18], abrasive particle detection [35], and instantaneous torque method [36].

With the rapid advancement and extensive application of artificial intelligence (AI) technology, fault diagnosis methods that rely on signal processing, feature extraction, and machine learning (ML) have garnered increasing attention and research [37,38]. During the operation of diesel engines, the measurement of vibration signals provides a convenient means to directly observe and analyze the dynamic variations of the equipment. Hence, vibration signal analysis has emerged as the predominant approach in the field of diesel engine fault diagnosis. Flett and Bone [39] introduced a novel approach for fault diagnosis of the diesel engine valve train. Aiming to address common issues in the valve train, such as valve spring deformation and abnormal clearance, the study employed the extraction of improved root mean square (RMS) to achieve accurate diagnosis results. Meanwhile, the naive Bayes classifier was applied for this purpose. Kowalski *et al.* [40] have proposed a novel intelligent fault diagnosis method for impulse diesel engines commonly used in marine and land transportation. The method is based on extreme learning machine (ELM) and demonstrates the ability to accurately classify 15 common fault states of diesel engines. A comparative analysis between SVM and KNN has been conducted to establish the superiority of the proposed method. Ramteke *et al.* [41] conducted a study where vibration and sound signals were collected to assess the wear status of diesel engines, employing fast Fourier transform (FFT) and short-time Fourier transform (STFT) for signal processing and extracting statistical feature parameters. Based on the features, ANN was used to effectively classify different types of wear faults.

Based on the above literature, it becomes evident that the fault diagnosis methods proposed in these studies lack significant accuracy. This can be attributed to the fact that infrared images are not subjected to enhancement techniques prior to feature extraction and fault classification. When acquiring infrared images, various factors such as circuit conditions, detector performance, space environment, and channel transmission errors can significantly impact the quality of the images. The issues often result in problems such as low contrast, blurred texture, and loss of details, which in turn adversely affect subsequent computer processing and recognition tasks. Image enhancement is a widely applied method in the preprocessing of infrared images. Its objective is to accentuate or refine certain features of the image, such as edges, contours, and contrast, with the aim of improving the visual impact of the image or making it more suitable for computer-based processing. Furthermore, it is important to

acknowledge that image enhancement does not fundamentally change the quantity of feature information conveyed by infrared images. Instead, it merely adjusts the dynamic range of feature information to a certain degree, facilitating the extraction of features and the identification of images by computer systems.

Image enhancement can be broadly categorized: including spatial domain and frequency domain, such as histogram equalization [42–44], grayscale transformation [45], spatial domain filtering [46,47], and frequency domain filtering [48,49]. The precise categorization is shown in Figure 3.

Currently, histogram equalization (HE) is widely employed as a prevalent technique in the field of image enhancement. The gray-scale histogram of an image is a representation of the distribution of gray-scale values, which provides an intuitive visualization of the frequency of each gray-scale level across all pixels in the image, effectively describing the overall gray-scale distribution. The primary objective of HE is to employ the cumulative distribution function to convert the grayscale histogram of the original image from a relatively concentrated grayscale interval to a uniform distribution across all grayscale ranges. By performing nonlinear stretching on the image, the pixel values are redefined to achieve a more balanced distribution of pixels within a specific grayscale range, as shown in Figure 4.

HE effectively enhances the contrast of the image [50]. However, the application of HE can significantly change the contrast of the entire image, potentially resulting in the loss of local detail information [51]. In recent years, the adaptive histogram equalization (AHE) has garnered significant attention and research interest [52,53]. AHE performs equalization according to the local regions of the image, avoiding the problem of excessive enhancement or distortion caused by global equalization, preserving image details, improving the effectiveness of image enhancement, and making it more conducive to the next step of feature extraction.

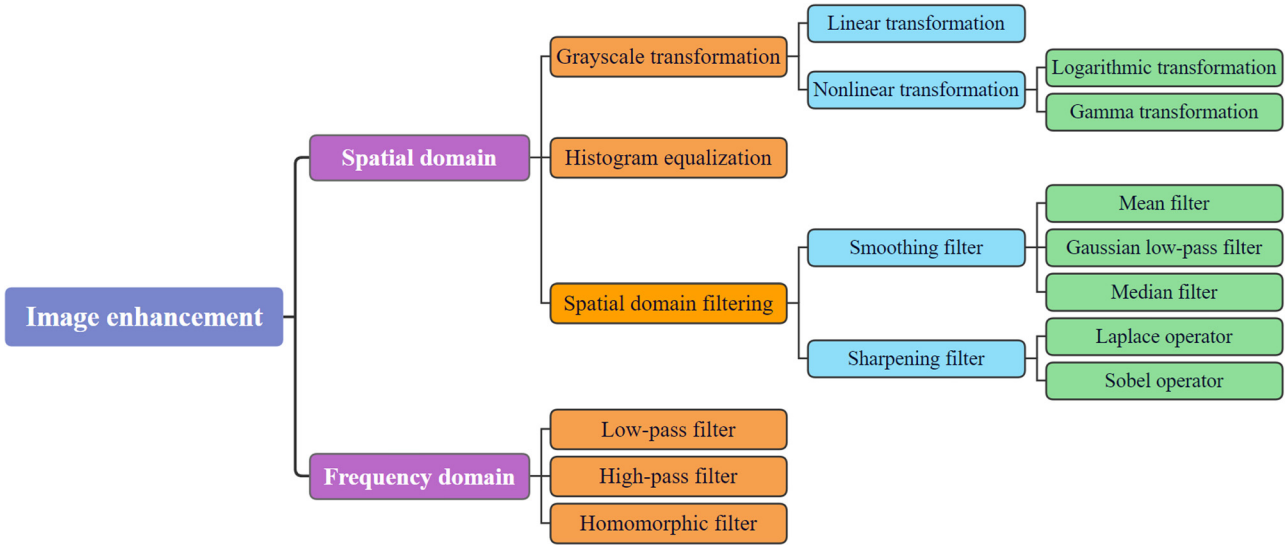

Figure 3: Classification of frequently employed techniques for image enhancement.

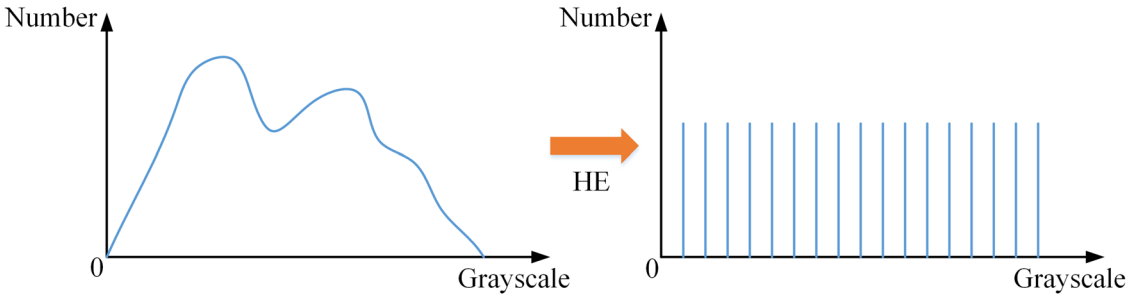

Figure 4: Schematic graph of histogram equalization.

Despite some progress in the field of diesel engine fault diagnosis, there are still some limitations associated with the application of conventional ML methods, including the inadequate capacity to extract intricate features from raw data and the inability to accurately identify interference information within the data. Up to now, research studies on fault diagnosis of mechanical equipment using IRT are still limited, and it is characterized by four main problems:

1) The majority of existing research on fault diagnosis of diesel engines primarily concentrates on the analysis of a single fault mode. However, in the practical operation of diesel engines, the state often involves a complex fault mode where multiple faults occur simultaneously. Therefore, it is not practical to study diesel engine fault diagnosis only in single fault mode.
2) Most of the existing fault diagnosis methods based on IRT need to manually extract the features of infrared images. Because infrared images contain rich and complex status information, the process of feature selection needs a lot of professional field prior knowledge and engineering experiments, which greatly affects the efficiency of fault diagnosis and the final classification results.
3) The application of IRT in the domain of mechanical equipment fault diagnosis is currently constrained to conventional rotating machinery, such as rolling bearings, rotor systems, and shafts. However, there exists a dearth of both theoretical and applied investigations pertaining to reciprocating mechanical equipment, such as diesel engines.
4) In practical industrial applications, faults may manifest under varying temperature conditions. When there is a variation in the temperature distribution of mechanical equipment, the corresponding distribution of fault information on infrared images also changes. This, in turn, poses a challenge in extracting features that remain unaffected by temperature fluctuations. However, the existing fault diagnosis methods that rely on IRT fail to take into account the impact of temperature fluctuations on the outcomes of fault diagnosis.

To address the research gap, this article concentrates on typical fault modes of diesel engines. It employs an infrared thermal camera to capture infrared images under different fault modes and subsequently applies deep learning (DL) methods for extracting fault characteristics from the infrared images at different temperatures. The primary advantage of DL resides in its inherent capability to automatically extract features from images, thereby overcoming the limitations associated with manual feature extraction methods. Additionally, DL exhibits strong image-processing capabilities [54]. In the context of mechanical equipment fault diagnosis, there exist four commonly used DL models: deep neural network (DNN) [55], recurrent neural network (RNN) [56,57], stacked automatic encoder (SAE) [58], and convolutional neural network (CNN) [59]. In the aforementioned models, CNN is capable of automatically and efficiently extracting deep-level features from images, which is achieved through key operations such as convolution and pooling [60], which help to mitigate information loss that may occur during manual processing. In the meantime, CNN employs the approach of local connection and weight sharing, which serves two purposes: first, it reduces the number of weights, thereby facilitating network training and optimization [61]. Secondly, it decreases the model's complexity, consequently mitigating the risk of overfitting. CNN exhibits significant advantages in the analysis of two-dimensional images, as it possesses the capability to extract various image features, encompassing color, texture, shape, and image topology. The proposed method exhibits strong robustness and computational efficiency when applied to image processing, particularly in tasks involving the identification of displacement, scaling, and other types of distortion invariance. Therefore, this article employs CNN to extract features from infrared images across various fault states.

3 Methodology

3.1 Adaptive histogram equalization

AHE is a sophisticated image processing technique to enhance the contrast of images. Unlike global HE, which aims to equalize the entire image, AHE operates by equalizing the local image based on the grayscale distribution within specific regions of the image, achieving a more effective enhancement effect. Therefore, AHE is considered more appropriate for enhancing the local contrast of images and capturing finer image details. The fundamental concept of AHE involves the partitioning of an image into multiple regions, with the application of HE to each local region, which mitigates the issue of excessive enhancement resulting from global HE, while simultaneously preserving the local characteristics of images. The following is the specific application process of AHE.

1) The image should be divided into multiple not overlapping local regions. The dimensions of each region can be modified based on the specific attributes of the application scenario and the image. Typically, the

dimensions of each region are 8 × 8, 16 × 16 or 32 × 32 pixels. As AHE operates in local regions, each local region can be processed independently. The presence of overlapping regions aids in preserving the image's continuity and smoothness, while also preventing the formation of conspicuous boundaries.

2) The grayscale histogram needs to be calculated for each region. To obtain the grayscale histogram for each region, it is necessary to calculate it separately for each region. The calculation method for the grayscale histogram can be represented by the following equation:

$$h(k) = \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} w(i,j)\delta(k - f(i,j)), \tag{1}$$

where $h(k)$ denotes the count of pixels with a grayscale level of k, $w(i, j)$ represents the weight assigned to pixel (i, j), $f(i, j)$ represents the grayscale level of pixel (i, j), and $\delta(x)$ is the Kronecker function. Specifically, when x equals 0, $\delta(x)$ is equal to 1; otherwise, its value is 0. N and M correspond to the width and height of the image, respectively.

3) The cumulative distribution function needs to be calculated for each region. For the grayscale histogram of each region, the cumulative distribution function can be calculated using the following equation:

$$c(k) = \sum_{i=0}^{k} h(i). \tag{2}$$

In Eq. (2), the variable $c(k)$ denotes the cumulative sum of the quantity of pixels having a grayscale level equal to or less than k.

4) Calculate the mapping function for each respective region. The calculation of the mapping function for each region's cumulative distribution function is expressed as follows:

$$s(k) = \frac{L-1}{NM} \sum_{i=0}^{k} h(i). \tag{3}$$

Among the variables used in the context, $s(k)$ denotes the grayscale level of a pixel after mapping, where k represents the original grayscale level. L represents the total number of grayscale levels present in the image.

5) HE is conducted for each region. For each region, the pixels contained within it are assigned a grayscale value through the utilization of a mapping function.
6) The processed regions should be merged into a single image.

The size of the local region in AHE is set to 8 × 8 in the article.

3.2 Convolutional neural network

CNN is composed of a series of convolution layers, activation layers, batch normalization layers, pooling layers, and fully connected layers. It can automatically extract features from the input data and apply the features to perform tasks such as classification, recognition, and prediction. It has achieved good results in image recognition, speech recognition, natural language processing, and other fields. In the CNN model, the convolution layer extracts features through a sliding filter (a set of convolution kernels), the pooling layer reduces the dimension of features through down-sampling, and the fully connected layer maps features to specific output classes. The basic process of CNN is shown in Figure 5, and the following sections will provide a detailed description of the different layers of CNN.

3.2.1 Convolutional layer

Convolution operation is a fundamental operation in CNN, wherein a set of convolutional kernels is employed to

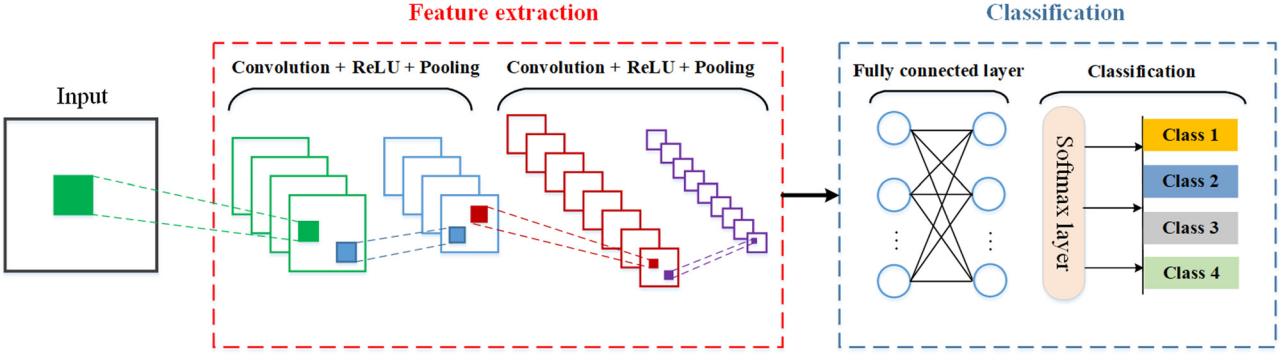

Figure 5: Typical architecture of CNN.

conduct convolution calculations on input data, resulting in the generation of feature maps. Each new feature map is obtained by convolving the local features of the input data, a process referred to as "local perception." During the process of obtaining the feature map through convolution operation, the convolution kernel remains constant in order to achieve "weight sharing." The mathematical representation of the convolution operation is as follows:

$$\mathbf{x}_i^n = \mathbf{c}_n * \mathbf{x}_i{}^{n-1} + \mathbf{b}_i^n, \tag{4}$$

where the variable $\mathbf{x}_i^n$ denotes the ith output map of the nth layer, $\mathbf{x}_i^{n-1}$ represents the ith output map of the (n−1)th layer, and $\mathbf{c}_n$ signifies the convolution kernel of the nth layer. The symbol "*" denotes the convolution operation, while $\mathbf{b}_i^n$ represents the bias vector.

3.2.2 Batch normalization layer

The inclusion of a batch normalization layer in CNN is intended to enhance both the training speed and overall performance of the network. By applying input normalization to each layer, the stability of the input distribution is enhanced, thereby mitigating the issues of gradient vanishing and exploding. Consequently, the approach facilitates the training process by accelerating its convergence. The process of batch normalization can be represented as follows [22]:

$$\mathbf{x}^{n-1} = \frac{\mathbf{x}^{n-1} - \mu_P}{\sqrt{\sigma_P^2 + \varepsilon}}, \tag{5}$$

$$\mathbf{x}^n = \gamma^n \mathbf{x}^{n-1} + \beta^n, \tag{6}$$

where $\mu_P = E[\mathbf{x}^{n-1}]$, $\sigma_P = \mathrm{Var}[\mathbf{x}^{n-1}]$. The constant ε is introduced to avoid division by zero in the denominator. On the other hand, the coefficients γ and β are used to scale and shift the normalized data, respectively. These coefficients need to be learned in order to properly adjust the data.

3.2.3 Activation layer

The activation layer is typically positioned following the convolutional layer. By incorporating nonlinear mapping functions such as Sigmoid, Tanh, ReLU, *etc*., the CNN's expression and generalization capabilities can be enhanced, allowing it to effectively address intricate nonlinear problems. Compared to activation functions such as Tanh and Sigmoid, the ReLU offers several advantages. Firstly, ReLU effectively mitigates the issue of gradient disappearance, ensuring more stable training of neural networks. Additionally, ReLU exhibits faster calculation speed and better convergence, further enhancing its appeal as an activation function. The mathematical representation of the output of ReLU can be expressed using the following equation:

$$f(x) = \begin{cases} x, & x \geq 0 \\ 0, & x < 0 \end{cases}, \tag{7}$$

where let x represent the input value and $f(x)$ denote the corresponding output value. When the input value is greater than or equal to zero, the output value is equal to the input value. Conversely, when the input value is less than zero, the output value is zero.

3.2.4 Pooling layer

The pooling layer serves as a crucial component in the down-sampling process, aiming to decrease the dimensions of images or feature maps while preserving the essential characteristics. This reduction in size not only helps to minimize computational expenses but also mitigates the risk of overfitting. The pooling operations commonly employed in various applications encompass maximal pooling and average pooling. The encoding of positional information in features makes the maximal value of the feature map more informative in terms of information content, as opposed to the average value. Therefore, when conducting pooling operations, it is common to use maximal pooling, which can be expressed as follows:

$$\mathbf{x}^n = \max_{(i-1)W+1 \leq j \leq iW} \{\mathbf{x}^{n-1}(j)\}, \tag{8}$$

where W represents the width of the maximal pooling area.

3.2.5 Fully connected layer

The fully connected layer is commonly deployed as either the output layer or an intermediate layer within a neural network to facilitate dimensional transformation. It is employed for the purpose of classifying or regressing input features. In the CNN model, the outputs of the convolutional layer, pooling layer, and activation function ReLU are typically linked *via* one or more fully connected layers in order to generate G-dimensional vectors for the purpose of pattern recognition. Here, G represents the total number of classes. The fully connected layer is characterized by a high number of parameters, which can lead to overfitting. Hence, it is common practice to employ regularization techniques, such as L1 and L2 regularizations, in conjunction to minimize the number of parameters in the fully connected layer.

3.3 SR

SR is a model used for multi-classification tasks, where it addresses the classification problem by mapping the input vector to a probability vector. The application of this approach offers several benefits, including rapid training speed and robust interpretability of output results. Consequently, it has found extensive application in various domains such as image classification, natural language processing, recommendation systems, face recognition, *etc.* [62]. Meanwhile, in the context of multi-classification problems, SR serves as a classifier that is frequently integrated with other models, such as CNN and RNN, as the output layer in DL. The primary function of SR is to determine the fault state by computing the probability that the samples to be classified are associated with each fault class label. The particular model is depicted in Figure 6.

The subsequent section outlines the procedural steps and mathematical derivation of the SR theory.

1) *Data Input*. Assuming a dataset with m samples, each containing n features, the input data are represented by $\boldsymbol{X}$. The dimension of $\boldsymbol{X}$ is (m, n).
2) *Provide explicit values for the weights and biases*. The weight matrix, denoted as $\boldsymbol{W}$, is employed in the context to represent the dimensions of (n, k), where k signifies the number of classes. Additionally, the bias vector, denoted as $\boldsymbol{b}$, possesses a dimension of k.
3) *Constructing a linear model*. For every given sample $\boldsymbol{X}$, the score under each class is calculated as $\boldsymbol{Z} = \boldsymbol{WX} + \boldsymbol{b}$. The dimension of $\boldsymbol{Z}$ is represented as (m, k).
4) *The Softmax function is substituted*. To transform the scores of each sample within each class into probability values, the softmax function is applied, denoted as y = softmax ($\boldsymbol{Z}$). The Softmax function is defined by the following equation:

$$\text{softmax}(\mathbf{Z}_i) = \frac{e^{\mathbf{z}_i}}{\sum_{j=1}^{k} e^{\mathbf{z}_j}}, \tag{9}$$

where i denotes the score of the current sample belonging to the ith class, and j represents all classes.

5) The introduction of the loss function is discussed. Cross entropy is commonly employed as the loss function, and it can be mathematically expressed as follows:

$$L = -\frac{1}{m}\sum_{i}^{m}\sum_{j}^{k} y_{ij} \log \hat{y}_{ij}, \tag{10}$$

where the variable y represents the true label, while the variable $\hat{y}$ represents the predicted probability value.

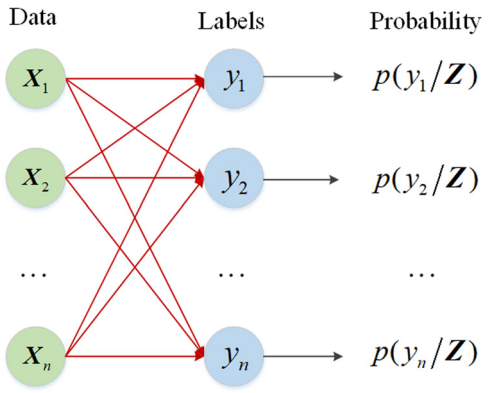

Figure 6: SR model.

6) *Update parameters*. The gradient of the loss function with respect to the weight and offset is calculated, and the parameters are updated using the gradient descent method. The process can be expressed as follows:

$$\Delta\mathbf{W} = -\alpha\frac{\partial L}{\partial \mathbf{W}}, \Delta\mathbf{b} = -\alpha\frac{\partial L}{\partial \mathbf{b}}. \tag{11}$$

In Eq. (11), the symbol α denotes the learning rate, which serves the purpose of regulating the magnitude of each iteration's step in the gradient descent algorithm.

7) *Pattern recognition*. For the given input sample, the probability value for each class should be calculated, and the class with the highest probability value should be selected as the predicted result.

4 The proposed system for fault diagnosis of diesel engines

The diesel engine fault diagnosis system proposed in the article is depicted in Figure 7. The system comprises four modules, namely, infrared image acquisition, image preprocessing, CNN-based adaptive feature extraction, and SR classifier. First, it is necessary to gather infrared images of the diesel engine in various conditions, including normal condition (NC), single cylinder misfire (SCM), multi-cylinder misfire (MCM), and air filter blockage (AFB). Before feeding the infrared images into the feature extraction module, it is necessary to preprocess them, which involves several steps, such as grayscale processing, capturing the region of interest (ROI) in the images, and applying AHE algorithm to enhance the contrast of the ROI. Finally, the image feature parameters extracted by CNN are divided into

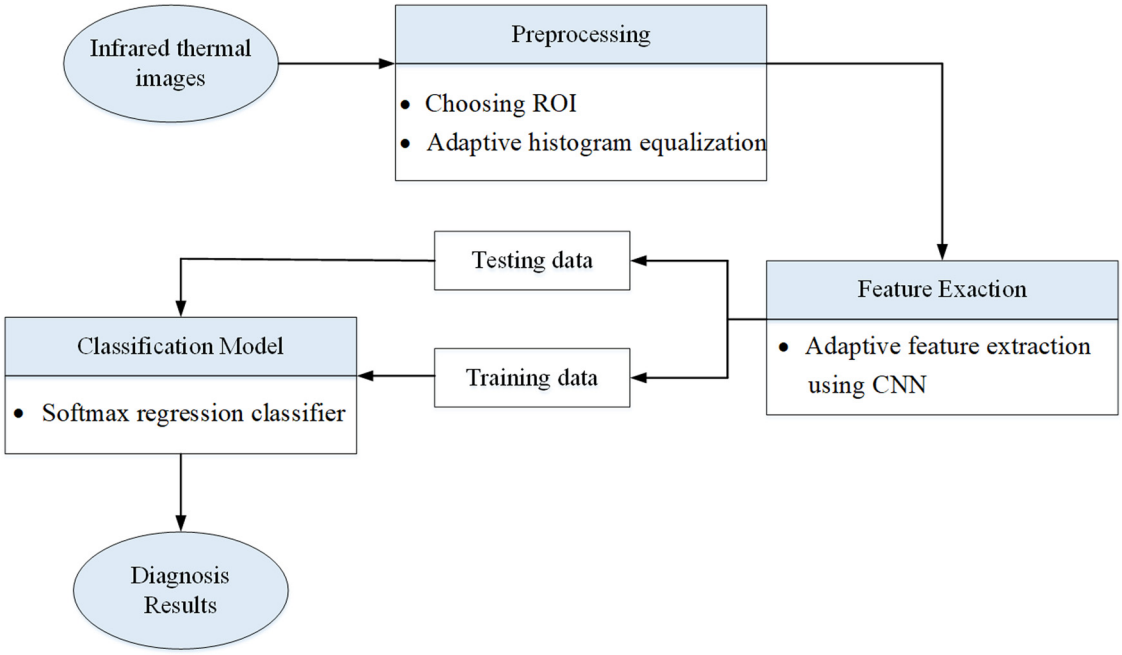

Figure 7: Fault diagnosis system of the diesel engine.

training and testing sets. These sets are then directly inputted into the fault pattern recognition process of the SR classifier.

4.1 The proposed IRT-CNN-based fault diagnosis method

This section aims to present the proposed method for fault diagnosis of diesel engines, which is based on IRT and CNN. CNN possesses two primary advantages in the extraction of image feature parameters. First, CNN has the capability to autonomously extract deep features from images without relying on prior knowledge or human subjective experience in feature design and extraction. This ability reduces the influence of human involvement in the process. Secondly, in contrast to traditional fully connected networks, CNN employs techniques such as local connection, weight sharing, and spatial pooling. On one hand, the reduction in the number of weights simplifies the training and optimization process of the network. On the other hand, it also decreases the complexity of the model, thereby mitigating the risk of overfitting and enhancing the extraction of translation invariant features. The flow chart illustrating the IRT-CNN-based fault diagnosis method for diesel engines, as proposed in the article, is presented in Figure 8.

As shown in Figure 7, the methodology comprises three primary stages. In the initial phase, infrared image data is gathered for the diesel engine across various fault states. In the subsequent phase, CNN is employed to develop a fault diagnosis model for diesel engines. The model aims to extract fault characteristics from infrared images. Finally, the fault features collected under various fault states are fed into the SR classifier to perform fault pattern recognition of the diesel engine. It is noteworthy that the application of IRT-CNN based fault diagnosis method, as proposed in the article, does not necessitate any prior understanding of the fault mechanism or parameter configurations of the diesel engine. This method exhibits robust self-learning capabilities and is easily applicable, making it highly adaptable for fault diagnosis of diesel engines.

4.2 Infrared image acquisition

When analyzing the infrared image signals, the variation in temperature of the diesel engine cylinder head serves as an indicator not only of the combustion conditions within the cylinder, but also provides valuable insights into the operational status of the ignition system, fuel supply system, and the reciprocating impact of the piston. Therefore, collecting infrared images of the diesel engine cylinder head for the purpose of fault pattern recognition holds significant relevance and practical importance. In this study, infrared image data acquisition is dependent on the utilization of a high-pressure common-rail diesel engine test bench. Figure 9

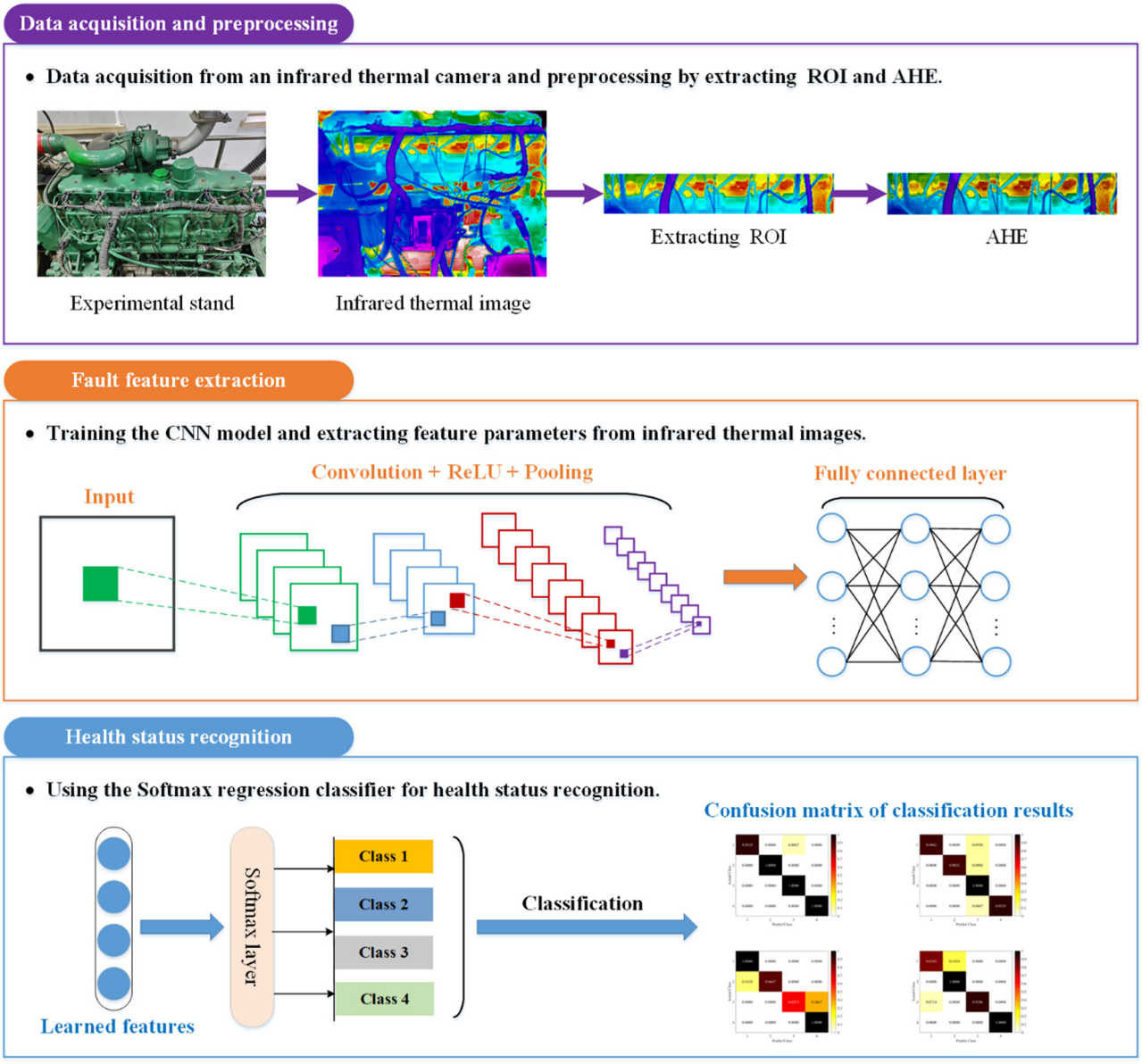

Figure 8: Procedure of the proposed method.

illustrates the fundamental structure and layout of the diesel engine condition monitoring system. This system primarily consists of a diesel engine and its control panel, an infrared thermal camera, and a data acquisition system. The diesel engine type employed in the experiment is CA6DF3-20E3, and the corresponding operational parameters are presented in Table 1.

As shown in Figure 9(a), the high-pressure common-rail diesel engine test bench comprises two main components, namely the diesel engine and the control panel. The control panel is responsible for managing the initiation, ignition, and shutdown of the diesel engine, while the accelerator pedal is used for executing acceleration and deceleration maneuvers. The control panel's dashboard enables real-time monitoring of various parameters such as speed, intake pressure, fuel rail pressure, and water temperature of the diesel engine. This allows for the detection of any sudden abnormalities or irregularities.

In this experimental study, the MAG32 infrared thermal camera was applied for the purpose of capturing infrared images, as depicted in Figure 9. The pertinent parameters of the infrared thermal camera are presented in Table 2.

The temperature measurement process of an infrared thermal camera can be influenced by various factors, including environmental temperature, measurement distance, and the emissivity of the object's surface. Infrared

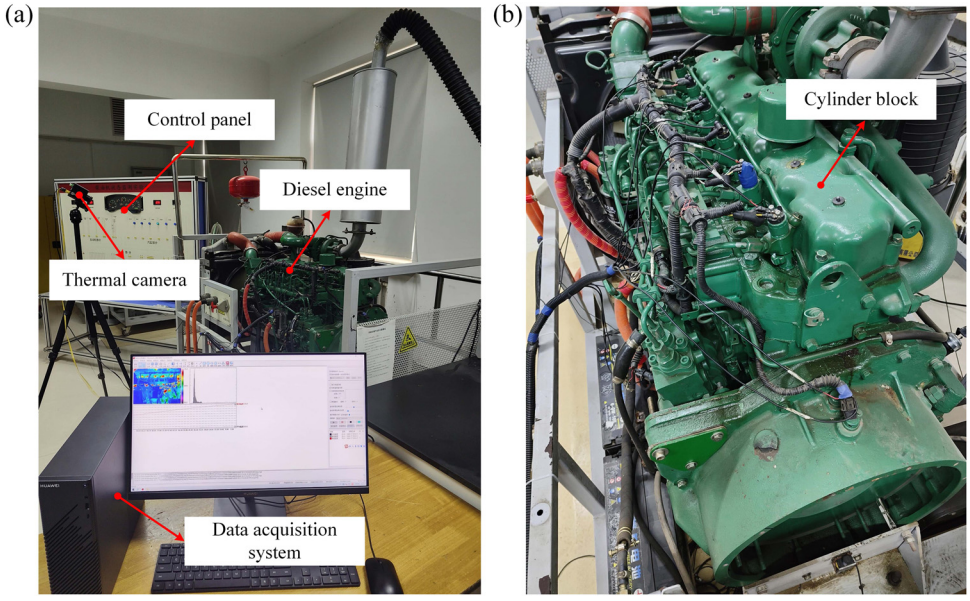

Figure 9: (a) IRT-based diesel engine condition monitoring system; (b) diesel engine test bench.

Table 1: Working parameters of the diesel engine

Items	Values
Common-rail mode	BOSCH electronic control common-rail system
Air intake mode	Supercharged intercooling
Rated power	155 kW
Rated speed	2300 rpm
Maximal torque	760 N m
Compression ratio	17.4
Total displacement	6.7 L
Number of cylinders	6
Number of valves per cylinder	2

Table 2: Parameters configuration of the infrared thermal camera

Configuration parameters	Values
Equipment type	MAG32, China
Infrared detector type	Uncooled focal plane
Image resolution	384 × 288
Frame rate	50 fps
Measuring range	–20°C–150°C
Environmental temperature	20°C
Measuring sensitivity	0.02°C
Emissivity	0.94
Test distance	1.2 m
Palette	High contrast

thermal cameras can be used for temperature correction to compensate for the above factors. The surface emissivity of an object is defined as the ratio of the radiant energy emitted by the object at a specific temperature T to the radiant energy emitted by a blackbody at the same temperature. The characterization of the thermal radiation characteristics of an object's surface and the calculation of thermal radiation energy transfer are crucial parameters. The object's thermal radiation and thermal conduction characteristics are of great significance for engineering design and research, which are influenced by various factors such as temperature, surface material, and spectral wavelength. The emissivity of the object surface ranges from 0 to 1, and any variations in emissivity within a specific wavelength and temperature range can be disregarded. The emissivity values of various common materials within a specific temperature range are presented in Table 3.

In the present study, the application of green paint on the surface of the engine cylinder block is investigated. With regard to the data presented in Table 3, the emissivity value is determined to be 0.94. During the experiment, four common fault states were simulated for the diesel engine, namely NC, SCM, MCM, and AFB. Among them, simulating single cylinder misfire and multi-cylinder misfire by disconnecting the cylinder ignition power cord, and installing an intake hood to simulate AFB. The malfunction of the diesel engine is shown in Figure 10.

Table 3: Emissivity of common materials in a certain temperature range

Materials and status	Temperature (°C)	Emissivity
Blackbody	All temperatures	1.00
Polished (oxidized) aluminum	50–500 (200–600)	0.04–0.06 (0.11–0.19)
Polished copper	115	0.023
Matte brass plate	50~350	0.22
Polished steel	100	0.07
Newly rolled steel	20	0.24
Polished iron	425–1025	0.14–0.38
Rusty iron plate	19	0.69
Water	0–100	0.95–0.96
Human skin	36	0.98
Coal (carbon)	25–30	0.93
Various colors of oil paint	Below 100°C	0.92–0.96

Set the engine speed to 500 revolutions per minute (rpm) and use the infrared thermal camera to acquire infrared images in the following manner.

Step 1. Choose one fault class from the four simulated fault states.

Step 2. Set the environmental parameters of the infrared thermal camera. Before commencing the experiment, it is observed that the temperatures of the diesel engine test bench and the surrounding environment are initially similar, which can be attributed to the gradual reduction in temperature difference between the diesel engine and the environment as a result of continuous heat transfer. Eventually, it is possible for the temperature to reach equilibrium. At present, the laboratory thermometer indicates a temperature of 20°C. The minimum temperature of the first infrared image, captured by the infrared thermal camera, is extracted for each of the four fault states, as depicted in Figure 11.

As shown in Figure 10, the minimum temperatures observed on the surface of the diesel engine cylinder remain consistently close to 20°C across all four fault states. In the experiment, the environmental temperature parameter of the infrared thermal camera is set to 20°C.

Step 3. At the constant speed of a diesel engine of 500 rpm, infrared images are collected every 10 s.

Step 4. The duration for image acquisition is set to 25 min in order to allow the highest temperature on the surface of the cylinder to reach a stable state. In the experiment, the temperature stabilizes at approximately 52°C. A total of 150 infrared images are collected for each fault class. The maximum temperature profile of the diesel engine cylinder surface the temperature is displayed in Figure 12.

Step 5. Cool the diesel engine to the environmental temperature and repeat the aforementioned steps for the remaining fault classes until the data collection process is concluded.

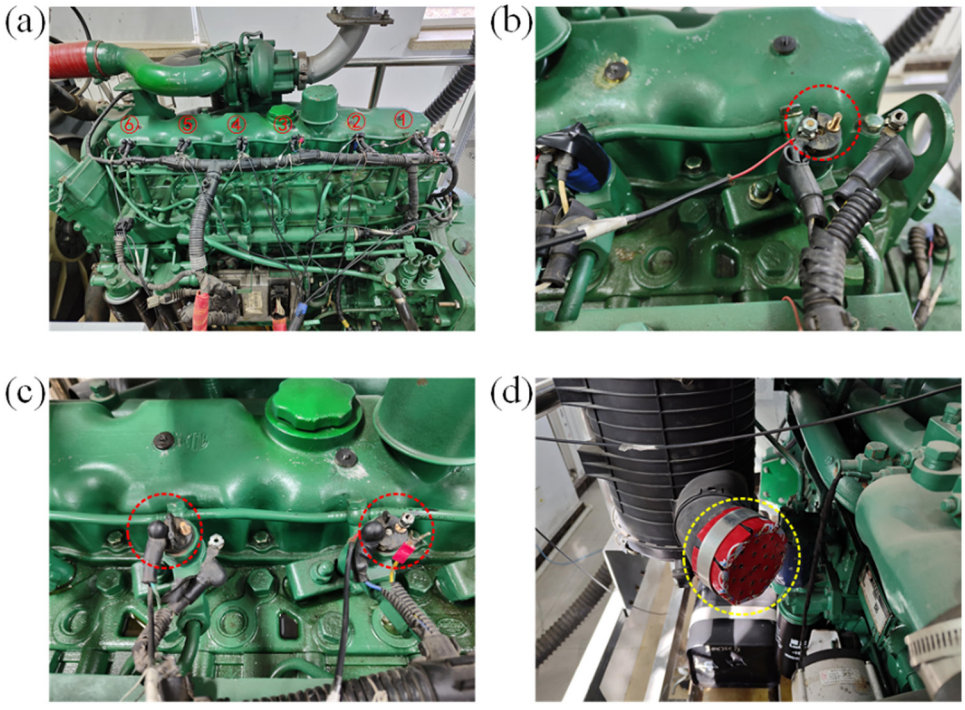

Figure 10: (a) NC, ①–⑥ are the six cylinders, respectively, (b) SCM, (c) MCM, (d) AFB.

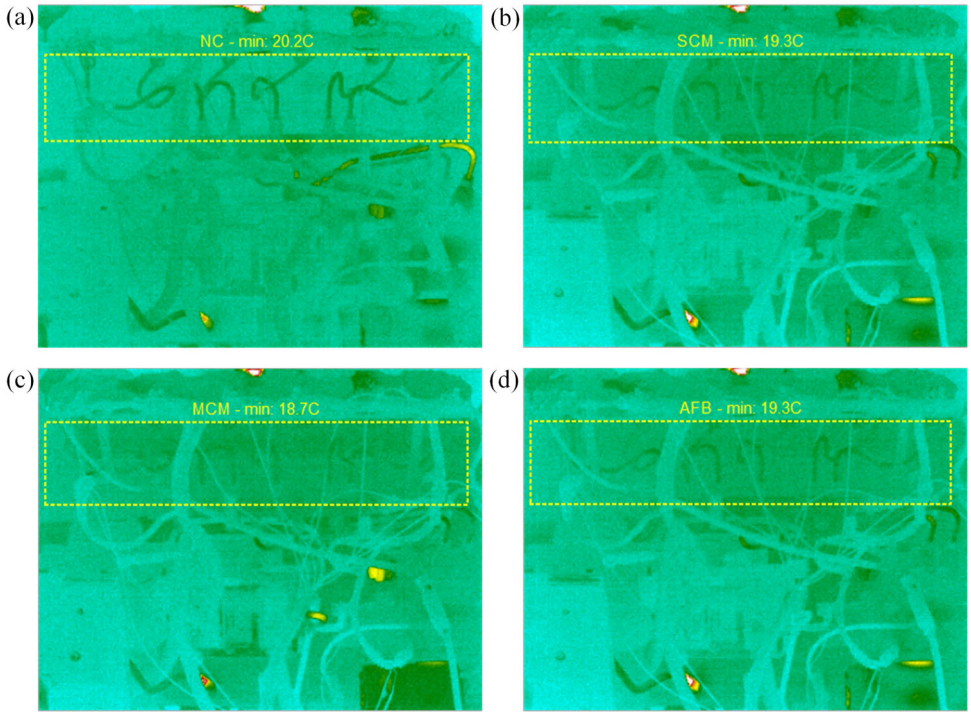

Figure 11: The minimum temperature profile of the diesel engine cylinder surface: (a) NC, (b) SCM, (c) MCM, and (d) AFB.

4.3 The process of infrared image enhancement

Before implementing image enhancement for infrared images, it is necessary to first extract the ROI from infrared images. This initial step serves to simplify the complexity of subsequent image processing tasks and effectively reduce the overall amount of image data. The research should primarily concentrate on the thermal dynamics of the diesel engine cylinder block, which experiences substantial variations in temperature. At the same time, it can also emphasize the primary information contained within the image, thereby facilitating subsequent image processing and analysis. The ROI plays a crucial role in an image as it encompasses significant information that requires careful consideration. Selecting the ROI can enhance the ability to capture pertinent information, optimize the speed and efficiency of image processing, and facilitate improved image recognition and classification by algorithms. Infrared images depicting the fault state of the diesel engine running steadily for a duration of 10 min are selected, as illustrated in Figure 13. The temperature unit is °C.

As shown in Figure 13(a), the infrared image reveals notable brightness and high temperature in the diesel engine cylinder region. Additionally, the brightness near the turbocharger is also prominent, indicating the highest temperature point in the entire image. In Figures 13(b) and 13(c), SCM and MCM occurred in the diesel engine, resulting in a decrease in the temperature of the corresponding cylinder block region. In Figure 13(d), it can be observed that the temperature in the vicinity of the turbocharger is significantly higher compared to other conditions. AFB can lead to suboptimal engine intake, which in turn can result in a dense fuel-air mixture, incomplete combustion, and inadequate airflow for turbocharger cooling. In addition, it should be noted that certain extraneous variables, such as the presence of power lines and experimental equipment, can introduce substantial interference to infrared images, thereby affecting the suitability for subsequent analysis. Therefore, the ROI extraction from infrared images is advantageous for subsequent feature extraction studies and serves to mitigate the impact of interfering factors. The resolution of infrared images obtained in the study is 384 × 288. After isolating the engine block region, the resolution is reduced to 355 × 49, as depicted in Figure 14.

After the ROI extraction from infrared images, the size of the images decreased significantly from 324 kb to 51.1 kb. The size reduction suggests that the ROI extraction can effectively reduce the amount of data that needs to be processed, leading to enhanced speed and efficiency in image processing. The extracted ROI from infrared images is inputted into the AHE-based image enhancement module, as shown in Figure 15.

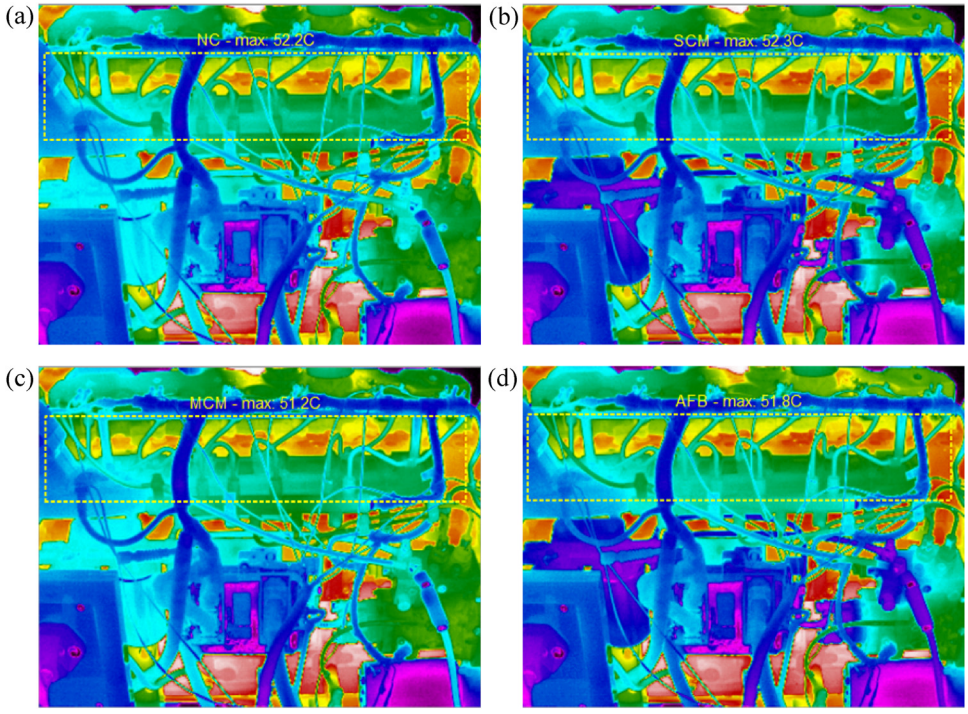

Figure 12: The maximal temperature profile of the diesel engine cylinder surface: (a) NC, (b) SCM, (c) MCM, and (d) AFB.

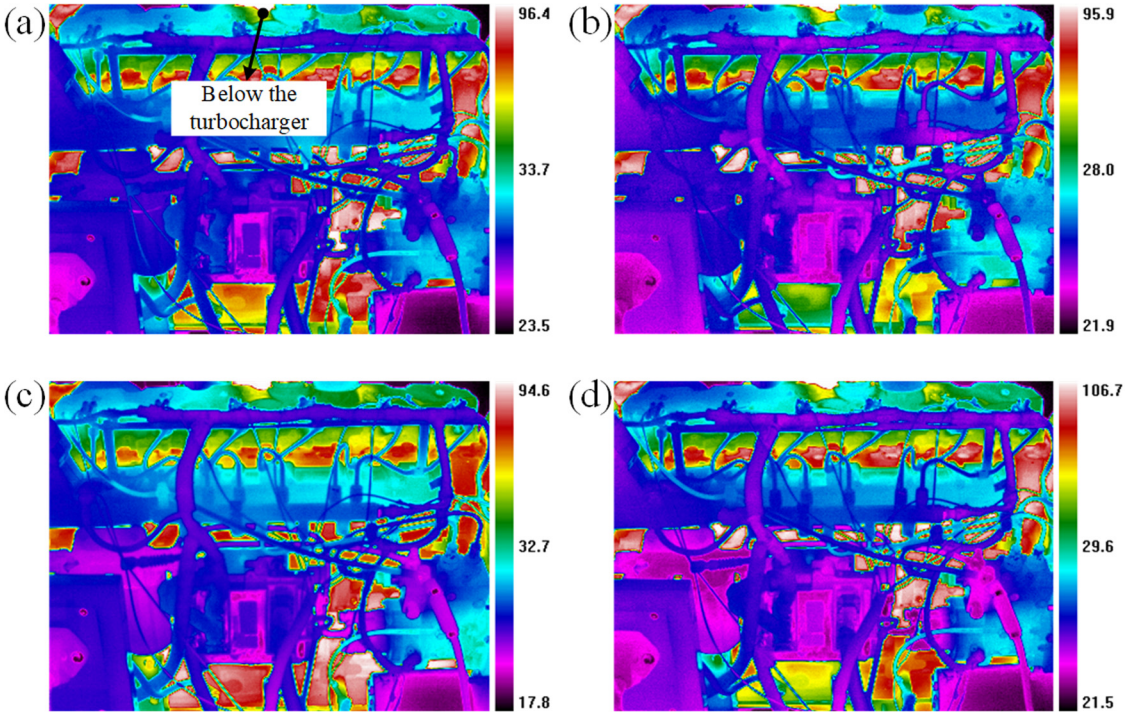

Figure 13: Infrared images of the diesel engine: (a) NC, (b) SCM, (c) MCM, (d) AFB.

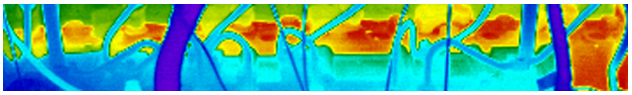

Figure 14: ROI of a diesel engine infrared image under NC.

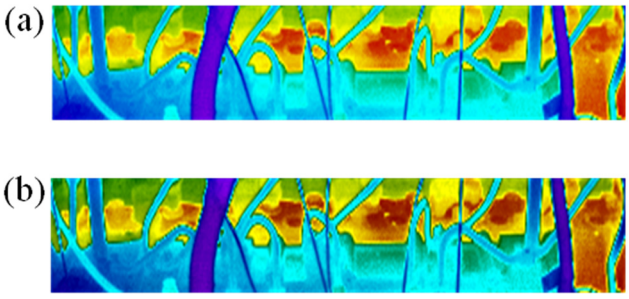

Figure 15: (a) ROI of infrared images under NC; (b) ROI of infrared images under NC after image enhancement.

The analysis of Figure 15 reveals that AHE has a notable impact on enhancing the visual quality of images. It effectively improves the color distribution and enhances the effect on feature extraction by adapting to the grayscale distribution in various regions.

5 Experimental study

5.1 Experiment dataset description

During the operation of the diesel engine, the surface temperature of the cylinder block experiences a certain degree of fluctuation. However, over time, it stabilizes at a temperature of approximately 52°C, after initially rising from 20°C. We collected 150 infrared images for each fault class. The article employs four commonly used DL models, namely CNN, DNN, RNN, and SAE, to analyze infrared images and identify fault modes of diesel engines.

In the present experimental study, two self-made experimental datasets, namely dataset 1 and dataset 2, after ROI extraction and AHE-based image enhancement, are employed to validate the efficacy and precision of the proposed method. Among the datasets, dataset 1 comprises 600 infrared image samples, with 150 samples collected from each of the four fault classes. On the other hand, dataset 2 is a sample grid that encompasses five time regions: region 1 (0–5 min), region 2 (5–10 min), region 3 (10–15 min), region 4 (15–20 min), and region 5 (20–25 min). Each time region consists of 120 samples (four fault classes × 30 samples). For instance, Figure 16 displays the infrared images depicting four different fault states of the diesel engine in region 2.

As reflected in Figure 16, it is difficult to visually distinguish the fault class when analyzing infrared images under different fault states. Due to the phenomenon of thermal conduction, the variations observed in infrared images of different fault classes are minimal. In the present study, a partitioning scheme was employed where 80% of the available data samples were allocated for training purposes, while the remaining 20% were reserved for evaluating the model's generalization ability and real-world performance. Tables 4 and 5 provide detailed descriptions of individual regions of dataset 1 and dataset 2, respectively.

Table 4: Detailed description of dataset 1

Fault class	Class label	Number of training set images	Number of test set images
NC	1	120	30
SCM	2	120	30
MCM	3	120	30
AFB	4	120	30

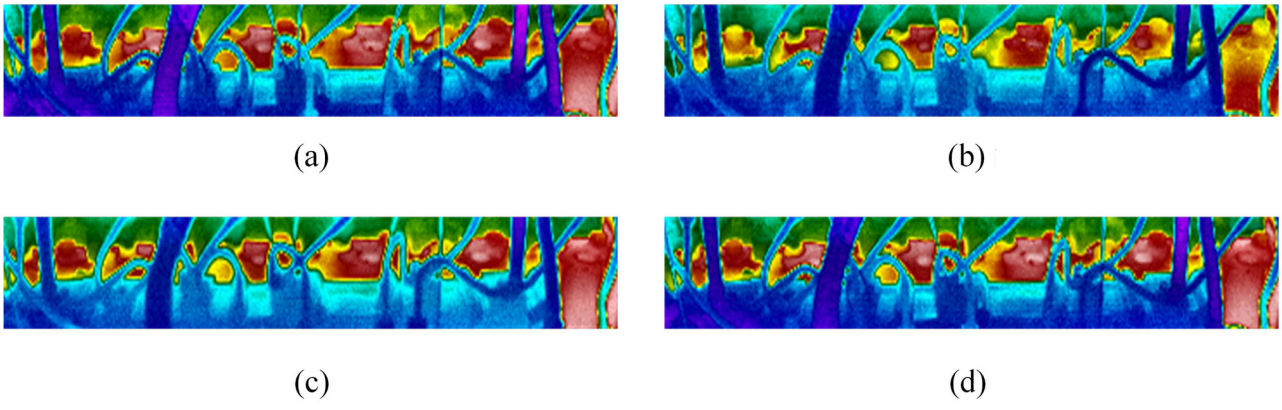

Figure 16: Infrared images of the diesel engine after AHE-based image enhancement: (a) NC, (b) SCM, (c) MCM, and (d) AFB.

Table 5: Detailed description of individual regions in dataset 2

Fault class	Class label	Number of training set images	Number of test set images
NC	1	24	6
SCM	2	24	6
MCM	3	24	6
AFB	4	24	6

5.2 Parameter settings of the involved DL methods

In this article, the AlexNet model in CNN is employed to extract fault features present in infrared images of diesel engines. AlexNet has significantly enhanced the accuracy of image recognition by leveraging several advantages. These include a deep network structure, a substantial number of convolutional layers and pooling layers, the application of the ReLU activation function, and the incorporation of dropout technology. The advancements have demonstrated the superiority of CNN in the field of image recognition, establishing AlexNet as a significant milestone in the field of DL. We used the pre-trained AlexNet model and made certain parameter adjustments to it according to the characteristics of the dataset employed.

AlexNet necessitates an input image size of 227 × 227 × 3, encompassing 5 convolutional layers, 3 pooling layers, and 3 fully connected layers. The filter consists of convolutional kernels of sizes 11 × 11, 5 × 5, and 3 × 3. Subsequently, a fully connected layer is employed to compute the probability of four classes and facilitate pattern recognition. The detailed structure is presented in Table 6. To showcase the effectiveness of the CNN-based method, various DL methods such as DNN, RNN, and SAE, which have been prominent in recent years, were compared and analyzed. These methods were then applied to the task of fault pattern recognition of diesel engines. In the context of RNN, LSTM is used to perform image classification. On the other hand, the DNN model employs MLP for the purpose of fault pattern recognition. The specific parameter configurations for the aforementioned methods are presented in Table 7.

6 Result analysis

6.1 Fault diagnosis results of diesel engines under dataset 1

The objective of the article is to evaluate the effectiveness of the proposed CNN-based fault diagnosis method for

Table 6: Parameter settings of AlexNet

Layer type	Number of kernels	Size of input map	Size of output map	Size of kernel	Stride	Padding
Input layer	/	227 × 227 × 3	/	/	/	/
Convolutional layer 1	96	227 × 227 × 3	55 × 55 × 96	11 × 11	4 × 4	0
Pooling layer 1	1	55 × 55 × 96	27 × 27 × 96	3 × 3	2 × 2	/
Convolutional layer 2	256	27 × 27 × 96	27 × 27 × 256	5 × 5	1 × 1	2 × 2
Pooling layer 2	1	27 × 27 × 256	13 × 13 × 256	3 × 3	2 × 2	/
Convolutional layer 3	384	13 × 13 × 256	13 × 13 × 384	3 × 3	1 × 1	1 × 1
Convolutional layer 4	384	13 × 13 × 384	13 × 13 × 384	3 × 3	1 × 1	1 × 1
Convolutional layer 5	256	13 × 13 × 384	13 × 13 × 256	3 × 3	1 × 1	1 × 1
Pooling layer 3	1	13 × 13 × 256	6 × 6 × 256	3 × 3	2 × 2	/
Fully connected layer 1	/	6 × 6 × 256	4096	/	/	/
Fully connected layer 2	/	4096	2048	/	/	/
Fully connected layer 3	/	2048	4	/	/	/
Output layer	/	4	4	/	/	/

Table 7: Parameter settings of comparison methods

Comparison method	Network size	Activation function	Classifier
MLP	[52185, 512, 256, 128, 4]	ReLU	Softmax
LSTM	[52185, 100, 100, 4]	Sigmoid	Softmax
SAE	[52185, 1024, 512, 128, 4]	ReLU	Softmax

Note: The input image size of the above three methods is 355 × 49 × 3.

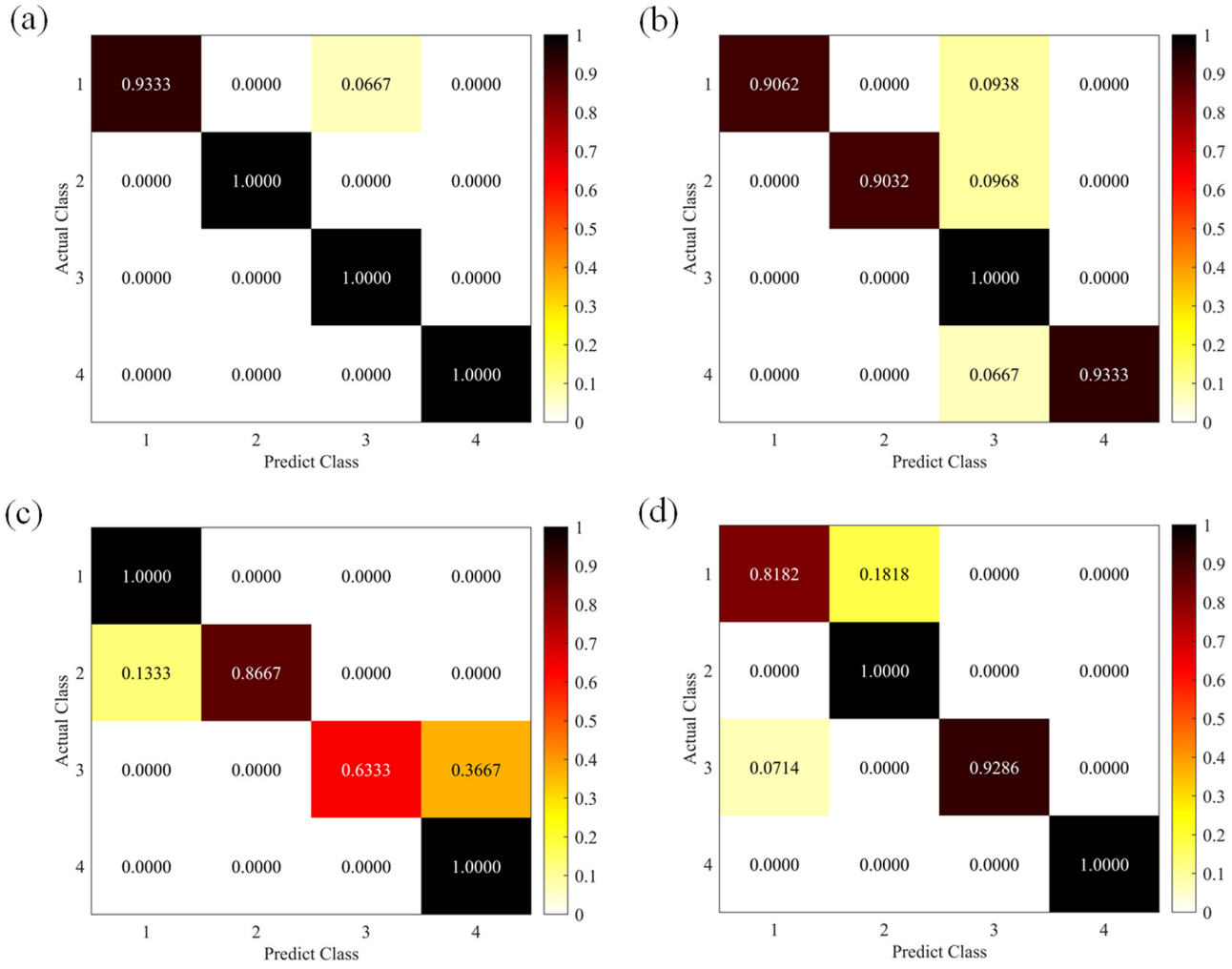

Figure 17: Classification results of four methods: (a) CNN; (b) MLP; (c) LSTM; (d) SAE.

diesel engines. However, it should be noted that this investigation does not take into account the impact of temperature variations on the surface of the diesel engine cylinder block. For the purpose of conducting a comparative analysis, this study employed various DL methods, including MLP, LSTM, and SAE. These methods were deployed to dynamically extract features of infrared images captured in mixed temperature regions and accurately classify the fault state of the diesel engine. The confusion matrix depicting the fault diagnosis results of the aforementioned four methods can be observed in Figure 17.

From the analysis of Figure 17, it is evident that the CNN-based method exhibits the highest fault diagnosis accuracy, regardless of temperature variations during diesel engine operation. Moreover, it demonstrates the capability to effectively differentiate between the four fault states of the diesel engine. When diagnosing diesel engine faults, MLP, LSTM, and SAE exhibit a higher number of misclassifications compared to the CNN-based method. Additionally, the classification accuracy is significantly lower than that of CNN-based methods. Simultaneously, to mitigate the influence of randomness in DL methods, each method performs 10 times of training and testing on dataset 1. This is done to determine the mean and standard deviation of the diagnostic accuracy, as presented in Table 8.

From the analysis of Table 8, it is evident that the CNN-based fault diagnosis method exhibits the highest

Table 8: Classification results of dataset 1

Methods	Average classification accuracy	Standard deviation
CNN	97.67%	0.0072
MLP	87.67%	0.0934
LSTM	83.33%	0.0456
SAE	88.50%	0.0423

classification accuracy and the lowest standard deviation, which achieves an accuracy rate that is 10.36% higher than the second-highest method based on SAE. The average testing accuracy of the LSTM model is the lowest, measuring only 83.33%. In relation to the stability of algorithms, the CNN-based method proposed in the article demonstrates a significant advantage over the other three DL methods. It is worth noting that MLP exhibits the highest standard deviation of 0.0934. Furthermore, MLP, LSTM, and SAE have a certain degree of mistaken identification between classes 2 and 3. In contrast, CNN demonstrates 100% accuracy in correctly identifying SCM and MCM. In conclusion, the CNN-based method proposed in this article demonstrates superior feature extraction capabilities in the recognition of diesel engine fault states.

6.2 Fault diagnosis results of diesel engines under dataset 2

In the engineering application of diesel engines, it is important to acknowledge that functional failures can potentially arise at any stage of operation. Therefore, with the objective of mitigating potential premature failures, this article seeks to showcase the efficacy of the method proposed in the research by segmenting the complete operational process of the diesel engine into five distinct regions. Simultaneously, the fault diagnosis results are compared with DL methods such as MLP, LSTM, and SAE to showcase the superior performance of the CNN-based fault diagnosis method proposed in the article. For the purpose of fault pattern recognition, the aforementioned four DL methods were employed on datasets from five different time regions. Each method was subjected to ten training and testing sessions, and the corresponding classification results are presented in Table 9 and Figure 18, respectively.

From the analysis of Table 9 and Figure 18, it is evident that among the five time regions of dataset 2, CNN-based fault diagnosis method exhibits the highest average classification accuracy. Furthermore, when considering all four methods, CNN also achieves the highest overall classification accuracy, reaching an impressive 93.58%. MLP demonstrates the second-highest average classification accuracy among the four methods, achieving an accuracy of 90%. The remaining two approaches, namely LSTM and SAE, exhibit comparatively lower average classification accuracy of 88.67 and 80.67%, respectively. On the contrary, the CNN-based method proposed in the article exhibits strong algorithm stability, as evidenced by its consistently low standard deviation of average classification accuracy. Among the four DL methods evaluated, the CNN-based method consistently achieves the lowest standard deviation of 0.0172 across five time regions and overall. The standard deviation of the average accuracy of SAE is the highest, reaching a value of 0.0654.

Furthermore, it is evident that during the initial time region, the classification accuracy of the four DL methods is relatively low. This can be attributed to the fact that in the early stage of diesel engine operation, the temperature of the cylinder and the environmental temperature are similar. Consequently, the training data for the model is inadequate, leading to interference in subsequent fault pattern recognition. The analysis results presented above indicate that the CNN-based fault diagnosis method proposed in the article is capable of effectively extracting fault features from infrared images of diesel engines.

The conclusion is supported by the fact that all four methods employed in the study use the SR classifier for final pattern recognition. Furthermore, by conducting a comparative analysis, it becomes evident that the surface temperature of the diesel engine cylinder block exhibits a gradual increase across the dataset of five consecutive time

Table 9: Classification results of dataset 2

Methods		Time regions					Average
		0–5 min	5–10 min	10–15 min	15–20 min	20–25 min	
CNN	Accuracy	70.42%	97.50%	100%	100%	100%	93.58%
	Std	0.0657	0.0204	0	0	0	0.0172
MLP	Accuracy	66.25%	92.07%	95.00%	96.67%	100%	90.00%
	Std	0.0684	0.0472	0.0520	0.0583	0	0.0452
LSTM	Accuracy	55.42%	93.75%	96.67%	97.50%	100%	88.67%
	Std	0.0813	0.0208	0.0583	0.0534	0	0.0428
SAE	Accuracy	36.67%	92.08%	98.33%	88.33%	87.92%	80.67%
	Std	0.1302	0.0292	0.0276	0.0717	0.0683	0.0654

Note: Std represents the standard deviation.

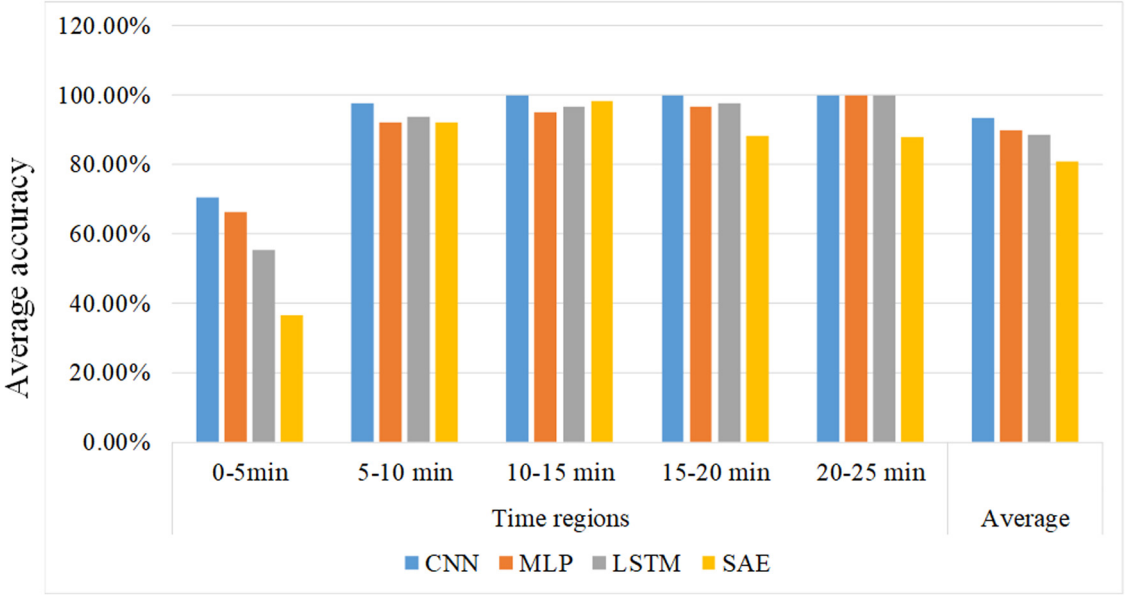

Figure 18: Diagnostic accuracy of CNN, MLP, LSTM, and SAE in five time regions.

regions. It is noteworthy that CNN demonstrates the highest accuracy in fault diagnosis and the lowest standard deviation of average accuracy. The aforementioned statement suggests that the CNN-based method, as proposed in the article, exhibits superior algorithm stability and anti-interference capability when it comes to fault diagnosis of diesel engines caused by temperature fluctuations.

6.3 Effect evaluation of image enhancement

In this article, the image pre-processing stage involved the application of AHE for infrared image enhancement. As depicted in Figure 14, image enhancement has resulted in a notable improvement in contrast for infrared images. To examine the impact of image enhancement on the feature extraction process of CNN, the features obtained after the final fully connected operation of CNN are chosen for visualization, as shown in Figure 19.

It can be clearly seen from Figure 19 that after AHE-based image enhancement, the feature distribution becomes more concentrated, with the distribution range of Dimension 1 and Dimension 2 reduced by half, and the feature separation degree has been improved, thereby facilitating subsequent feature extraction and pattern recognition.

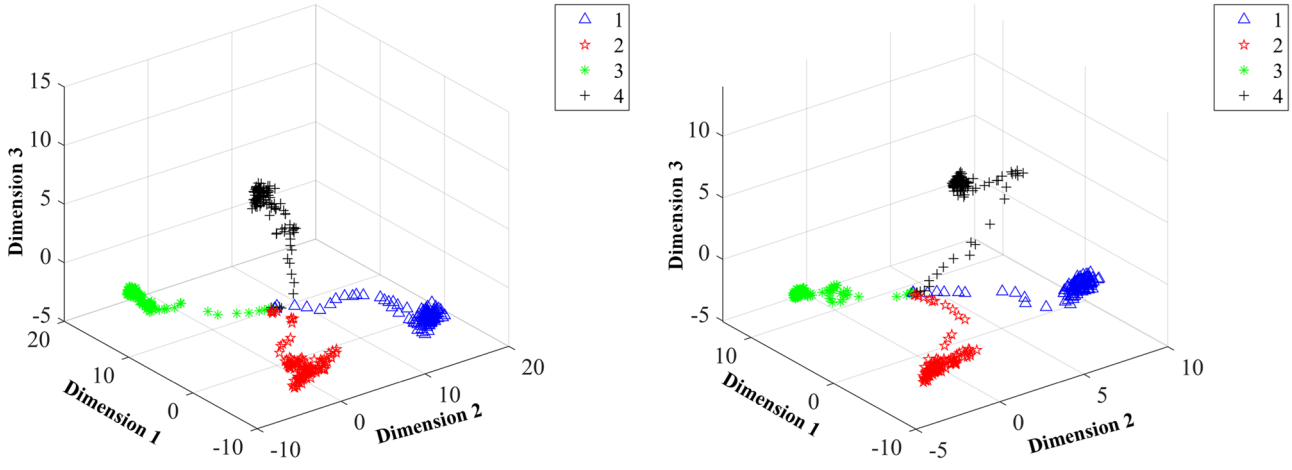

Figure 19: Comparison of feature visualization: (a) pre- and (b) post-AHE.

Table 10: Classification results of dataset 3

Methods	AHE	Average accuracy	Comparison
CNN	Y	97.67%	1.65%
	N	96.08%	
MLP	Y	87.67%	11.80%
	N	78.42%	
LSTM	Y	83.33%	2.66%
	N	81.17%	
SAE	Y	88.50%	1.72%
	N	87.00%	

To provide additional evidence on the impact of image enhancement on fault state recognition of diesel engines, we conducted training and testing using various methods including CNN, MLP, LSTM, and SAE on dataset 3 obtained from dataset 1 without undergoing any image enhancement. The network parameters remained constant, and the classification results are presented in Table 10 and Figure 20.

It is worth noting that the raw data used in the research process of this article is consistent with a previous article published by the author [63]. It is reasonable to say that the fault diagnosis results before applying AHE-based image enhancement in Table 10 of this article should be consistent with the results for Dataset 3 in Table 10 of the previous article. However, due to the unpredictability of DL methods, there may be significant differences in the results of each training and testing. Therefore, it can be explained that in this article, after retraining and testing the dataset 10 times, the results have shown some differences compared to the previous article. However, this difference is relative and does not affect the scientificity and accuracy of the results in this article. The results in Table 10 fully demonstrate the effectiveness of the AHE-based image enhancement method.

In Table 10 and Figure 20, the variable *Y* denotes the application of AHE for image enhancement, while the variable *N* signifies the absence of AHE in image enhancement. Table 10 demonstrates that when using the above four DL methods for feature extraction, AHE-based image enhancement can effectively improve fault diagnosis accuracy. Among the various models considered, the MLP-based method exhibited the most significant enhancement in fault diagnosis accuracy, achieving a notable increase of 11.80%. The fault diagnosis accuracy of the other three DL methods has also shown improvement to varying degrees following the application of AHE.

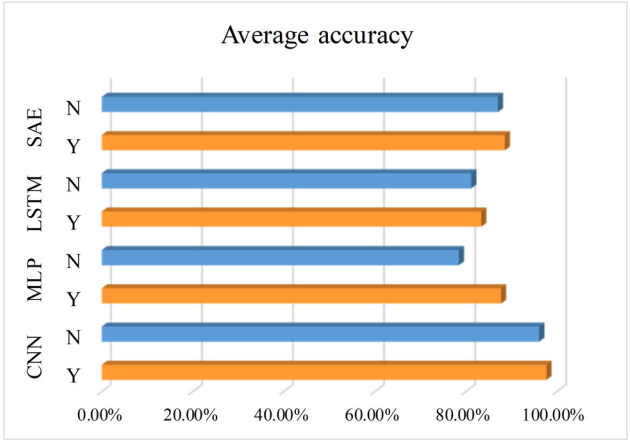

Figure 20: Comparison of diagnostic accuracy of CNN, MLP, LSTM, and SAE.

6.4 Generalization ability verification of CNN

To elaborate on the generalization ability [64] of the CNN for infrared image classification under study, we have set up a situation where the training dataset is unbalanced. For dataset 1, we randomly deleted 20% of the infrared images under class label 1, which means there are still 120 infrared images left under this label. Among them, the training set and test set contain 96 and 24 images, respectively, and the infrared images under the other class labels remain unchanged. Set the epochs to 10 because in

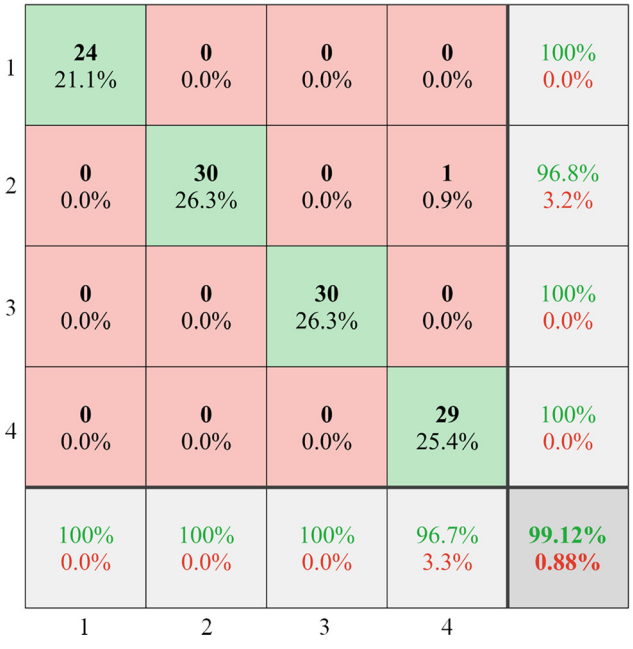

Figure 21: Confusion matrix of fault diagnosis results.

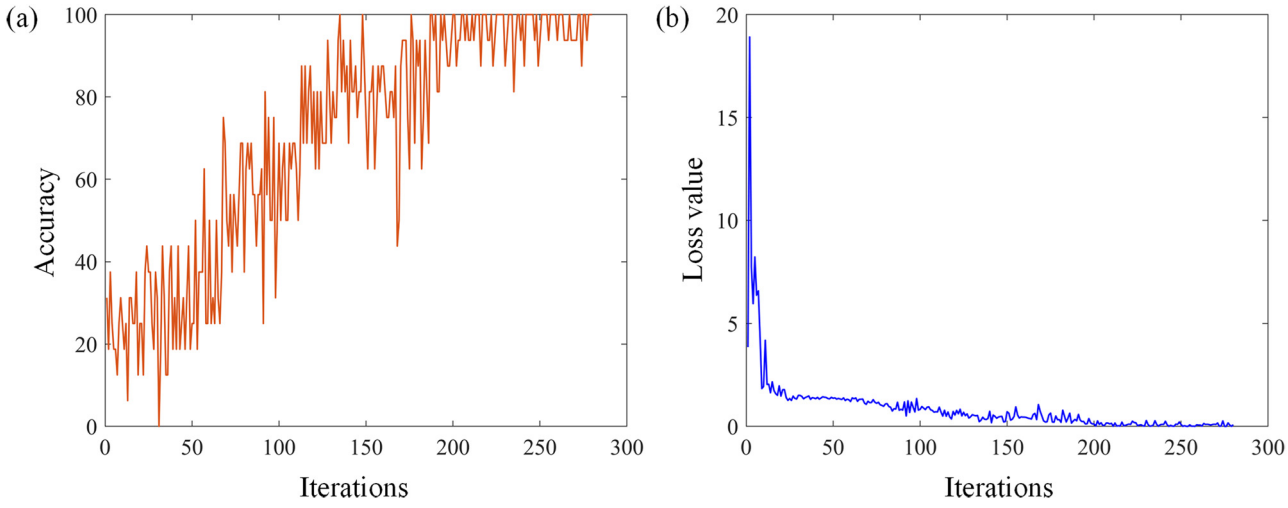

Figure 22: The variation curves of accuracy and loss value: (a) accuracy, (b) loss value.

this case, the training accuracy curve and loss value curve of CNN have become stable, and then record the entire training and testing process time. The confusion matrix of the highest fault diagnosis accuracy of 10 times of training and test results is taken and presented, as shown in Figure 21.

The variation curves of fault diagnosis accuracy and loss value during the training process are shown in Figure 22.

From Figure 21, it can be seen that even with reduced training data under class label 1, the fault diagnosis accuracy can still reach 99.12%, with an average accuracy of 97.28% and a standard deviation of 0.018. Furthermore, all test data under classs label 1 have been correctly identified. Therefore, the above facts are sufficient to prove the good generalization ability of the proposed CNN for infrared image classification.

In terms of hardware configuration, a computer of Windows 11 is used in this study, with 13th Gen Intel(R) Core(TM) i5-13600KF 3.50 GHz CPU, NVIDIA GeForce RTX 4060 GPU, and the RAM is 32GB DDR5. Moreover, the total number of iterations is 280, taking 25 s. The number of parameters of the CNN used in this article is about 60 M. Therefore, we can conclude that the method proposed in this article has the acceptable feasibility of online monitoring and can identify the real-time status of diesel engines accurately. In addition, due to the relatively large number of parameters, the model needs to be compressed and pruned in subsequent embedding tasks, and the hardware configuration of the computer needs to be improved.

7 Conclusions

This article introduces a highly efficient fault diagnosis approach for diesel engines, using the combination of IRT and CNN. In order to address four prevalent fault states of diesel engines, we have collected the corresponding infrared images. The effectiveness of the proposed method is verified in comparison to DL methods such as MLP, LSTM, and SAE. Simultaneously, the article investigates the impact of temperature variations on the fault diagnosis efficacy of DL methods. To summarize, the article makes the following key contributions:

1) This study proposes an effective fault diagnosis method for diesel engines by combining IRT with CNN. The proposed method is capable of automatically extracting fault features from infrared images and demonstrates excellent algorithm stability and anti-interference ability against temperature fluctuations.
2) The AHE-based image enhancement, proposed in the article, demonstrates its effectiveness in enhancing the fault diagnosis accuracy. Moreover, it exhibits a certain level of adaptability to DL methods, including CNN, MLP, LSTM, and SAE.
3) The experimental results demonstrate that the proposed CNN has a good generalization ability for infrared image classification and acceptable feasibility of online monitoring, and exhibits superior fault diagnosis performance and effectively identifies mixed faults (MCM) when compared to MLP, LSTM, and SAE.

In future research endeavors, the primary focus will be on addressing the issue of low accuracy in fault diagnosis during the early stages of diesel engine operation. This will involve conducting research on few-shot image classification, as acquiring an adequate amount of data is challenging.

Funding information: The research work was funded by the National Natural Science Foundation of China (No. 71871219).

Author contributions: Conceptualization: Rongcai Wang and Xisheng Jia; methodology: Rongcai Wang and Hao Yan; writing – original draft preparation: Rongcai Wang; writing – review and editing: Enzhi Dong and Yuan Li; and supervision: Zhonghua Cheng and Xisheng Jia. All authors have accepted responsibility for the entire content of this manuscript and approved its submission.

Conflict of interest: The authors state no conflict of interest.

Data availability statement: All data generated or analyzed during this study are included in this published article.

References

[1] Li ZX, Jiang Y, Duan ZH, Peng ZX. A new swarm intelligence optimized multiclass multi-kernel relevant vector machine: An experimental analysis in failure diagnostics of diesel engines. Struct Health Monit. 2018;17(6):1503–19.

[2] Hou LS, Zou JQ, Du CJ, Zhang JD. A fault diagnosis model of marine diesel engine cylinder based on modified genetic algorithm and multilayer perceptron. Soft Comput. 2020;24:7603–13.

[3] Yan J, Liu TL, Ye XY, Jing QZ, Dai YN. Rotating machinery fault diagnosis based on a novel lightweight convolutional neural network. PLoS One. 2021;16(8):e0256287.

[4] Xue F, Zhang WM, Xue F, Li DD, Xie SL, Fleischer J. A novel intelligent fault diagnosis method of rolling bearing based on two-stream feature fusion convolutional neural network. Measurement. 2021;176:109226.

[5] Zhang GY, Wang Y, Li XM, Tang BP, Qin Y. Enhanced symplectic geometry mode decomposition and its application to rotating machinery fault diagnosis under variable speed conditions. Mech Syst Signal Process. 2022;170:108841.

[6] Toutountzakis T, Tan CK, Mba D. Application of acoustic emission to seeded gear fault detection. NDT E Int. 2005;38:27–36.

[7] Cheng CC, Yang SY, Lee D. Novel real-time temperature diagnosis of conventional hot-embossing process using an ultrasonic transducer. Sensors. 2014;14(10):19493–506.

[8] Raposo H, Farinha JT, Fonseca I, Ferreira LA. Condition monitoring with prediction based on diesel engine oil analysis: A case study for urban buses. Actuators. 2019;8(1):14.

[9] Glowacz A, Glowacz Z. Diagnostics of stator faults of the single-phase induction motor using thermal images, MoASoS and selected classifiers. Measurement. 2016;93:86–93.

[10] Taheri-Garavand A, Ahmadi H, Omid M, Mohtasebi SS, Mollazade K, Smith AJR, et al. An intelligent approach for cooling radiator fault diagnosis based on infrared thermal image processing technique. Appl Therm Eng. 2015;87:434–43.

[11] Glowacz A, Glowacz Z. Diagnosis of the three-phase induction motor using thermal imaging. Infrared Phys Technol. 2017;81:7–16.

[12] Duan LX, Yao MC, Wang JJ, Bai TB, Zhang LB. Segmented infrared image analysis for rotating machinery fault diagnosis. Infrared Phys Technol. 2016;77:267–76.

[13] Zabihi-Hesari A, Ansari-Rad S, Shirazi FA. Fault detection and diagnosis of a 12-cylinder trainset diesel engine based on vibration signature analysis and neural network. Proc Inst Mech Eng C-J Mech Eng Sci. 2019;233(6):1910–23.

[14] Gu C, Qiao XY, Li HY, Jin Y. Misfire fault diagnosis method for diesel engine based on MEMD and dispersion entropy. Shock Vib. 2021;2021:9213697.

[15] Bi XY, Cao SQ, Zhang DM. Diesel engine valve clearance fault diagnosis based on improved variational mode decomposition and bispectrum. Energies. 2019;12(4):661.

[16] Ciabattoni L, Ferracuti F, Freddi A, Monteri A. Statistical spectral analysis for fault diagnosis of rotating machines. IEEE Trans Ind Electron. 2018;65(5):4301–10.

[17] Zhao HP, Zhang JJ, Jiang ZN, Wei DH, Zhang XD, Mao ZW. A new fault diagnosis method for a diesel engine based on an optimized vibration mel frequency under multiple operation conditions. Sensors. 2019;19(11):2590.

[18] Yan H, Bai HJ, Zhan XB, Wu ZH, Wen L, Jia XS. Combination of VMD mapping MFCC and LSTM: A new acoustic fault diagnosis method of diesel engine. Sensors. 2022;22(21):8325.

[19] Janssens O, Loccufier M, Van DWR, Hoecke SV. Data-driven imbalance and hard particle detection in rotating machinery using infrared thermal imaging. Infrared Phys Technol. 2017;82:28–39.

[20] Ciampa F, Mahmoodi P, Pinto F, Meo M. Recent advances in active infrared thermography for non-destructive testing of aerospace components. Sensors. 2018;18(2):609.

[21] Laborda A, Robinson A, Wang SC, Zhang Y, Reed P. Fatigue assessment of multilayer coatings using lock-in thermography. Mater Des. 2018;141:361–73.

[22] Li YB, Du XQ, Wan FY, Wang XZ, Yu HC. Rotating machinery fault diagnosis based on convolutional neural network and infrared thermal imaging. Chin J Aeronaut. 2020;33(2):427–38.

[23] Li X, Shao HD, Jiang HK, Xiang JW. Modified Gaussian convolutional deep belief network and infrared thermal imaging for intelligent fault diagnosis of rotor-bearing system under time-varying speeds. Struct Health Monit. 2022;21(2):339–53.

[24] Wang RC, Zhan XB, Bai HJ, Dong EZ, Cheng ZH, Jia XS. A review of fault diagnosis methods for rotating machinery using infrared thermography. Micromachines. 2022;13(10):1644.

[25] Younus AM, Widodo A, Yang BS. Evaluation of thermography image data for machine fault diagnosis. Nondestruct Test Eval. 2010;25(3):231–47.

[26] Younus AMD, Yang BS. Intelligent fault diagnosis of rotating machinery using infrared thermal image. Expert Syst Appl. 2012;39(1):2082–91.

[27] Widodo A, Satrijo D, Prahasto T, Lim GM, Choi BK. Confirmation of thermal images and vibration signals for intelligent machine fault diagnostics. Int J Rotating Mach. 2012;2012:847203.

[28] Bagavathiappan S, Saravanan T, George NP, Philip J, Jayakumar T, Raj B. Condition monitoring of exhaust system blowers using infrared thermography. Insight Non-Destr Test Cond Monit. 2008;50(9):512–5.
[29] Janssens O, Schulz R, Slavkovikj V, Stockman K, Loccufier M, Walle RVD, et al. Thermal image based fault diagnosis for rotating machinery. Infrared Phys Technol. 2015;73:78–87.
[30] Li YB, Wang XZ, Si SB, Du XQ. A new intelligent fault diagnosis method of rotating machinery under varying-speed conditions using infrared thermography. Complexity. 2019;2019:2619252.
[31] Dadon I, Koren N, Klein R, Bortman J. A realistic dynamic model for gear fault diagnosis. Eng Fail Anal. 2018;84:77–100.
[32] Xu XJ, Zhao ZZ, Xu XB, Yang JB, Chang LL, Yan XP, et al. Machine learning-based wear fault diagnosis for marine diesel engine by fusing multiple data-driven models. Knowl Syst. 2020;190:105324.
[33] Taghizadeh-Alisaraei A, Ghobadian B, Tavakoli-Hashjin T, Mohtasebi SS, Rezaei-asl A, Azadbakht M. Characterization of engine's combustion-vibration using diesel and biodiesel fuel blends by time-frequency methods: A case study. Renew Energy. 2016;95:422–32.
[34] Vattulainen J, Nummela V, Hernberg R, Kytölä J. A system for quantitative imaging diagnostics and its application to pyrometric in-cylinder flame-temperature measurements in large diesel engines. Meas Sci Technol. 2000;11(2):103–19.
[35] Macián V, Tormos B, Ruiz S, Miró G. Low viscosity engine oils: Study of wear effects and oil key parameters in a heavy duty engine fleet test. Tribol Int. 2016;94:240–8.
[36] Guerrero DP, Jiménez-Espadafor FJ. Torsional system dynamics of low speed diesel engines based on instantaneous torque: Application to engine diagnosis. Mech Syst Signal Process. 2019;116:858–78.
[37] Lei YG, Yang B, Jiang XW, Jia F, Li NP, Nandi AK. Applications of machine learning to machine fault diagnosis: A review and roadmap. Mech Syst Signal Process. 2020;138:106587.
[38] Wang B, Jahanshahi H, Volos C, Bekiros S, Khan MA, Agarwal P, et al. A new RBF neural network-based fault-tolerant active control for fractional time-delayed systems. Electronics. 2021;10(12):1501.
[39] Flett J, Bone GM. Fault detection and diagnosis of diesel engine valve trains. Mech Syst Signal Process. 2016;72-73:316–27.
[40] Kowalski J, Krawczyk B, Woźniak M. Fault diagnosis of marine 4-stroke diesel engines using a one-vs-one extreme learning ensemble. Eng Appl Artif Intell. 2017;57:134–41.
[41] Ramteke SM, Chelladurai H, Amarnath M. Diagnosis and classification of diesel engine components faults using time–frequency and machine learning approach. J Vib Eng Technol. 2022;10:175–92.
[42] Jebadass JR, Balasubramaniam P. Low contrast enhancement technique for color images using interval-valued intuitionistic fuzzy sets with contrast limited adaptive histogram equalization. Soft Comput. 2022;26:4949–60.
[43] Rahman H, Chandra G. Tripartite sub-image histogram equalization for slightly low contrast gray-tone image enhancement. Pattern Recognit. 2023;134:109043.
[44] Ren X, Lai S. Medical image enhancement based on laplace transform, sobel operator and histogram equalization. Acad J Comput Inf Sci. 2022;5(6):48–54.
[45] Chen YF, Lin JH, Zhang SQ, Chen CM. Image hiding algorithm based on block random scrambling and grayscale transformation. 2010 International Conference on Display and Photonics; 2010.
[46] Bahraini T, Hamedani T, Mohammad S, Sadoghi H. Edge preserving range image smoothing using hybrid locally kernel-based weighted least square. Appl Soft Comput. 2022;125:109234.
[47] Yan WQ, Xu GH, Du YH, Chen XB. SSVEP-EEG feature enhancement method using an image sharpening filter. IEEE Trans Neural Syst Rehabil Eng. 2022;30:115–23.
[48] Huang ZH, Zhang YZ, Li Q, Zhang TX, Sang N, Hong HY. Progressive dual-domain filter for enhancing and denoising optical remote-sensing images. IEEE Geosci Remote Sensing Lett. 2018;15(5):759–63.
[49] Kim S, Kang W, Lee E, Paik J. Wavelet-domain color image enhancement using filtered directional bases and frequency-adaptive shrinkage. IEEE Trans Consum Electron. 2010;56(2):1063–70.
[50] Thillainayagi R, Kumar KS. Bi-dimensional empirical mode decomposition based contrast enhancement technique for UAV thermal images. IETE J Res. 2019;68(1):630–7.
[51] Kim YT. Contrast enhancement using brightness preserving bi-histogram equalization. IEEE Trans Consum Electron. 1997;43(1):1–8.
[52] Acharya UK, Kumar S. Genetic algorithm based adaptive histogram equalization (GAAHE) technique for medical image enhancement. Optik. 2021;230(1):166273.
[53] Dar KA, Mittal S. An enhanced adaptive histogram equalization based local contrast preserving technique for HDR images. IOP Conf Ser: Mater Sci Eng. 2021;1022:012119.
[54] Li YB, Gu JX, Zhen D, Xu MQ, Ball A. An evaluation of gearbox condition monitoring using infrared thermal images applied with convolutional neural networks. Sensors. 2019;19(9):2205.
[55] Jia F, Lei YG, Lin J, Zhou X, Lu N. Deep neural networks: A promising tool for fault characteristic mining and intelligent diagnosis of rotating machinery with massive data. Mech Syst Signal Process. 2016;72-73:303–15.
[56] Zhu JJ, Jiang QS, Shen YH, Qian CH, Xu FY, Zhu QX. Application of recurrent neural network to mechanical fault diagnosis: a review. J Mech Sci Technol. 2022;36:527–42.
[57] Rajchakit G, Sriraman R, Boonsatit N, Hammachukiattikul P, Lim CP, Agarwal P. Exponential stability in the Lagrange sense for Clifford-valued recurrent neural networks with time delays. Adv Differ Equ. 2021;2021:256.
[58] Luo SY, Huang XF, Wang YZ, Luo RM, Zhou Q. Transfer learning based on improved stacked autoencoder for bearing fault diagnosis. Knowl Syst. 2022;256:109846.
[59] Li X, Zhang W, Ma H, Luo Z, Li X. Domain generalization in rotating machinery fault diagnostics using deep neural networks. Neurocomputing. 2020;403:409–20.
[60] Mao WT, Feng WS, Liang XH. A novel deep output kernel learning method for bearing fault structural diagnosis. Mech Syst Signal Process. 2019;117:293–318.
[61] Wen L, Li XY, Gao L, Zhang YY. A new convolutional neural network-based data-driven fault diagnosis method. IEEE Trans Ind Electron. 2018;65(7):5990–8.
[62] Lei YG, Jia F, Lin J, Xing SB, Ding SX. An intelligent fault diagnosis method using unsupervised feature learning towards mechanical big data. IEEE Trans Ind Electron. 2016;63(5):3137–47.
[63] Wang RC, Jia XS, Liu ZC, Dong E, Li S, Cheng Z. Conditional generative adversarial network based data augmentation for fault diagnosis of diesel engines applied with infrared thermography and deep convolutional neural network. Eksploat Niezawodn. 2024;26(1):175291.
[64] Rajchakit G, Agarwal P, Ramalingam S. Stability analysis of neural networks. Singapore: Springer; 2021.