IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2017.Doi Number

Image classification for Automobile pipe joints surface defect detection using Wavelet decomposition and Convolutional neural network

Zeqing Yang[1,2], Mingxuan Zhang[1], Chao Li[1], Zhaozong Meng[1], Yue Li[1], Yingshu Chen[1,2], Libing Liu[1,2]

[1]School of Mechanical Engineering, Hebei University of Technology, Tianjin 300130, China

[2]National Engineering Research Center for Technological Innovation Method and Tool, Hebei University of Technology, Tianjin 300130, China

Corresponding author: Yingshu Chen (e-mail: chenyingshu@hebut.edu.cn).

This work was supported in part by National Natural Science Foundation of China under Grant 52175461, Top Young Talents Project of Hebei Province under Grant 210014, Intelligent Manufacturing Project of Tianjin under Grant 20201199.

ABSTRACT The surface defect detection of automobile pipe joints based on computer vision faces technical challenges. The tiny-sized and smooth surfaces with processing textures will undermine the defect detection accuracy. In order to solve this problem, a new method was proposed, which combines wavelet decomposition and reconstruction with the canny operator to detect defects, and then uses the multi-channel fusion convolutional neural network to identify the types of defects. Firstly, illumination compensation technology is used to obtain a more uniform gray distribution of the original image. Then, the wavelet decomposition and reconstruction are used to remove noises and processing textures. Furthermore, the defect regions are segmented using the canny operator and hole filling from the image. Finally, the multi-channel fusion convolutional neural network of decision-level is used to identify the surface defect types. This method provides an idea for the surface defects detection of automobile pipe joints with serious interference, such as smooth surface, random noises, and processing textures. The experimental results reveal that the method can effectively eliminate the influence of uneven illumination, random noises, and processing textures and achieve high defect classification accuracy.

INDEX TERMS Automobile pipe joint, surface defect detection, wavelet decomposition and reconstruction, multi-channel fusion convolutional neural network.

I. INTRODUCTION

Automobile pipe joints are precision metal parts used as the engine's air pipes and oil pipes. The surface defects directly affect the sealing performance, assembly accuracy, and service life of the whole pipeline and even affect the driving safety of the automobiles in severe cases. Therefore, defect detection is crucial to control product quality effectively.

Common defect types of automobile pipe joints include scratches, pits, and burrs. The causes of its defects are as follows:(1) surface damage caused by changes in hardness and stress state in the surface structure due to grinding heat and force; (2) surface damage caused by abrasion of the machining tool; (3) mechanical damage caused by collision and scratching.

At present, off-line manual inspection is still used to detect the surface defects of engine pipe joints. This kind of long-term repeated measurement is easily affected by personnel fatigue and subjective judgment, resulting in low efficiency and accuracy [1]. The application of machine vision inspection technology to the production line of automobile engine pipe joints can detect and classify surface defects of parts and components and help improve the automation and intelligence level of equipment [2].

Machine vision inspection of the surface defects of automobile pipe joints faces the problems of low contrast between defective area and non-defective area, the high similarity between processing texture and acceptable defects, and low defect accuracy. All these problems can lead to

indistinguishability between the surface defects and other areas, making the accuracy of the surface defect detection method based on machine vision so low that it is challenging to meet the requirements of the actual part manufacturing detection process. These problems will cause difficulties in subsequent defect detection and classification, so a reasonable image preprocessing process is necessary to remove various interferences and improve the efficiency and accuracy of defect detection.

Convolutional neural network (CNN) is a deep learning method that is widely used to solve complicated problems, which overcomes the limitations of traditional machine learning methods [3]. CNN has been widely used in machine vision, especially in the fields of image recognition and image classification [4]. Compared with traditional methods, CNN can automatically extract and learn deep and specific features to update model parameters [5]. Its expression of the object is more efficient and accurate, and its robust ability is better than traditional pattern recognition methods.

However, the single CNN faces the challenge of high variance in the prediction results brought by the random training process, so sometimes it cannot meet the surface defect detection accuracy requirements. Therefore, the fusion network method, which can reduce prediction variance compared to the single network, has been widely used. The typical fusion network methods include pixel-level fusion, feature-level fusion, and decision-level fusion [6]. Pixel-level fusion refers to directly processing the pixels of an image to obtain a fused image. It can retain more information with high accuracy but low efficiency, poor analytical ability, and weak anti-interference ability [7]. Feature-level fusion is to process the features extracted from the source image information and generates a fused feature vector [8]. The fused feature information is identified and classified. The advantages are high processing speed and a small amount of calculation, but the information loss is excellent. The decision-level fusion is the process of judging and reasoning the image [9]. The advantages of decision-level fusion are fault tolerance, openness, short processing time, low data requirements, and strong analytical ability. The surface defects of automobile pipeline joints are minor and contain less feature information. Using the feature-level fusion method will cause information loss, which will weaken the ability of CNN defect classification. The decision-level fusion retains the integrity of the feature extraction information. It is fused on the classification results of a single CNN, which is more suitable for the surface defect classification of automobile pipe joints. However, the requirements of decision-level fusion for preprocessing are high [10]. Appropriate preprocessing methods can avoid wasting the energy of feature extraction on filtering interference information and can underline the features of the defects themselves, helping classification and fusion.

In this paper, new defect detection and classification method are proposed, aiming at the problems of uneven illumination, random noise, and processing texture in the images of automobile pipeline joints. The defect detection and classification of automobile pipe joints includes image preprocessing, initial defect location, and defect type identification. Light source illumination compensation and wavelet denoising can reduce the uneven distribution of grayscale and improve image quality. Canny edge operator combined with hole filling is used to initially locate the defects, which improves the efficiency of subsequent defect identification. Multi-channel fusion CNN of decision-level is used to identify and classify defects.

Summing up, the contributions of this paper are as follows:

• The image denoising method based on illumination compensation and second-order wavelet decomposition is proposed, which can effectively remove the processing texture and random noise of the parts in the image.

• The multi-channel fusion CNN of decision-level is proposed to identify the surface defect types of parts, which has higher classification accuracy than a single network.

II. REELATED WORK

Researchers have conducted much research on the visual inspection and classification of product surface defects. In order to improve the detection accuracy, a filtering algorithm is introduced to remove interference noise in the detection of surface defects of parts, which can effectively improve the feature extraction effect [11]. Li et al. [12] used Fourier transform and Butterworth high-pass filter to effectively remove random texture and background noise, which solved the problem of random texture mixing of surface defects of small-sized annular parts. Yang et al. [13] proposed a magnetic tile defect detection method based on stationary wavelet transform (SWT), which uses a nonlinear image enhancement algorithm to achieve target defect enhancement and solves the problem of magnetic tile surface under different lighting conditions.

Median filtering has a good denoising effect on images containing uniform salt and pepper noise. It can effectively protect the edges of the image after denoising and try to avoid blurring [14]. Wavelet filtering is widely used in time-frequency analysis and multi-scale analysis [15]. It can effectively filter random noise mixed in high-frequency signals and distinguish defects and interference points in the image [16]. Image denoising methods combining median filtering and wavelet transform are generally used for images with high-frequency and salt-and-pepper noise. Lin et al. [17] proposed Gaussian mixture model estimation thresholds to determine noise-free wavelet coefficients. Finally, the denoised image is obtained through wavelet reconstruction, which can effectively remove the mixed noise in the complex background. M. Olfa et al. [18] analyzed the characteristics of Ultrasound images according to Bayesian maximum posterior probability. They proposed an image denoising algorithm based on wavelet transform and bilateral filtering, which can remove the speckle noise in the high-frequency

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2022.3178380, IEEE Access

and low-frequency components of the image. J. L. Song et al. [19] proposed an image denoising method based on the curvature change model and wavelet transform, which successfully removed the noise in the high-frequency components of the original image.

In order to improve the detection classification rate, typical defect detection classification methods are used, such as digital morphology and deep neural networks. Tsai et al. [20] proposed a machine vision-based detection method for minor defects on the machining surface of the circular parts marking texture. This method is based on digital morphology, introduces any size-shaped structural elements (SE), and performs morphological operations. Successfully removed the influence of round part machining traces on defect detection and strengthened the contour characteristics of the defect. Experimental results show that using the method for image preprocessing is of great help in detecting various minor defects such as scratches, bumps, and edge bursts on the surface of round parts. It is helpful for subsequent classifiers to classify defect types. Guo et al. [21] proposed a surface defect detection method for wind turbine blades that combines Haar-AdaBoost and CNN. Haar-AdaBoost is used for area search, and then CNN performs defect detection in this area. The actual data of the wind power plant was used to compare the method with support vector machine (SVM) and neural network. The test results show that the method has higher accuracy and stronger robustness. Nevertheless, the single network has sometimes been challenging to fit the required requirements. The channel fusion method has been applied gradually. He et al. [22] proposed a detection method of multi-classifier fusion to detect steel surface defects. In this method, the classification priority network (CPN) is used as the framework, and multiple convolutional neural network (MG-CNN) is used as the backbone network. This method achieves a 94% detection rate of surface defects in hot-rolled strips, reflecting the advantages of the fusion network.

Based on the existing research, this paper provides a method to detect the disturbance of tiny surface defects, such as uneven illumination, random noise, and processing texture. The method can solve the surface defect detection problem of automobile pipeline joints with random noise and processing texture, which combines light source illumination compensation, wavelet decomposition reconstruction, Canny edge detection, and multi-channel fusion convolutional neural network.

III. SURFACE DEFECT DETECTION AND CLASSIFICATION METHOD

Figure 1 shows the process of surface defects detection method for automobile pipe joints. In this method, the preprocessing process removes the noise and processing texture of surface defects in the captured image. Canny edge operator combined with hole filling is used to extract defect features, and the defects are initially located. Then a three-channel fusion convolutional neural network model of decision-level is designed. The same coarsely located defect image is input into three different structures. The pre-trained CNN is processed to obtain three classification matrices, and then the three classification matrices are fused. A new fused classification matrix is obtained to classify defects accurately.

The process of preprocessing is shown in Figure 2. The preprocessing technology of light source illumination compensation and wavelet denoising can reduce the uneven distribution of grayscale and improve image quality.

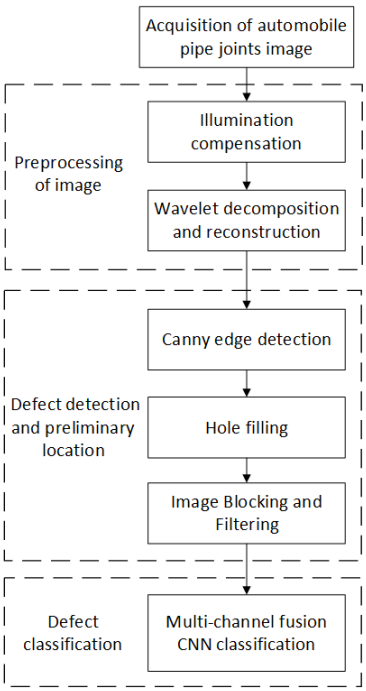


FIGURE 1. Defect detection and classification process.

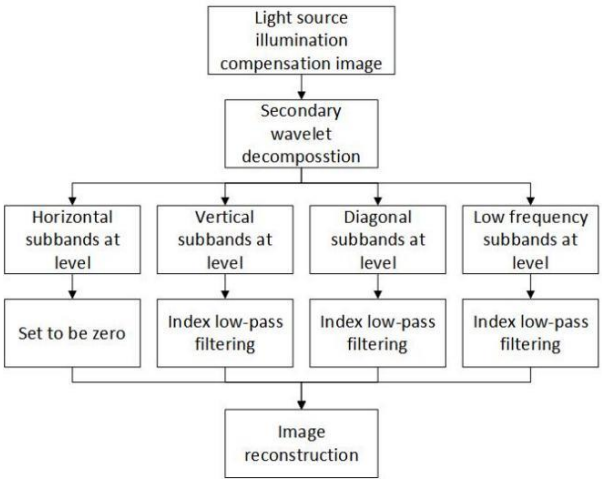


FIGURE 2. Image preprocessing process.

A. COMPENSATION OF LIGHT SOURCE ILLUMINATION

In the process of surface defect inspection, the image collected by an industrial camera would appear uneven grayscale, which will affect the performance of surface

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2022.3178380, IEEE Access

defect inspection. There are two main reasons for the uneven distribution of grayscale on the surface image of automobile pipe joints. The bending structure of the part results in uneven reflection, and the unequal distance between the light source and the part results in luminance unevenness.

The uneven grayscale caused by uneven reflection mainly exists between each column of pixels in the image of automobile pipeline joints. The compensation method is as Equations (1) to (4):

$$f_{cor}(x,y) = f(x,y) * a(x) \tag{1}$$

$$a(x) = f_{\text{mean}} \div f_m(x) \tag{2}$$

$$f_{mean} = \frac{\sum_{x=0}^{\mathrm{M}} \sum_{y=0}^{\mathrm{N}} F(x,y)}{\mathrm{MN}} \tag{3}$$

$$f_m(x) = \frac{\sum_{y=0}^{\mathrm{N}} f(x,y)}{\mathrm{N}} \tag{4}$$

Where the number of columns in the image $f(x,y)$ is M, and the number of rows is N, $x \in (0,\mathrm{M}), y \in (0,\mathrm{N})$. The image after compensation of reflection unevenness is $f_{cor}(x,y)$.

The uneven grayscale caused by uneven luminance mainly exists between each row of pixels in the image of automobile pipeline joints. The compensation method is as Equations (5) and (6):

$$f_{col}(x,y) = f(x,y) + b(y) \tag{5}$$

$$b(y) = \frac{|f(\mathrm{M}) - f(1)|}{\mathrm{M}} \times y \tag{6}$$

Where the number of columns in the image $f(x,y)$ is M, and the number of rows is N, $x \in (0,\mathrm{M}), y \in (0,\mathrm{N})$. The image after compensation of luminance unevenness is $f_{col}(x,y)$.

Figure 3 compares the average gray value in each column of pixels of automobile pipe joints image before and after compensation of reflection unevenness. Figure 4 compares the average gray value in each row of pixels before and after the compensation of the luminance unevenness. The red * line indicates the gray value before the compensation. The blue line represents the compensated gray value. The position with a high gray value indicates that the color of the image is bright, and a position with a low gray value indicates that the color is dark. The gray curve before compensation is concave, indicating that the gray distribution is uneven. The compensated gray value curve is relatively flat, indicating that the overall gray of the image is relatively uniform.

Figure 5 compares automobile pipe joints before and after the illumination compensation. After illumination compensation, the overall brightness distribution of the image is more even. Compared with Figure 5(a), Figure 5(b) no longer shows a state where the pixels in the two side columns are brighter than the pixels in the middle column. At the same time, compared with Fig. 5(a), Fig. 5(b) no longer presents a state in which the pixels in the upper row are brighter than the pixels in the lower row.

B. IMAGE WAVELET DENOISING

Although the illumination compensation has been carried out, it is still difficult to detect surface defects. Defect detection and classification of processing texture interference caused by parts processing. The processing texture in the image is the main influencing factor that interferes with the feature extraction of defects [23]. Therefore, the method of this paper enhances the complementary image through discrete wavelet transform, weakens the processed texture, and removes the interference of uneven background in the image, thus enhancing the image [13].

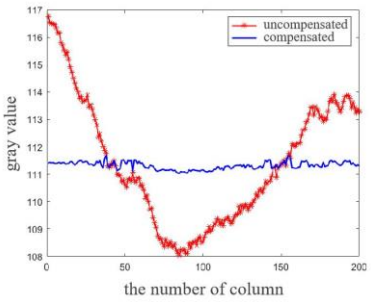


FIGURE 3. Contrast before and after the reflection unevenness compensation.

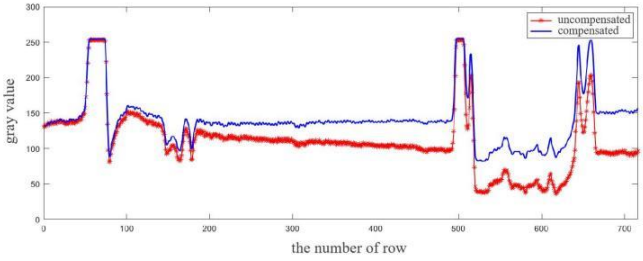


FIGURE 4. Contrast before and after the intensity unevenness compensation.

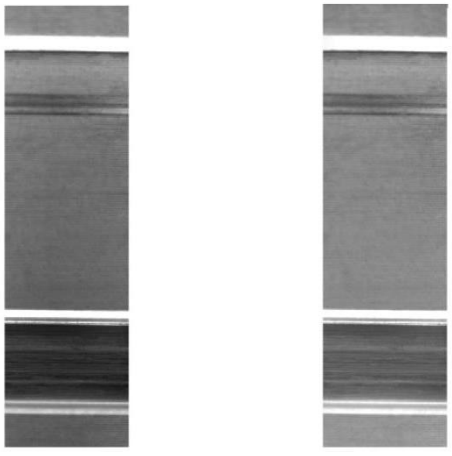


FIGURE 5. Compare images before and after source illumination compensation: (a) uncompensated image, (b) compensated image.

The surface image of automobile pipe joints often contains much information, such as structure, texture, noise, and defect information. Therefore, image decomposition is

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2022.3178380, IEEE Access

essential for extracting useful information and removing interference information. In detecting surface defects of automobile pipe joints, processing texture affects the accurate extraction of defect features. If the texture information in images can be decomposed and removed, the success rate of defect feature extraction will be improved. The image $f(x,y)$ can be decomposed and expressed as Equation (7):

$$f(x,y) = u(x,y) + v(x,y) + e(x,y) \quad (7)$$

Where $u(x,y)$ is the defect information, $v(x,y)$ is the texture information, and $e(x,y)$ includes the background structure and noise information.

As shown in Figure 6, it is the first-level wavelet decomposition of the defect image. Where Figure 6 (a) is an approximation coefficient (low-frequency coefficients), Figure 6 (b) is a horizontal detail coefficient, Figure 6 (c) is a vertical detail coefficient, and Figure 6 (d) is a diagonal detail coefficient.

Figure 6 shows that the approximate coefficient contains the primary information of the defect image, and the horizontal detail coefficient contains most of the texture information of the image. Therefore, the processing texture can be removed by zeroing the horizontal detail coefficient. The remaining coefficients still contain noise information.

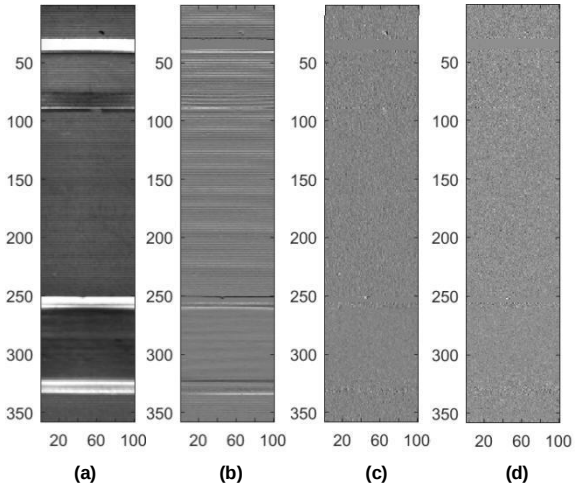


FIGURE 6. Defect image wavelet decomposition map: (a) approximation coefficients, (b) horizontal detail coefficients, (c) vertical detail coefficients, (d) diagonal detail coefficients.

The image $f(x,y)$ can be decomposed by second-level wavelet decomposition to get $A_2(x,y)$, $d_1^h(x,y)$, $d_1^v(x,y)$, $d_1^d(x,y)$, $d_2^h(x,y)$, $d_2^v(x,y)$, $d_2^d(x,y)$. They represent the approximation coefficient of the second-level wavelet decomposition, the first-level horizontal detail coefficient, the first-level vertical detail coefficient, the first-level diagonal detail coefficient, the second-level horizontal detail coefficient, the second-level vertical detail coefficient, and the second-level diagonal detail coefficient. Among, the horizontal detail coefficient represents texture information, which is indicated as Equation (8):

$$v(x,y) = d_1^h(x,y) + d_2^h(x,y) \quad (8)$$

The coefficient matrix $v(x,y)$ is zeroed to remove the texture information. In addition, the low-pass filter is used to denoise $A_2(x,y)$, $d_1^v(x,y)$, $d_1^d(x,y)$, $d_2^v(x,y)$, $d_2^d(x,y)$. The reconstructed image after texture and noise removal is obtained after wavelet reconstruction. The new signal without texture and noise is shown as Equation (9):

$$f_{DWT}(x,y) = u(x,y) + e'(x,y) \quad (9)$$

Where $u(x,y)$ is the defect information, and $e'(x,y)$ is the background structure information after denoising.

Figure 7 is shown the pit defect original image, first-level wavelet decomposition processing image, second-level wavelet decomposition processing image, and third-level wavelet decomposition processing image. According to the image, it can be found that the processing texture of Figure 7(c) is weakened. The defect location is distinct. It illustrates the effectiveness of the algorithm in order to explain further the rationality of the selection of the wavelet decomposition's level. The defect example image is subjected to the first-level wavelet decomposition and the third-level wavelet decomposition. These results are represented in Figures 7(b) and 7(d). After the first-level wavelet decomposition processing, the outline of the defect remained good, but the processed texture is more prominent. The requirement of weakening the texture cannot be completely realized. The defect in the third-level wavelet decomposition image is excessively denoised, affecting the defect contour's definition. The effect of processing texture removal is feeble. Therefore, it can be concluded that the selected second-level wavelet decomposition is better than the others.

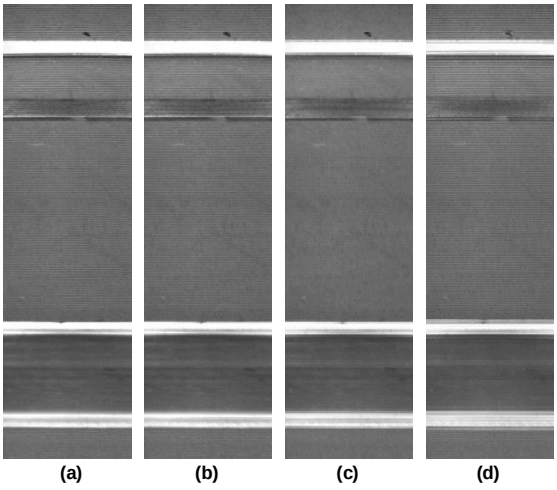


FIGURE 7. Comparison of different decomposition layer processing results: (a) original image, (b) first-level, (c) second-level, (d) third-level.

C. PRELIMINARY DEFECT LOCATION

The Canny edge detection operator is an optimal edge detector [26] that uses the gray image as input to generate the output image. The intensity of the edge position tracked is discontinuous. It is an edge detection technology based on gradient transformation, which has the advantages of high positioning accuracy and can suppress false edges [27]. The

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2022.3178380, IEEE Access

Canny edge detection operator extracts useful structural information from different angles of the image [28]. The amount of data to be processed is reduced dramatically [29].

The surface defect is divided by combining the Canny algorithm and hole filling. Figure 8 shows the examples of defect images. Figure 9 shows the edge feature extraction results of the image reconstructed by the wavelet transform after the Canny edge operator. It can be found that the defect edge is accurately extracted. Figure 10 shows the defect shapes after the hole filling. The defects are effectively segmented.

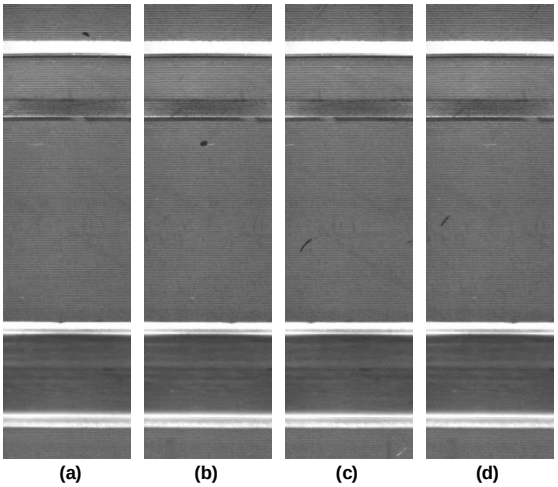


FIGURE 8. Examples of defect images: (a) pit1, (b) pit2, (c) scratch1, (d) scratch2.

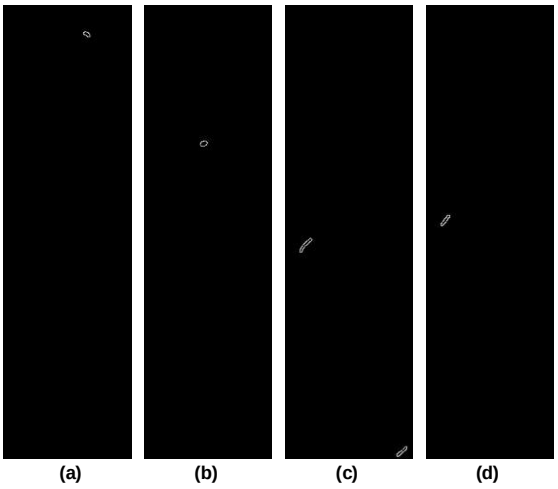


FIGURE 9. Defect image canny edge extraction results: (a) pit1, (b) pit2, (c) scratch1, (d) scratch2.

Figure 11 shows the original image of the scratch defect, the image of the defect segmentation, and the block of the image. The detection area of the part is mainly a convex outer ring, and the concave inner ring is a non-detection area. The portions in the red frame in Figure 11(a) are the defect detection area called the AOI (Area of Interest). The rest is the non-defect detection area. Figure 11(b) shows the defect segmentation result, which can be seen that the shape of the defect in the whole picture is small. The AOI is divided into several pieces of fixed size to improve the classification efficiency of defect detection. Set the width and height of the image block to 68×68 pixels, and divide the AOI into blocks starting from the origin (the upper left corner). The AOI with the size of 198×476 pixels is divided into image blocks of 3 columns by 7 rows. Figure 11(c) is the image block result of the defect detection area.

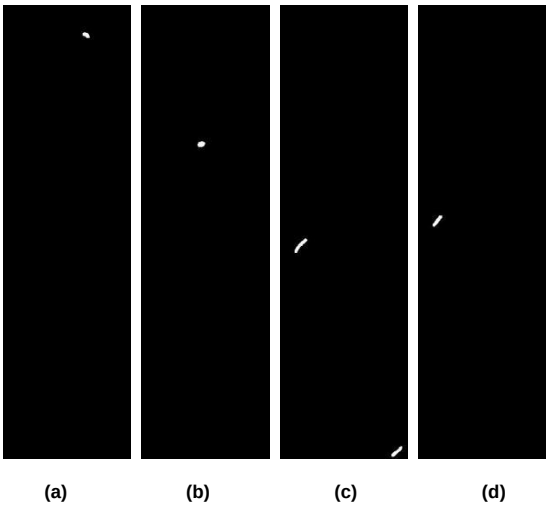


FIGURE 10. Defect segmentation results after hole filling: (a) pit1, (b) pit2, (c) scratch1, (d) scratch2.

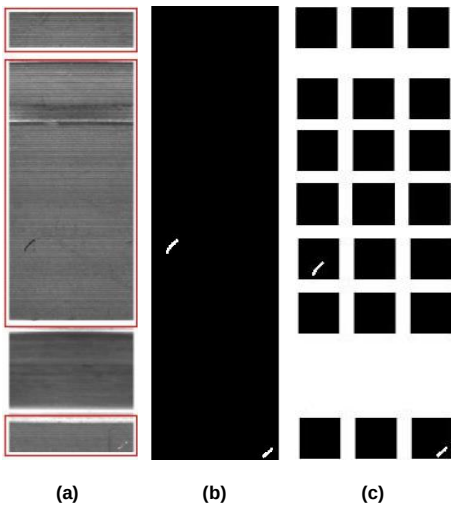


FIGURE 11. Schematic diagrams of image clipping effect: (a) Scratch defect example map original image, (b) Defect feature extraction map, (c) image segmentation map.

The ideal non-defective block can be removed by rough classification to achieve preliminary defect location and improve the calculation speed. The rough classification is mainly calculated based on the pixels of the image. The image block can be roughly separated by detecting the number of non-zero pixels (defective pixels). If the number of times that the non-zero pixel was detected is zero, the

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2022.3178380, IEEE Access

image block is an entirely defect-free entire area. Nevertheless, not all image blocks with non-zero pixels have surface defects. For example, the type is shown in Figure 12. At this time, there are non-zero pixels, but they are qualified parts. These non-zero pixels may be caused by noise or other disturbances that have not been completely removed. So set a threshold of T. When the number of non-zero pixels in the image block is greater than T, the image block is considered defective. During the detection process, T=10.

FIGURE 12. The type of good image block.

D. DEFECT CLASSIFICATION

Decision-level fusion has the characteristics of fault tolerance, openness, low data requirements, and strong analytical capabilities. At the same time, the information loss problem caused by feature-level fusion methods is avoided. Decision-level fusion is used in this paper to classify surface defects in automotive pipe joints. Figure 13 shows the network model of n convolutional neural networks in decision-level fusion. The image blocks shown in figure 11(c) are used as the input of the fusion network. Each network extracts the feature of the input image and recognizes the pattern, and then a classification matrix will be obtained. The matrix is the basis of the image classification. The network of different architectures will get different classification matrices, which have different judgment bases. Then the weighted fusion method is used to fuse the classification matrix. The final classification matrix will be obtained. The matrix is used as the final classification criterion for image classification.

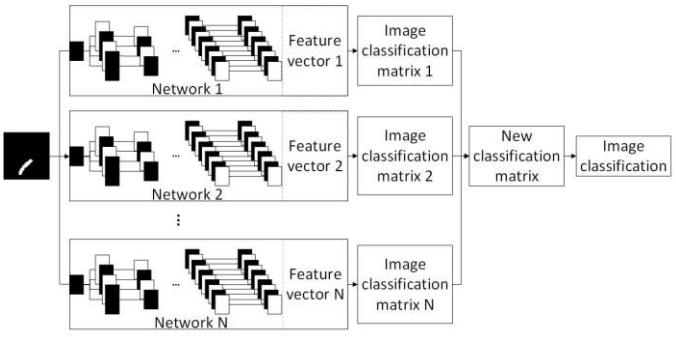


FIGURE 13. Model of convolutional neural network in decision-level fusion.

Decision-level fusion is performed to achieve accurate classification of defect categories. The same picture is input into three pre-trained networks of different structures for processing. Each network will have a Softmax layer to generate a classification vector of $X_i = \{o_1, o_2, \ldots, o_j\}$, where i is the network number in the converged network, and j is the number of categories. Each classification vector generated by the network is given a classification weight vector $C_i = \{c_1, c_2, \ldots, c_j\}$ to achieve higher accuracy. Therefore, the new classification vector is $C_1X_1 + C_2X_2 + \ldots + C_iX_i = \{m_1, m_2, \ldots m_j\}$. According to the new classification vector, the final result of the defect detection classification will be obtained.

This paper designs three convolutional neural networks with different architectures for decision-level fusion to better realize defect type classification, as shown in Figure 14.

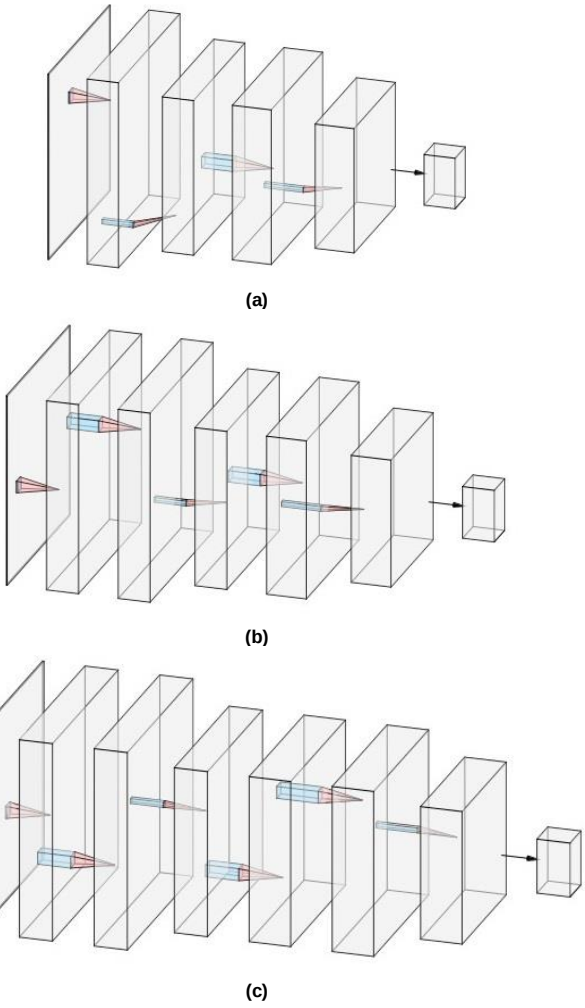


FIGURE 14. Network architecture: (a) Net 1, (b) Net 2, (c) Net 3.

Their network architectures are as follows:

Network 1: A typical net-5 network has an input layer, two convolutional layers, two mean-pooling layers, a fully connected layer, two tanh activation functions, and one output layer, as shown in Figure 14 (a). This network uses fewer convolutional layers and uses mean-pooling to classify defects based on shape features as much as possible.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2022.3178380, IEEE Access

Network 2: The network is designed to consist of an input layer, three convolutional layers, a mean-pooling layer, a max-pooling layer, and three ReLU activation functions, a fully connected layer, and an output layer, as shown in Figure 14 (b). Compared with Network 1, this network increases the number of convolutional layers. It replaces an average pooling layer with a max-pooling layer, which increases the proportion of defect outline detail features in prediction.

Network 3: The network is designed with an input layer, four convolution layers, two max-pooling layers, two ReLU activation functions, a fully connected layer, and an output layer, as shown in Figure 14 (c). This network uses the largest number of convolutional layers and uses max-pooling, which has the most robust feature extraction ability of detail to predict defect classification.

The different depths of these three convolutional neural networks bring different degrees of feature extraction capabilities. So convolutional neural network of decision-level fusion can make predictions through features at different scales.

IV. EXPERIMENTS AND RESULTS

A. EXPERIMENTAL DEVICE AND DATA SET

The surface defect detection system based on machine vision designed for automobile pipe joints is shown in Figure 15 and Table I. The system consists of a light source, a lens, a CCD image sensor, an image acquisition card, a computer image processing system, and a part positioning device. According to the shape and size of the tested part, the CCD image sensor selects the black and white industrial camera of model MV-EM120M. It has a resolution of 1280*960 and a pixel size of 3.75μm*3.75μm. Proportional determination of lens focal length and object distance ratio is confirmed by the area of the CCD pixel and the area of the part. Computer series M3520-MPW2 industrial lens is selected, and the focal length is 35mm. It can be manually adjusted to ensure the object image is as complete as possible.

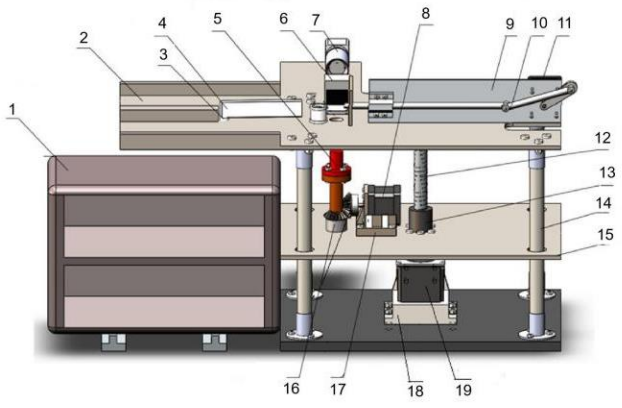


FIGURE 15. **Automobile pipe joint surface defect detection system.**

The experimental samples are from a particular type of automobile pipe joints in an automobile parts production factory. Surface images are collected for 100 parts. A total of 900 images are selected as training samples for the network. A total of 289 samples are selected as testing samples with typical defects, as shown in Table II. No defect, pit, and scratch are experimentally validated for specific classification, as shown in Figure 16. This experiment environment was carried out in the MATLAB R2016a. Iteration 500 is chosen as the number of network training.

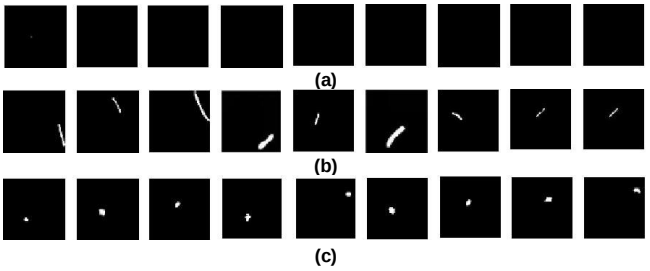


FIGURE 16. **Sample images: (a) no defect; (b) scratch; (c) pit;**

TABLE I
SERIAL NUMBER AND NAME OF COMPONENT IN DETECTION SYSTEM

Number	Component name
1	Control cabinet
2	Mounting plate II
3	Motor
4	Baffle
5	Detecting set
6	Light source
7	CCD
8	Stepping motor I
9	Aluminum plate
10	Push-pull rod
11	Stepping motor II
12	Ball screw
13	Screw nut
14	Pillar
15	Mounting plate I
16	Bevel gear
17	Pedestal I
18	Pedestal II
19	Stepping motor III

TABLE II
THE NUMBER OF DIFFERENT TYPES OF SAMPLE

Type	Number	
	Training sample	Testing sample
No defect	300	100
Pit	300	71
Scratch	300	118
Total	900	289

B. EVALUATION OF IMAGE WAVELET DENOISING

Take the defect image in Figure 6 (a) as an example. In order to verify the denoising and de-texturing effects of the discrete wavelet reconstruction proposed in this paper, the denoised image processing by the four denoising methods of Gaussian filtering, mean filtering, wavelet filtering, and median

filtering are compared. The processing results were obtained as shown in Figure 17.

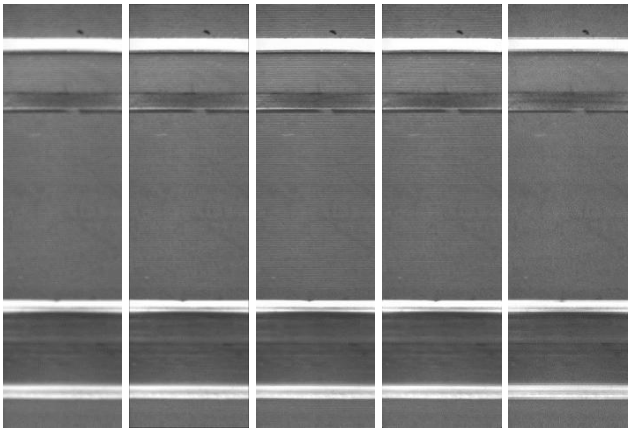


FIGURE 17. The filtered images of different methods (a) Gaussian filtering, (b) Mean filtering, (c) Wavelet filtering, (d) Median filtering, (e) Method of this paper.

Mean Square Error (MSE), Peak signal-to-noise ratio (PSNR) [31], and structural similarity index (SSIM) [32] are selected to evaluate the denoising performance of different denoising methods. The calculation formulas of these evaluation indicators for the image $f(x,y)$ and its processed image $\hat{f}(x,y)$ are as Equations (10) to (12):

$$\mathrm{MSE} = \frac{1}{\mathrm{MN}}\sum_{x=0}^{\mathrm{M}}\sum_{y=0}^{N}(f(x,y)-\hat{f}(x,y))^2 \tag{10}$$

$$\mathrm{PSNR} = 10\,lg\left(\frac{255^2-1}{\mathrm{MSE}}\right) \tag{11}$$

$$\mathrm{SSIM} = \frac{(2\mu_f\mu_{\hat{f}}+\mathrm{C}_1)(2\sigma_{f,\hat{f}}+\mathrm{C}_2)}{(\mu_f{}^2+\mu_{\hat{f}}{}^2+\mathrm{C}_1)(\sigma_f{}^2+\sigma_{\hat{f}}{}^2+\mathrm{C}_2)} \tag{12}$$

Where the number of columns in the image $f(x,y)$ and $\hat{f}(x,y)$ is M, and the number of rows is N, $x \in (0,\mathrm{M}), y \in (0,\mathrm{N})$. μ_f, $\mu_{\hat{f}}$, $\sigma_f{}^2$ and $\sigma_{\hat{f}}{}^2$ are the mean and variance of $f(x,y)$ and $\hat{f}(x,y)$ respectively. $\sigma_{f,\hat{f}}$ is the covariance of $f(x,y)$ and $\hat{f}(x,y)$. C_1 and $\mathrm{C}_2$are the constants.

Since there is no standard image of the automobile pipe joints, the original image that needs to be noise-reduced is used as the standard image. Therefore, it is necessary to consider the physical truth when using the evaluation indexes to evaluate noise reduction. The more significant the difference between the original and denoised images, the better the denoising effect is. In other words, the filtering effect is more apparent. In terms of the MSE, the bigger it is, the better the result is. The PSNR and SSIM need to be as small as possible, indicating a significant difference between the processed image and the original one. The results of different noise reduction algorithms using the evaluation indexes are evaluated in Table III.

It can be seen from Table III that the MSE values are ranked from largest to smallest. They are the method in the paper, Gaussian filtering, mean filtering, median filtering, and wavelet filtering, respectively. The PSNR is the opposite. The rank of SSIM is Gaussian filtering, the method in the paper, mean filtering, median filtering, and wavelet filtering from small to large. It can be found from the above results that the proposed method, Gaussian filtering and mean filtering, perform better in the value of the evaluation index.

TABLE III
THE EVALUATION INDEXES OF DIFFERENT FILTERING ALGORITHMS' NOISE REDUCTION RESULTS

Method	MSE	PSNR	SSIM
Gaussian filtering	120.1401	27.3339	0.4838
Mean filtering	70.8303	29.6286	0.7768
Wavelet filtering	10.9965	37.7183	0.9453
Median filtering	44.2324	31.6734	0.8066
Method of this paper	129.6748	27.0022	0.6552

In order to further illustrate the denoising effect, the gradient image is used to indicate the effect of denoising and removing processing texture. The 3-Dimensional gradient images of the unprocessed image and the denoised image processed by different methods are respectively shown in Figure 18.

It can be seen from Figure 18. that the gaussian filtering, mean filtering, and the method of this paper have the best smoothing effect on the processing texture. It dramatically reduces the interference of the processing texture. Nevertheless, the Gaussian and mean filtering methods also have a sizeable smoothing effect on the defect, making feature extraction difficult. The median filtering method can effectively protect the defect features, but the interference of the processing texture is still severe. Unfortunately, the effect of wavelet filtering is Worst. The method in the paper can effectively remove the processing texture. At the same time, the characteristics of the defect feature are maintained. The contrast between the background position and the defect position is obvious. In summary, the denoising method in the paper has the best effect. It effectively protects the defect features and has a good application effect on solving the detecting problems of automobile pipe joints.

C. EVALUATION OF DEFECT CLASSIFICATION

To judge whether each network has good classification performance. In this paper, the stability of each network is judged by loss function and the training accuracy [33]. The cross-entropy loss function is a kind of smoothing function that applies cross-entropy in information theory to classification problems [34]. Its formula is represented as Equation (13):

$$Loss = -p\,ln(z) - (1-p)\,ln(1-z) \tag{13}$$

Where z is the network node output, and p is the correct output. According to the definition of cross-entropy, it is known that minimizing cross-entropy is equivalent to the minimum observed value and the relative entropy of the estimated value. In other words, it is the Kullback-Leibler divergence of the probability distribution. It is a proxy loss that provides unbiased estimation. The cross-entropy loss function is the most widely used in neural network classification.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2022.3178380, IEEE Access

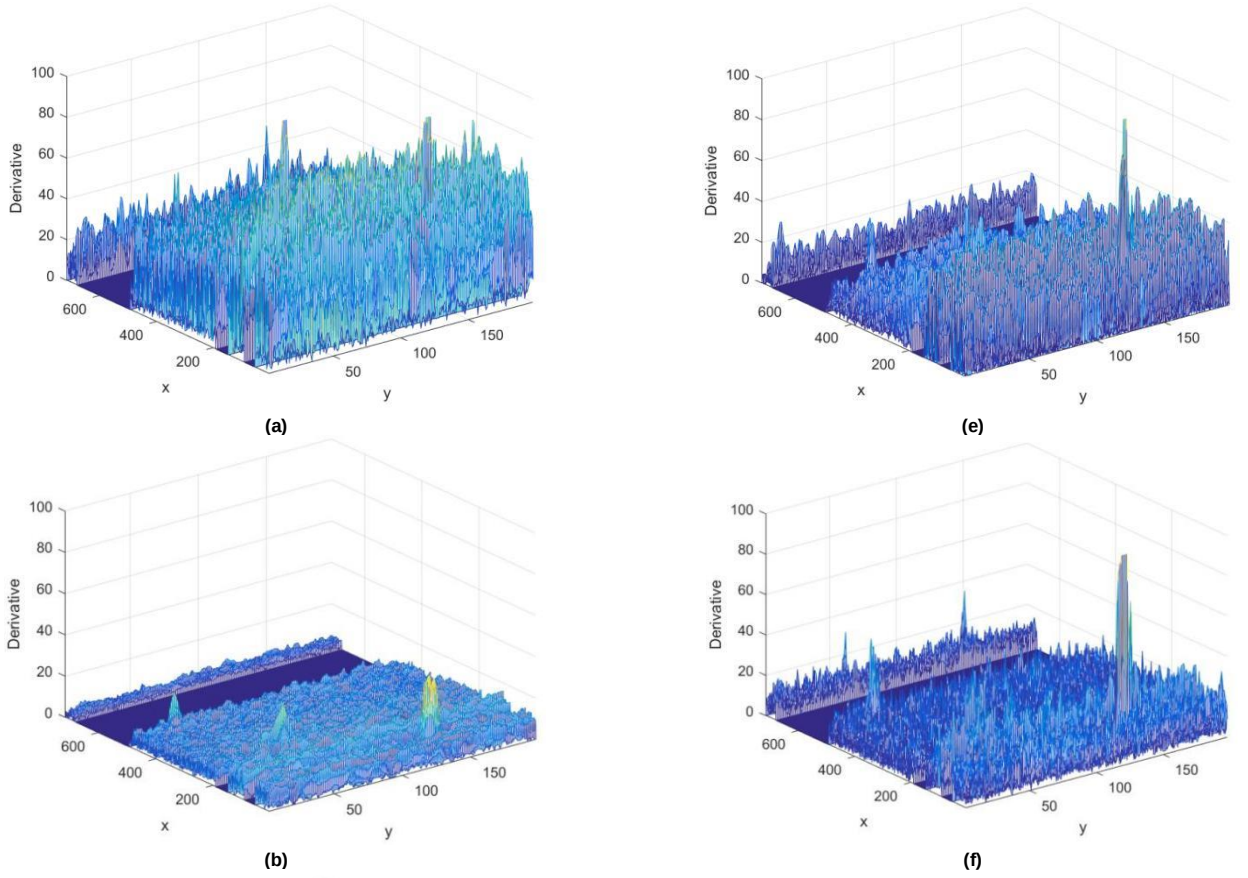


FIGURE 18. Three-dimensional gradient image: (a) Unprocessed, (b) Gaussian filtering, (c) Mean filtering, (d) Wavelet filtering, (e) Median filtering, (f) Method of this paper.

Figure 19 shows the accuracy and loss function of Net 1, Net 2, Net 3, and Fusion Net. It can be found that the loss function value decreases gradually and tends to be stable. The accuracy of network training increases gradually with the number of iterations. In addition, the accuracy of Fusion Net is higher than those of Net 1, Net 2, Net 3, and the loss function of the fusion network has a faster decline rate than others.

The method in this paper is compared with the traditional SVM and single-channel convolutional neural network to predict the results of the test set samples. The results are shown in Table IV.

In order to evaluate the quality of the classification method, the accuracy, precision, and recall rate of the commonly used evaluation indicators of these methods were calculated.

TABLE IV
THE RESULTS OF DIFFERENT CLASSIFICATION METHODS

Method	No defect		Pit		Scratch	
	True	False	True	False	True	False
SVM	86	14	80	38	68	3
Net 1	99	1	94	24	70	1
Net 2	100	0	93	25	71	0
Net 3	100	0	87	31	71	0
Fusion Net	100	0	100	18	71	0

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2022.3178380, IEEE Access

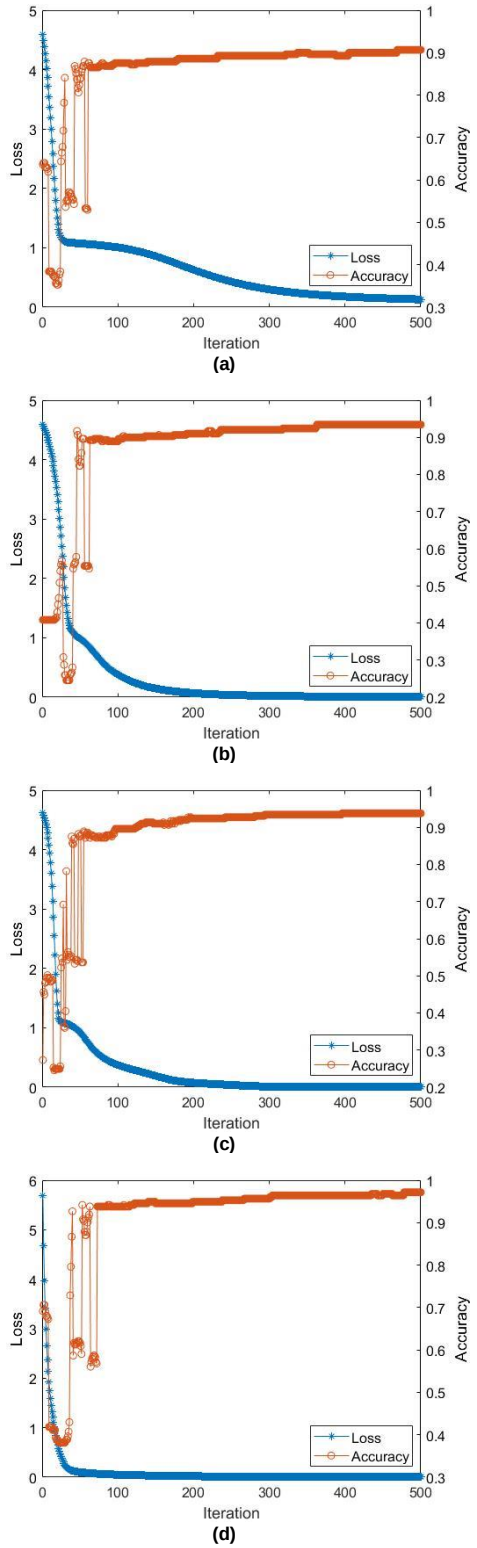


FIGURE 19. Accuracy and loss function curve of networks: (a) Net 1, (b) Net 2, (c) Net 3, (d) Fusion Net.

Accuracy is generally used to assess the global effect of the model. It can be indicated as the proportion of correct predictions to the total number. Precision is the proportion of the number of labels predicted to be the same as the real label. Recall is a measure of coverage, which is the proportion of the number of real labels predicted as real labels to real labels. The results are shown in Table V.

TABLE V
COMPARISON OF EFFECTS OF DIFFERENT CLASSIFICATION METHODS

Method	Accuracy	Precision			Recall		
		No defect	Pit	Scratch	No defect	Pit	Scratch
SVM	83.1%	86.0 %	68.2 %	96.3%	90.7 %	89%	66.3%
Net 1	90.6%	99.0 %	79.7 %	98.4%	100 %	97.1 %	84.7%
Net 2	93.4%	100 %	78.8 %	100%	100 %	100 %	83.8%
Net 3	93.7%	100 %	73.4 %	100%	100 %	100 %	78.8%
Fusion Net	97.1%	100 %	84.9 %	100%	100 %	100 %	89.0%

The P-R(Precision-Recall) Curve's horizontal axis is the Recall, and the vertical axis is the Precision. The fuller the P-R curve of a network model, the better its performance and the higher the classification accuracy. Quantitative description means that the larger the area under the P-R curve, the better the performance of the model and the higher the classification accuracy.

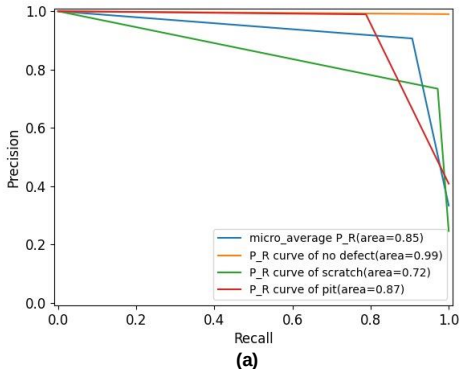
Figure 20 shows the P-R Curve of Net 1, Net 2, Net 3, and Fusion Net. The P-R curves of Net 1, Net 2, and Net 3 for each classification are relatively full, and the area under the curve exceeds 0.7 to achieve a higher value. The areas under the p-r curves of the Fusion network for each kind of sample are larger than any single network, indicating that the fusion network has better classification ability.

The ROC (Receiver Operating Characteristic) Curve's horizontal axis is the FPR (False Positive Rate), and the vertical axis is the TPR (True Positive Rate). Similar to the P-R curve, the shape of the ROC curve can qualitatively describe the performance of the network model. The introduction of AUC (Area under ROC Curve) can be used to quantitatively analyze the model, which refers to the size of the area under the ROC curve, which can be obtained by integrating along the horizontal axis of the ROC curve. The value of AUC. The larger the value of AUC, the better the model's performance.

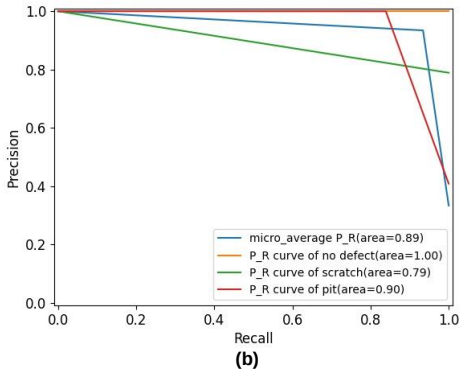
Figure 21 shows the ROC Curve of Net 1, Net 2, Net 3, and Fusion Net. The AUC of the Fusion network for each sample is larger than any single network, which also shows that the fusion network has better classification ability.

From the above table and figures, this paper compares the accuracy, recall, precision, True positive rate, and false positive rate of the traditional classification method SVM, the single convolutional neural classification network (Net 1, Net 2, Net 3), and the Fusion Net. It can be found that, in the categories of no defect, pit, and scratch, the classification accuracy of the SVM classification is mostly around 83%. There is no strong classification advantage. The network with

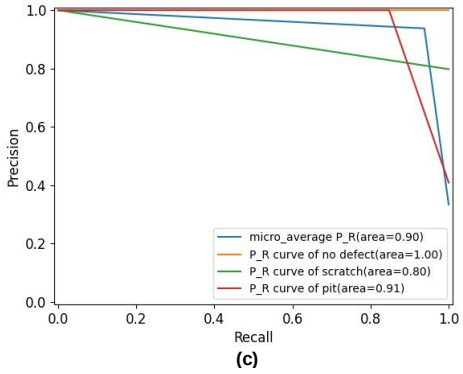
This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2022.3178380, IEEE Access



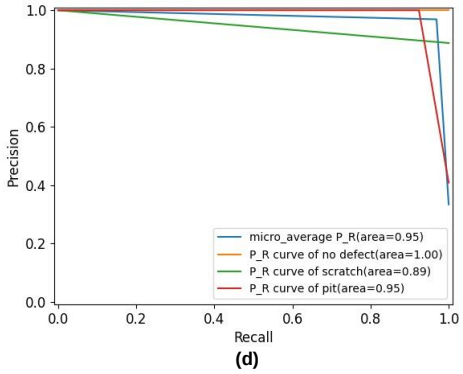
(a)



(b)

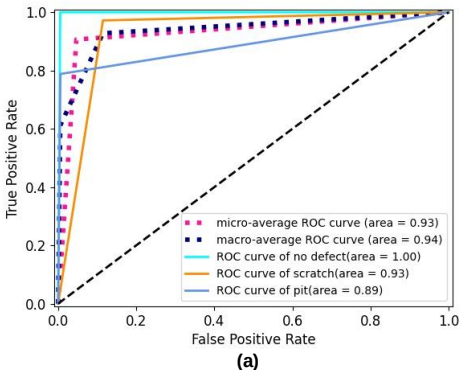


(c)

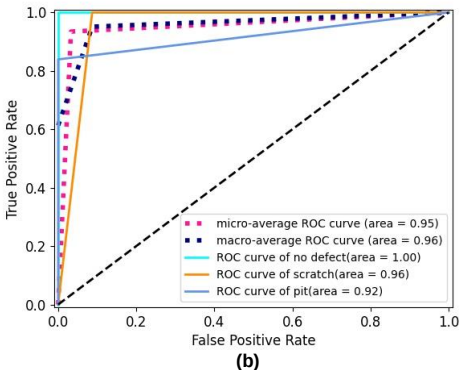


(d)

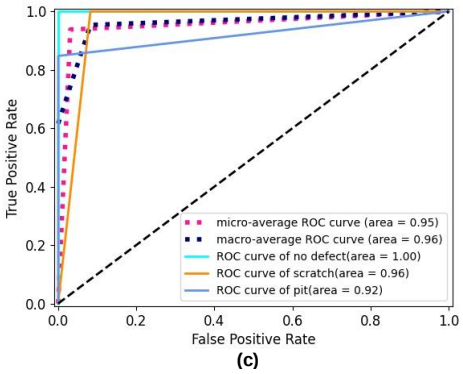
FIGURE 20. Precision-Recall curve of networks: (a) Net 1, (b) Net 2, (c) Net 3, (d) Fusion Net.



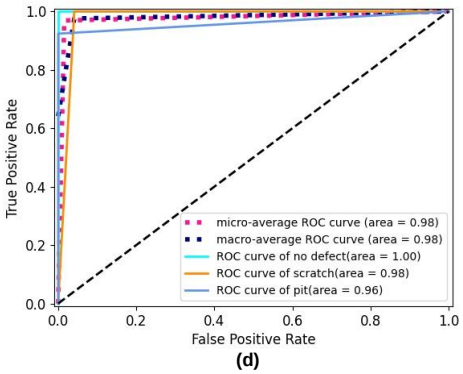
(a)



(b)



(c)



(d)

FIGURE 21. ROC curve of networks: (a) Net 1, (b) Net 2, (c) Net 3, (d) Fusion Net.

different architectures also perform in single deep convolution networks. Net 1 has the highest precision for pit and the highest recall for scratch, and Net 2 and Net 3 perform better on other predictions. Fusion Net combines the advantages to represent the best classification effect, superior to the traditional classification method. In addition, in Fusion Net, the precision of scratch is higher than that of the pit, and the recall of the pit is higher than that of the scratch, indicating that some pits are predicted to be scratches.

V. CONCLUSION

This paper proposes the problem of surface defect detection and classification of automobile pipe joints. The Compensation of light source illumination and wavelet decomposition transformation are combined to improve image quality. The Canny edge operator is combined with hole filling to segment the defect. The three-channel fusion CNN of decision-level is used to classify the segmented defect. The results show that this method effectively eliminates processing texture and noise interference on defect segmentation and can accurately segment defects. The classification result of the three-channel fusion CNN is superior to SVM and the single CNN, which significantly improves the accuracy of defect type identification. The results show that the fusion network can better balance the classification of different sizes and shapes, and its effect is better than single networks'. This method has application value in parts processing and production. According to the classification result, the process parameters can be optimized, and the factors that affect the quality of parts in the machining process can be adjusted in time. This method provides a theoretical basis for surface defect detection and classification system design.

References

[1] C. Xu, L. Li, J. Li and C. Wen, "Surface Defects Detection and Identification of Lithium Battery Pole Piece Based on Multi-Feature Fusion and PSO-SVM," in *IEEE Access*, vol. 9, pp. 85232-85239, 2021, doi: 10.1109/ACCESS.2021.3067641.

[2] L. Xu, S. Lv, Y. Deng and X. Li, "A Weakly Supervised Surface Defect Detection Based on Convolutional Neural Network," in *IEEE Access*, vol. 8, pp. 42285-42296, 2020, doi: 10.1109/ACCESS.2020.2977821.

[3] S. Indolia, A. Goswami, S. Mishra and P. Asopa, "Conceptual understanding of convolutional neural network-a deep learning approach." *Procedia computer science*, vol. 132, pp. 679-688, 2018.

[4] L. Song, W. Lin, Y. Yang, X. Zhu, Q. Guo and J. Xi, "Weak Micro-Scratch Detection Based on Deep Convolutional Neural Network," in *IEEE Access*, vol. 7, pp. 27547-27554, 2019, doi: 10.1109/ACCESS.2019.2894863.

[5] K. Phil, "Matlab deep learning with machine learning, neural networks and artificial intelligence." *Apress*, IN, New York, 2017.

[6] J. Zhong, B. Yang, G. Huang, F. Zhong and Z. Chen, "Remote sensing image fusion with convolutional neural network." *Sensing and Imaging*, vol. 17, no. 1, pp. 1-16, 2016.

[7] X. Jia, Z. Deng, F. Min and D. Liu, "Three-way decisions based feature fusion for Chinese irony detection." *International Journal of Approximate Reasoning*, vol. 113, pp. 324-335, 2019.

[8] Peng, Y., Liao, M., Song, Y., Liu, Z., He, H., Deng, H., & Wang, Y. "FB-CNN: Feature Fusion-Based Bilinear CNN for Classification of Fruit Fly Image," in *IEEE Access*, vol. 8, pp. 3987-3995, 2020, doi: 10.1109/ACCESS.2019.2961767.

[9] X. Xu, L. Wang, X. Chen and B. Liu, "Large group emergency decision-making method with linguistic risk appetites based on criteria mining." *Knowledge-Based Systems*, vol. 182, 2019, Art. no. 104849.

[10] G. Bongiovì, S. Grazioso, and S. Jimenez. "Concept selection of the automated inspection and maintenance test unit for the EU DEMO using a novel fuzzy-based decision support tool." *Fusion Engineering and Design*, vol. 148, 2019, Art. no. 111324.

[11] M. Wu, "Wavelet transform based on Meyer algorithm for image edge and blocking artifact reduction." *Information Sciences*, vol. 474, pp. 125-135, 2019.

[12] H. Zhang, L. Yu and U. Hassler, "An experimental and analytical study of micro-laser line thermography on micro-sized flaws in stitched carbon fiber reinforced polymer composites." *Composites Science and Technology*, vol. 126, pp. 17-26, 2016.

[13] C. Yang, P. Liu, G. Yin, "Defect detection in magnetic tile images based on stationary wavelet transform." *Ndt & E International*, vol. 83, pp. 78-87, 2016.

[14] Y. Zhu and C. Huang "An improved median filtering algorithm for image noise reduction." *Physics Procedia*, vol. 25, pp. 609-616, 2012.

[15] C. Guobin, Z. Sun and L. Zhang, "Road Identification Algorithm for Remote Sensing Images Based on Wavelet Transform and Recursive Operator," in *IEEE Access*, vol. 8, pp. 141824-141837, 2020, doi: 10.1109/ACCESS.2020.3012997.

[16] D. Pankaj, K. Narayanankutty, and D. Govind. "Image Denoising Using Total Variation Wavelet Galerkin Method." *Procedia computer science*, vol. 143, pp. 481-492, 2018.

[17] L. Lin, "An effective denoising method for images contaminated with mixed noise based on adaptive median filtering and wavelet threshold denoising." *Journal of Information Processing Systems*, vol. 14, no. 2, pp. 539-551, 2018.

[18] M. Olfa and K. Nawres, "Ultrasound image denoising using a combination of bilateral filtering and stationary wavelet transform," International Image Processing, *Applications and Systems Conference*, pp. 1-5, 2014, doi: 10.1109/IPAS.2014.7043258.

[19] J. Song, M. Chen, C. Jiang, Y. Huang, Q. Liu and Y. Meng, "Research on Image Denoising Method Based on Wavelet Transform," *2018 37th Chinese Control Conference (CCC)*, pp. 7354-7358, 2018, doi: 10.23919/ChiCC.2018.8482633.

[20] D. Tsai and D. Molina, "Morphology-based defect detection in machined surfaces with circular tool-mark patterns." *Measurement*, vol. 134, pp. 209-217, 2019.

[21] J. Guo, C. Liu, J. Cao and D. Jiang, "Damage identification of wind turbine blades with deep convolutional neural networks." *Renewable Energy*, vol. 174, pp. 122-133, 2021.

[22] D. He, K. Xu and P. Zhou, "Defect detection of hot rolled steels with a new object detection framework called classification priority network." *Computers & Industrial Engineering*, vol. 128, pp. 290-297, 2019.

[23] S. Balasubramani and N. Balaji, "Investigations of vision inspection method for surface defects in image processing techniques-a review." *Advances in Natural and Applied Sciences*, vol. 10, no. 6 SE, pp. 115-120, 2016.

[24] J. Xiao and S. You, "Denoising method of engine surface defect image based on wavelet transform." *Surface technology*, vol. 47, no. 12, pp. 328-333, 2018.

[25] T. Czimmermann, G. Ciuti, M. Milazzo, M. Chiurazzi, S. Roccella, C. Oddo and P. Dario, "Visual-based defect detection and classification approaches for industrial applications — a survey." *Sensors*, vol. 20, no. 5, Mar. 2020, Art. no. 1459.

[26] J. Bezdek, R. Chandrasekhar and Y. Attikouzel, "A geometric approach to edge detection," in *IEEE Transactions on Fuzzy Systems*, vol. 6, no. 1, pp. 52-75, Feb. 1998, doi: 10.1109/91.660808.

[27] Ş. Öztürk and A. Bayram, "Comparison of HOG, MSER, SIFT, FAST, LBP and CANNY features for cell detection in histopathological images." *Helix*, vol. 8, no. 3, pp. 3321-3325, 2018.

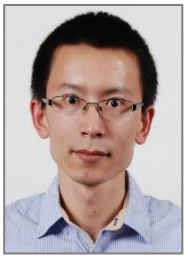
[28] F. Al-Hafiz, S. Al-Megren and H. Kurdi, "Red blood cell segmentation by thresholding and Canny detector." *Procedia Computer Science*, vol. 141, pp. 327-334, 2018.

[29] C. Raju, G. Raju and V. Gottumukkala, "Comparative Studies on Cell Segmentation by Fuzzy Logic and Canny Edge." *ijarece*, vol. 5, no. 5, pp. 1555-1559, 2016.
[30] A. Tanchenko, "Visual-PSNR measure of image quality." *Journal of Visual Communication and Image Representation*, vol. 25, no. 5, pp. 874-878, 2014.
[31] Z. Wang, A. Bovik, H. Sheikh and E. Simoncelli, "Image quality assessment: from error visibility to structural similarity," in *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600-612, April 2004, doi: 10.1109/TIP.2003.819861.
[32] Y. Li, H. Huang, Q. Xie, L. Yao, and Q. Chen, "Research on a surface defect detection algorithm based on MobileNet-SSD." *Applied Sciences*, vol. 8, no. 9, 2018, Art. no. 1678.
[33] H. Nhat-Duc, Q. Nguyen and V. Tran, "Automatic recognition of asphalt pavement cracks using metaheuristic optimized edge detection algorithms and convolution neural network." *Automation in Construction*, vol. 94, pp. 203-213, 2018.

Zeqing Yang received the Ph.D. degree from Hebei University of Technology, China. She is currently an associate Professor with the Department of Instrumentation engineering, Hebei University of Technology. Her current research interests include online detection and error compensation of CNC equipment, digital integrated measurement and control and digital twin operation and maintenance monitoring of complex equipment, visual inspection and pattern recognition.

Mingxuan Zhang received the B.E. degree from Jilin University, Changchun, China, in 2020. He is currently pursuing the M.E. degree with School of Mechanical Engineering, Hebei University of Technology, Tianjin, China. His current research interests include machine learning and computer vision.

Chao Li received the M.S. from Hebei University of Technology, Tianjin, China, in 2020. Her research interests include machine learning and computer vision.



Zhaozong Meng received the B.S. and M.S. degrees in measurement and control technology and instrument from Sichuan University, Chengdu, China in 2006, and Beihang Univerisity, Beijing, China in 2009, respectively, and his Ph.D. degree in computer science from University of Huddersfield, West Yorkshire, UK, in 2014. He worked as a research associate with the University of Manchester, UK, and research fellow with university of Southampton, UK in 2014-2016 and 2016-2018, respectively. Since 2018, he has been a lecturer with Hebei University of Technology, Tianjin, China. Since 2020, he has been promoted an associate professor. His research interests include novel sensing techniques, wearable devices and body area network, industrial IoT and cyber-physical systems.

Yue Li received the B.E. and M.S. from Hebei University of Technology, Tianjin, China, in 2018 and 2021. Her research interests include machine learning and online detection of CNC equipment.

Yingshu Chen received the Ph.D. degree from Hebei University of Technology, Tianjin, China. She is currently a senior lab master of the Department of Instrumentation engineering, Hebei University of Technology. Her research interests include: online inspection and error compensation of CNC machining, measuring and embedded control technology, intelligent manufacturing and intelligent equipment technology. As a main researcher, she has participated in several research and development projects.

LIBING LIU is currently a Full Professor with the School of Measurement and Control Technology and Instrument, Hebei University of Technology, China. She has 168 scientific papers and guides about 200 masters and doctors. Her research interests are mainly in CNC Manufacturing and CNC Equipment Technology, Intelligent Manufacturing and Intelligent Equipment Technology, Digital Integrated Monitoring and Control Technology, Intelligent Perception Technology of Complex System, Digital Twin of Complex Products and Big Data Technology.