*International Journal of Innovative Computing, Information and Control*
Volume **20**, Number **3**, June **2024**

ICIC International ©2024 ISSN 1349-4198
pp. **677–692**

# BM3D WITH ADAPTIVE SEARCH NEIGHBORHOOD AND APPLICATION TO RAIL SURFACE DEFECT DETECTION

Tao Lin[1,2], Yingkun Hou[1,2,*], Hao Hou[3], Zekun Lv[4] and Xiaoya Dai[1,2]

[1]School of Information Science and Engineering
Guilin University of Technology
No. 319, Yanshan Street, Guilin 541006, P. R. China
{ 2531732885; 2281991672 }@qq.com

[2]School of Information Science and Technology
Taishan University
No. 525, Dongyue Street, Taian 271000, P. R. China
*Corresponding author: ykhou@tsu.edu.cn

[3]School of Computer Science and Engineering
Nanjing University of Science and Technology
No. 200, Xiaolingwei Street, Nanjing 210094, P. R. China
haohou0611@gmail.com

[4]School of Big Data and Digital Media
Chongqing Telecommunication Polytechnic College
No. 5, Xuefu Avenue, Chongqing 402247, P. R. China
1310671265@qq.com

Received July 2023; revised November 2023

Abstract. *Rail surface defect images are often affected by noise, which degrades image quality and hinders the detection of helpful information. To improve the accuracy of rail surface defect detection, denoising the rail surface defect image has become essential work. The BM3D algorithm is an excellent image-denoising method, and it still shows a strong denoising performance in the image-denoising of rail surface defects. However, the BM3D algorithm ignores the relationship between the internal morphological components of the reference block and the search neighborhood when matching similar blocks, resulting in insufficient block-matching accuracy, which may affect the performance of the denoising methods. To alleviate this problem, an improved BM3D image-denoising algorithm based on an adaptive search neighborhood is proposed. Different from the BM3D block matching process, we add the step of analyzing the morphological components of the reference block and adaptively assign the optimal search neighborhood range according to the morphological components of the reference block. The experimental results based on the dataset show that compared with the baseline method BM3D and other state-of-the-art methods, we have achieved higher PSNR and SSIM values and improved the accuracy of defect detection. Our study has important practical implications. Through noise reduction processing, we successfully eliminated noise interference in railway defect images, thereby improving the visibility of defective areas and helping to improve the safety and efficiency of railway transportation.*

**Keywords:** Image-denoising, BM3D, Morphological component analysis, Adaptive search neighborhood

1. **Introduction.** High-quality images are essential for accurate rail surface defect detection. However, the acquisition process can be influenced by factors such as equipment

---

DOI: 10.24507/ijicic.20.03.677

quality, acquisition methods, and signal interference, resulting in noise that can significantly impact detection accuracy [1]. Hence, it becomes imperative to employ noise reduction techniques specifically tailored for rail surface defect detection to mitigate the adverse effects of noise and enhance the overall accuracy of the detection system [2, 3, 4].

Although many research methods have emerged for the image-denoising problem of railway surface defects, it is common to use traditional image-denoising algorithms. Traditional denoising algorithms achieve better denoising effects in railway surface defect images. Traditional image-denoising algorithms can be divided into three categories according to different processing domains: spatial domain denoising algorithms [5], transform domain denoising methods [6] and hybrid domain denoising algorithms [7].

The spatial domain denoising algorithm is used to process the target pixel and its neighborhood features in the image. The main algorithms include mean filter [8], median filter [9], and bilateral filter [10]. Mean filters are effective for reducing low noise but perform poorly in strong noise. The median filter is effective in removing random noise, but it may not handle Gaussian noise well and could lead to image blurring. Bilateral filters can remove noise while preserving details, but they may introduce distortion in high-frequency image regions. One improved denoising algorithm is based on non-local mean filtering (NLM) [11], which utilizes redundant information in the image to denoise. However, it may encounter challenges when processing non-Gaussian noise and maintaining sharpness in image edges. Although spatial domain denoising methods have their advantages and disadvantages, in general, the spatial domain image-denoising algorithm has shown good results in the task of denoising rail defect images.

The transform domain image-denoising algorithm transforms the image into the transform domain for processing through various transformations. Common methods include discrete cosine transform (DCT) denoising algorithm [12], principal component analysis (PCA) denoising method and wavelet denoising algorithm [13]. Although these algorithms perform well in removing a large amount of noise in rail surface defect images, ringing effects caused by the Gibbs phenomenon may appear in some image defects [14]. These distorted signals seriously affect the subjective visual quality.

The hybrid domain denoising algorithm combines the advantages of spatial and transform domains, and in the denoising of railway track defect images, its denoising competitiveness is more prominent. The most typical method is BM3D [15], which uses non-local ideas for searching similar image blocks in the spatial domain and employs orthogonal transformation to enhance sparsity and collaborative filtering for better denoising. It is considered one of the best image-denoising filters. Due to BM3D's powerful denoising ability, improved algorithms based on it have been a research hotspot. Dabov et al. [16] proposed BM3D-SAPCA, using non-local image modelling, PCA, and local shape-adaptive anisotropic estimation, achieving breakthroughs in retaining image details and introducing fewer artifacts. [17] improved BM3D using SSIM optimization, enhancing the Wiener filter by maximizing structural similarity (SSIM) between real and estimated images. Hou and Shen [18] proposed SA-BM1-3D, an adaptive block-matching transform domain filtering algorithm based on morphology and size. It adjusts block sizes in different image regions to improve matching accuracy and enhance BM3D's performance. Despite BM3D's good denoising performance, it may produce thin oscillation lines and blocking artifacts, especially when handling complex detail information.

In addition to traditional image-denoising algorithms, deep learning-based methods are gradually applied to rail defect image-denoising [19]. DnCNN is a representative image-denoising algorithm based on deep learning [20]. Introducing residual learning and batch normalization technology effectively solves the gradient diffusion effect between network layer structures and significantly improves the denoising performance. Yang and Sun [21] proposed a convolutional neural network-based method for collaborative filtering in the transform (BM3D-Net). It proposes a deep learning-based approach for denoising images that combines the strengths of convolutional neural networks (CNNs) and the BM3D algorithm. The proposed method unfolds the computation pipeline of the BM3D algorithm into a CNN structure by modelling the block matching stage with “extraction” and “aggregation” layers, thus effectively integrating deep learning with the BM3D method idea. Deep learning has achieved remarkable performance in image-denoising tasks. However, these methods heavily rely on convolutional neural networks (CNNs) to extract features from input images [22]. This reliance results in significant hardware resource requirements for both training and inference. GPU acceleration can significantly reduce the training time of CNNs, but deep learning-based denoising algorithms also require large, diverse digital image datasets for training, which limits their generalization ability and transferability.

Although these methods have achieved good denoising effects on rail defect images, given the excellent denoising performance and ease of operation of the BM3D algorithm on rail surface defect images, we choose the BM3D algorithm to improve and propose a method based on adaptive search neighborhood BM3D image-denoising algorithm.

## 2. **Proposed Algorithm.**

The BM3D algorithm matches similar blocks by computing the Euclidean distance between the block reference block and the block to be matched in a fixed neighborhood. Although this approach increases the efficiency of BM3D in matching similar blocks, this approach ignores the relationship between the internal morphological components of the reference block and the search neighborhood, which leads to insufficient block matching accuracy and makes the image detail texture region too smooth. The algorithm proposed in this paper is improved based on the BM3D algorithm. We add the step of analyzing the morphological components of the reference block in the process of matching similar blocks in the algorithm and adaptively allocate the optimal search neighborhood range by judging the morphological components of the reference block. Finally, the matched similar 2D image blocks are stacked into 3D block groups to enhance the sparsity of image blocks in the transform domain and remove noise in collaborative filtering. Similar to the BM3D algorithm, the algorithm in this paper is also divided into two stages: the first stage is the basic estimation stage, which mainly performs preliminary

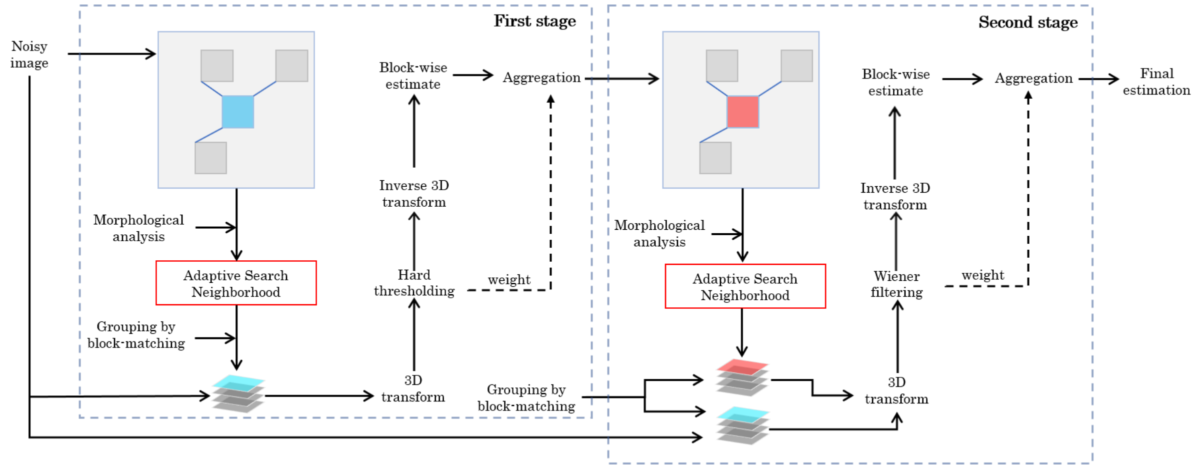


FIGURE 1. The overall framework of the proposed image-denoising algorithm, the red box is the adaptive search neighborhood step, which is the core of the improved algorithm

processing and noise reduction on the noisy image; the second is the final estimation stage, which focuses on the finer noise reduction of the initially processed images.

### 2.1. **First stage of denoising.**

2.1.1. *Block-grouping.* Given a defect image $\boldsymbol{I_r} \in \mathbb{Z}^{h\times w}$ on the surface of a railway track, a target block of size $L \times L$ is selected as the reference block $P$. Next, we perform a multi-scale analysis on the pixels inside the defect image $\boldsymbol{I_r}$ to determine the appropriate analysis parameter $m$. Meanwhile, we mark the pixel numbers in the reference block $P$ in order, then accumulate all the pixels in the reference block and calculate the average value $M$.

$$M = \frac{1}{L^2}\sum_{i=1}^{L^2} x(i) \tag{1}$$

where $x(i)$ represents the pixel values in the reference block, arranged by the tag sequence number.

The result of this average value is used to calculate the standard deviation $std$ of the pixels in the reference block, which is given by

$$std = \sqrt{\frac{1}{L^2-1}\sum_{i=1}^{L^2}\left(x(i)-M\right)^2} \tag{2}$$

By combining the standard deviation of the pixel with the previous multi-scale analysis result from $m$ in the reference block, we obtain the shape component analysis result in $M_{al}$ of the reference block.

$$M_{al} = m \cdot std \tag{3}$$

By comparing the results of the morphology analysis with preset thresholds $\Psi_1$ and $\Psi_2$, we classify the morphology types into reference block smooth morphology, contour morphology, and reference block texture morphology. Specifically, when $M_{al}$ is less than $\Psi_1$, we define it as a reference block with a smooth shape; when $M_{al}$ is greater than $\Psi_1$ but less than $\Psi_2$, we classify it as a contour shape; and when $M_{al}$ is greater than $\Psi_2$, it is classified as a reference block with a texture form. For reference blocks with morphology contour components, we choose a smaller search neighborhood range. For reference blocks with texture components, we select a medium-sized search neighborhood range. And for reference blocks with smooth morphology components, we choose a wider range of search neighborhoods.

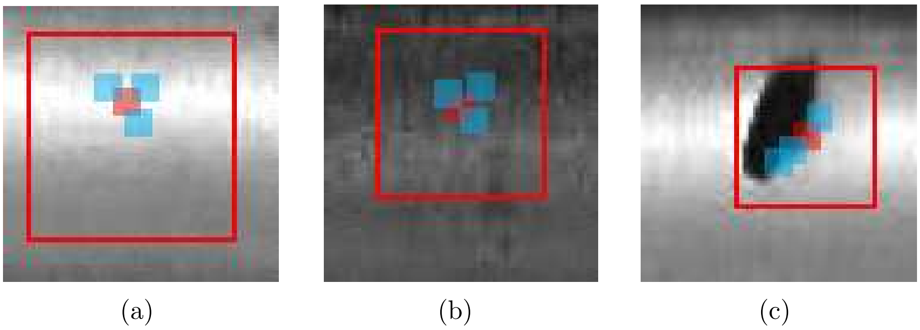

(a) (b) (c)

FIGURE 2. Illustration of image blocks with three morphological components, i.e., with (a) smooth, (b) texture, and (c) contour components, respectively. The red box represents the range of the search neighborhood.

Given the reference block $P$ as the center and using a sliding step size of $S^{ht}$, we search for a matching similar block $Q$ within the adaptive search neighborhood. The similarity between blocks $P$ and $Q$ can be measured by the Euclidean distance $d$, where $d$ is as the following:

$$d(P,Q) = \frac{\|Y_P - Y_Q\|^2}{L^2} \tag{4}$$

The smaller the Euclidean distance between two blocks, the higher their similarity. However, the Euclidean distance between similar blocks should be greater than zero but less than a threshold value $\tau$. Therefore, it is possible to determine whether a matching block is a similar block based on the Euclidean distance between the defined matching block and the reference block. When the distance between the matching block and the reference block is less than a fixed threshold value $\tau$, the matching block is considered a similar block. The set of similar blocks $K$ can be mathematically expressed as follows:

$$K = \{Q : d(P,Q) \le \tau\} \tag{5}$$

Next, stack these similar blocks into a 3D array, and denote it as block group $G_R$.

2.1.2. *Collaborative hard-threshold filtering.* The 3D transformation used in the BM3D algorithm is a separable 3D transformation. Firstly, the 3D image block group $G_R$ is subjected to a separable orthogonal transformation, which involves performing a bior1.5 wavelet transformation on the 2D matrix image blocks and a Haar transform on the third dimension of the 3D image block group. After this transformation, a hard thresholding filter is applied to setting some of the high-frequency coefficients of the transformed image to zero:

$$\gamma(x) = \begin{cases} 0, & \text{if } |x| \le \lambda_{3d} \\ x, & \text{otherwise} \end{cases} \tag{6}$$

where $\gamma$ is the coefficient hard thresholding results, $x$ are the transformed coefficients, and $\lambda_{3d}$ is the hard thresholding parameter.

The number of non-zero elements $N_z$ is calculated in the subband obtained after hard thresholding. Finally, an inverse transform is applied to reconstructing a new 3D matrix image block group $\hat{G}_R$ from the processed data. The mathematical formulation of this process can be expressed as follows:

$$\hat{G}_R = T_{3d}^{-1}\left(r\left(T_{3d}(G_R)\right)\right) \tag{7}$$

where $r$ represents a hard threshold operation, $T_{3d}$ and $T_{3d}^{-1}$ represent 3D transformation and 3D inverse transformation, respectively.

2.1.3. *Aggregation.* Restore each image patch within $\hat{G}_R$ to its initial position. Since a single pixel may have multiple image patches in common, the weight $w_a$ is estimated using the reciprocal of the non-zero number $N_z$:

$$w_a = \begin{cases} \dfrac{1}{N_z}, & \text{if } N_z \ge 1 \\ 1, & \text{otherwise} \end{cases} \tag{8}$$

Finally, a base estimate of the image is computed by taking a weighted average of all obtained overlapping block-wise estimates.

2.2. **Second stage of denoising.** The basic estimation phase has eliminated a considerable amount of noise; however, it has also resulted in the loss of some fine details. Therefore, the ultimate objective of denoising is to precisely eliminate noise while restoring certain intricate features of an image. Similar to the initial estimation stage, this procedure can be categorized into three phases – block matching, collaborative filtering, and aggregation. In contrast to the utilization of a hard-thresholded filter in the first stage, the collaborative filter is substituted with a Wiener filter.

2.2.1. *Block-grouping.* During our block-matching process on the preliminary denoised image of the first stage, we still perform morphological composition analysis on the reference block. Since part of the image information is bound to be smoothed during the denoising process of the first stage, to improve the accuracy of block matching, the search neighborhoods assigned to different morphological reference blocks are slightly smaller than in the previous stage. Unlike the basic estimation stage, where similar blocks are searched for on the basic estimation image to form a block group $G_a^{wie}$, a block group $G_b^{wie}$ is also formed on the noisy image.

2.2.2. *Collaborative Wiener filtering.* Perform separable 3D orthogonal transforms on the two block groups obtained. Specifically, apply the DCT cosine transform to the image matrix in two dimensions, followed by the Haar wavelet transform along the third dimension. This results in two transformed 3D block groups, denoted as $T_a^{wie}$ and $T_b^{wie}$. After transformation, the coefficients of a 3D matrix group are subjected to Wiener filtering for coefficient shrinkage.

$$W^{wie} = \frac{\left|T_a^{wie}\right|^2}{\left|T_a^{wie}\right|^2 + \sigma^2} \tag{9}$$

where $W^{wie}$ is the Wiener filter coefficient and $\sigma$ is the noise coefficient.

The pixel weights of overlapping image blocks need to be calculated, just like in the basic estimation phase. The weighting coefficients after 3D collaborative filtering depend on the Wiener filtering coefficient and the magnitude of the noise variance. The weight coefficient $w^{wie}$ can be expressed as

$$w^{wie} = \sigma^{-2} \left\|W^{wie}\right\|_2^{-2} \tag{10}$$

2.2.3. *Aggregation.* The final estimated value of pixels is obtained by taking the weighted average of pixels from all overlapping image blocks.

## 3. Experiments and Results.

As the surface of the railroad track is worn, abraded or smoothed due to long-term use, defects such as cracks, abrasions, flaws, rust and dents are present in railroad track defect images [23]. In general, the ratio of normal and defective photographs acquired may be in the range of 70 : 30 to 90 : 10 depending on the method of data acquisition, location and time of shooting [24]. In this study, the image noise of rail surface defects follows a Gaussian noise model. To better restore the various defects receiving noise interference, Gaussian white noise is superimposed on the original image as needed, and then various denoising algorithms are applied respectively. To ensure the stability and validity of experimental outcomes, all code executions were carried out on a high-performance computer that incorporates an Intel(R) Iris(R) Xe Graphics processor and an 11th Gen Intel(R) Core(TM) i5-1135G7 graphics card – chosen specifically for compatibility with the experimental setup.

### 3.1. Datasets.

The RSDDs dataset was employed in the experiment, which was collected by a visual inspection system for steel rail surfaces based on a hierarchical extractor. During the acquisition process, the noise is effectively suppressed through the processing of the fine extractor [25].

The RSDDs dataset consists of two subclass datasets: the first is the Type I RSDDs dataset, which contains 67 challenging images shot from high-speed rail tracks. To prevent overfitting, we need to perform data enhancement on the first subclass of RSDDs dataset. After brightness adjustment, we obtained 100 images. The second is a Type II RSDDs dataset captured from regular and heavy transport tracks, which includes 128 challenging images. Similarly, to prevent overfitting, we still need to perform data enhancement on the second subclass of the RSDDs dataset. After scaling and cropping, we obtained 200 images. Although there are slight differences between the two datasets, every image in both datasets has a defect. The defect location is marked by professional observers in the field of track surface inspection. The images in both datasets have a complex background and numerous details.

### 3.2. Evaluation metrics for experimental results.

To better evaluate the denoising performance of algorithms on rail surface defect images, this experiment employs peak signal-to-noise ratio and structural similarity as objective analysis indicators for evaluating image quality.

3.2.1. *Peak signal-to-noise ratio.* The peak signal-to-noise ratio (PSNR) is derived by comparing the relationship between the noisy image and the denoised image, and $I(a,b)$ and $F(a,b)$ denote the pixel values of the noisy image and the denoised image, respectively. $mse$ is the loss function between the noisy image and the denoised image, which can be derived from the following equation:

$$mse = \frac{1}{HW}\sum_{a=1}^{H}\sum_{b=1}^{W}\left[I(a,b) - F(a,b)\right]^2 \tag{11}$$

where $H$ and $W$ represent the length and width of the noisy and denoised images, respectively.

The result of mean squared error $mse$ can be used to obtain the PSNR.

$$PSNR = 10\lg\left(\frac{MAX_a^2}{mse}\right) \tag{12}$$

where $MAX_a$ represents the maximum pixel value of the image $F(a,b)$.

3.2.2. *Structural similarity.* The structural similarity index (SSIM) is a metric used to measure the similarity between two images:

$$SSIM(I,F) = \frac{(2u_p u_j + c_1)(2\sigma_{pj} + c_2)}{\left(u_p^2 + u_j^2 + c_1\right)\left(\sigma_p^2 + \sigma_j^2 + c_2\right)} \tag{13}$$

The mean values of images $I$ and $F$ are represented by $u_p$ and $u_j$, respectively. The variances of the two images are denoted as $\sigma_p^2$ and $\sigma_j^2$, while their covariance is given by $\sigma_{pj}$. The constants $c_1$ and $c_2$ are used to stabilize the relationship between the two images. When there are similarities between the images, higher values of both PSNR and SSIM indicate superior denoising effects.

### 3.3. Experimental results and analysis.

3.3.1. *Comparative analysis of denoising algorithms across varied noise intensities.* To better present the effect of image-denoising, we selected the rail surface defect images I-rail-2 and II-rail-2 as clean images in the datasets of two models of RSDDs, and added Gaussian white noise with standard deviation $\sigma$ of 5, 10, 25, and 50. Noise standard deviation $\sigma$ measures the intensity of noise in the image. The larger the standard deviation, the higher the noise intensity of the image and the more obvious the noise interference; the smaller the standard deviation, the lower the noise intensity of the image and the less noise interference. To better restore the characteristics of railway defect images that are easily affected by medium and low noise in reality, we set the noise standard deviation to 5, 10, 25 and 50 Gaussian white noise. Then the mean filter, median filter, NLM denoising algorithm, DnCNN denoising algorithm, BM3D algorithm, and BM3D algorithm based on adaptive search neighborhood were used to conduct denoising control experiments, and the values of PSNR and SSIM were calculated.

In the experimental comparison of I-rail-2 denoising, the improved denoising algorithm in this paper achieves better PSNR values than BM3D. In terms of recovering the image structure, the SSIM value of the algorithm in this paper is equal to that of the BM3D algorithm, which well protects the structural information of the image. In the low-noise case with $\sigma = 5$, the algorithm in this paper has a higher PSNR value and an optimal SSIM value. Compared with other algorithms, the algorithm in this paper has a competitive advantage. In addition to the DnCNN method, this paper achieves a more desirable denoising effect than other algorithms.

Table 1. Comparison of PSNR (dB) and SSIM of different denoising models for I-rail-2 image-denoising

| $\sigma$ | 5 | | 10 | | 25 | | 50 | |
|---|---|---|---|---|---|---|---|---|
| Metric | PSNR↑ | SSIM↑ | PSNR | SSIM | PSNR | SSIM | PSNR | SSIM |
| AverFilter | 35.87 | 0.9241 | 33.98 | 0.8789 | 28.64 | 0.6853 | 23.29 | 0.4199 |
| MedianFilter | 35.34 | 0.6932 | 32.92 | 0.8789 | 27.18 | 0.6082 | 21.73 | 0.3397 |
| NLM | 38.05 | 0.9448 | 34.70 | 0.8955 | 30.39 | 0.7981 | 30.39 | 0.6841 |
| DnCNN | 35.53 | 0.9050 | **36.23** | **0.9247** | **33.31** | **0.8813** | **30.75** | **0.8345** |
| BM3D | 39.27 | 0.9581 | 36.00 | 0.9194 | 33.02 | 0.8717 | 30.38 | 0.8235 |
| Ours | **39.31** | **0.9583** | 36.07 | 0.9199 | 33.07 | 0.8725 | 30.42 | 0.8260 |

In the comparison of II-rail-2 denoising experiments, the proposed algorithm achieves the best PSNR and SSIM values compared with other algorithms. In comparison with the BM3D algorithm, the PSNR is 0.12 dB higher on average, and the SSIM is the same as that of the BM3D denoising algorithm. In comparison with the deep learning method, the denoising effect of DnCNN is much worse than the algorithm of this paper in the low noise case of $\sigma = 5$. At $\sigma = 10, 25, 50$, the PSNR values of this algorithm are, on average, 0.24 dB higher than those of DnCNN.

Table 2. Comparison of PSNR (dB) and SSIM of different denoising models for II-rail-2 image-denoising

| $\sigma$ | 5 | | 10 | | 25 | | 50 | |
|---|---|---|---|---|---|---|---|---|
| Metric | PSNR↑ | SSIM↑ | PSNR↑ | SSIM↑ | PSNR↑ | SSIM↑ | PSNR↑ | SSIM↑ |
| AverFilter | 35.60 | 0.9456 | 33.53 | 0.9039 | 27.95 | 0.7342 | 22.57 | 0.4907 |
| MedianFilter | 36.85 | 0.9494 | 32.88 | 0.8832 | 26.61 | 0.6684 | 21.14 | 0.4106 |
| NLM | 39.30 | 0.9711 | 35.16 | 0.9301 | 28.40 | 0.7911 | 24.55 | 0.5974 |
| DnCNN | 38.02 | 0.9642 | 38.11 | 0.9668 | 33.99 | 0.9247 | 30.32 | 0.8490 |
| BM3D | 41.56 | 0.9830 | 38.20 | 0.9679 | 34.09 | 0.9269 | 30.45 | 0.8654 |
| Ours | **41.65** | **0.9833** | **38.33** | **0.9688** | **34.23** | **0.9298** | **30.58** | **0.8693** |

3.3.2. *Comparative analysis of denoising algorithms on various surface defect images.* To compare the denoising effects of different algorithms on images with surface defects on various railroad tracks, this experiment selected six clean images from two subclasses of the

RSDDs dataset and added Gaussian white noise with $\sigma = 10$ as the noisy images. After applying various denoising algorithms, the PSNR and SSIM were used for evaluation.

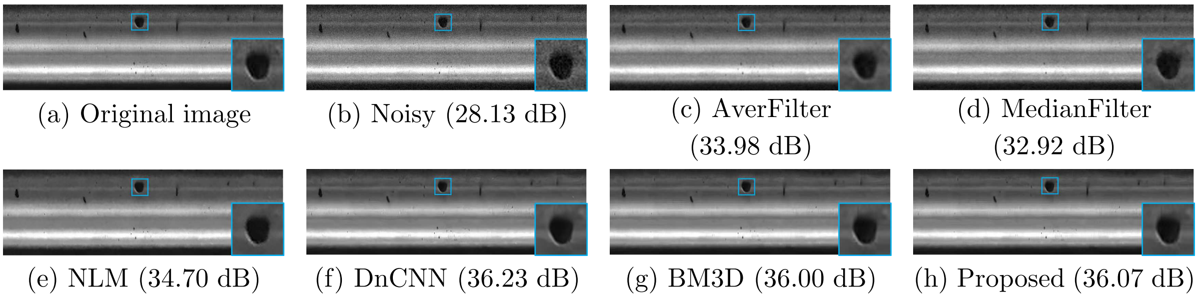
Analyzing the experimental results in the first type of sub-dataset of RSDDs, the algorithm in this paper and DnCNN have achieved a good denoising effect. The average PSNR of the algorithm in this paper reaches 35.68 dB, and the SSIM is as high as 0.9088. Compared with BM3D, the average PSNR of the algorithm in this paper is 0.04 dB higher.

Table 3. Comparison of PSNR (dB) of different algorithms at noise level $\sigma = 10$ on the first sub-dataset of RSDDs

| $\sigma$ | Image | I-rail-1 | I-rail-2 | I-rail-3 | I-rail-4 | I-rail-5 | I-rail-6 | Average |
|---|---|---|---|---|---|---|---|---|
| 10 | AverFilter | 33.84 | 33.98 | 33.64 | 33.73 | 33.32 | 34.01 | 33.75 |
| | MedianFilter | 32.94 | 32.92 | 32.69 | 32.81 | 32.39 | 32.94 | 32.78 |
| | NLM | 34.61 | 34.70 | 34.01 | 34.36 | 33.95 | 34.36 | 34.33 |
| | DnCNN | **36.10** | **36.23** | **35.61** | **35.80** | **35.46** | **35.86** | **35.84** |
| | BM3D | 35.92 | 36.00 | 35.39 | 35.61 | 35.26 | 35.68 | 35.64 |
| | Ours | 35.93 | 36.07 | 35.42 | 35.63 | 35.31 | 35.73 | 35.68 |

Table 4. Comparison of SSIM of different algorithms at noise level $\sigma = 10$ on the first sub-dataset of RSDDs

| $\sigma$ | Image | I-rail-1 | I-rail-2 | I-rail-3 | I-rail-4 | I-rail-5 | I-rail-6 | Average |
|---|---|---|---|---|---|---|---|---|
| 10 | AverFilter | 0.8767 | 0.8790 | 0.8526 | 0.8597 | 0.8613 | 0.8742 | 0.8672 |
| | MedianFilter | 0.8451 | 0.8476 | 0.8193 | 0.8259 | 0.8292 | 0.8423 | 0.8349 |
| | NLM | 0.8950 | 0.8956 | 0.8638 | 0.8710 | 0.8784 | 0.8880 | 0.8819 |
| | DnCNN | **0.9219** | **0.9247** | **0.9012** | **0.9042** | **0.9085** | **0.9163** | **0.9128** |
| | BM3D | 0.9183 | 0.9194 | 0.8961 | 0.8977 | 0.9053 | 0.9131 | 0.9083 |
| | Ours | 0.9184 | 0.9199 | 0.8969 | 0.8983 | 0.9063 | 0.9134 | 0.9088 |


(a) Original image (b) Noisy (28.13 dB) (c) AverFilter (33.98 dB) (d) MedianFilter (32.92 dB)

(e) NLM (34.70 dB) (f) DnCNN (36.23 dB) (g) BM3D (36.00 dB) (h) Proposed (36.07 dB)

Figure 3. Denoising results on image I-rail-2 by different methods (noise level $\sigma = 10$)

From the experimental results in the second sub-dataset of RSDDs, compared with other denoising methods, this paper achieves the best PSNR. The average PSNR value of the six pictures is as high as 37.81 dB, which is 0.06 dB higher than that of BM3D. SSIM is also slightly higher than BM3D.

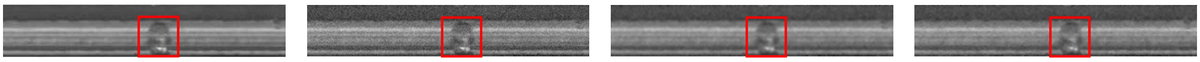
The experimental results from two sub-class datasets of the RSDDs dataset show that traditional denoising methods such as mean filtering, median filtering, and NLM denoising algorithm do not yield satisfactory results due to the abundant details and texture information in the surface defect images of railway tracks. Although the DnCNN achieves

TABLE 5. Comparison of PSNR (dB) of different algorithms at noise level $\sigma = 10$ on the second sub-dataset of RSDDs

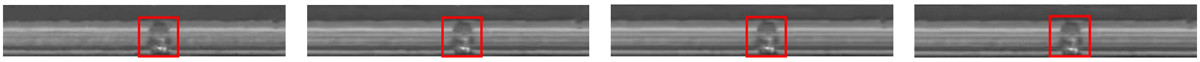
| $\sigma$ | Image | II-rail-1 | II-rail-2 | II-rail-3 | II-rail-4 | II-rail-5 | II-rail-6 | Average |
|---|---|---|---|---|---|---|---|---|
| 10 | AverFilter | 34.11 | 33.53 | 32.89 | 32.35 | 32.41 | 32.86 | 33.02 |
| | MedianFilter | 33.24 | 32.88 | 32.39 | 32.23 | 32.19 | 32.50 | 32.57 |
| | NLM | 35.55 | 35.16 | 35.00 | 34.59 | 35.02 | 35.03 | 35.05 |
| | DnCNN | 38.22 | 38.11 | 37.54 | 36.59 | 37.55 | 37.69 | 37.61 |
| | BM3D | 38.39 | 38.20 | 37.58 | 36.57 | 38.02 | 37.77 | 37.75 |
| | Ours | **38.44** | **38.33** | **37.63** | **36.61** | **38.03** | **37.85** | **37.81** |

TABLE 6. Comparison of SSIM of different algorithms at noise level $\sigma = 10$ on the second sub-dataset of RSDDs

| $\sigma$ | Image | II-rail-1 | II-rail-2 | II-rail-3 | II-rail-4 | II-rail-5 | II-rail-6 | Average |
|---|---|---|---|---|---|---|---|---|
| 10 | AverFilter | 0.9074 | 0.9039 | 0.9136 | 0.9051 | 0.8933 | 0.9090 | 0.9053 |
| | MedianFilter | 0.8852 | 0.8833 | 0.8983 | 0.8934 | 0.8804 | 0.8954 | 0.8893 |
| | NLM | 0.9306 | 0.9302 | 0.9443 | 0.9417 | 0.9380 | 0.9420 | 0.9378 |
| | DnCNN | **0.9702** | 0.9668 | **0.9725** | **0.9710** | **0.9709** | **0.9749** | **0.9710** |
| | BM3D | 0.9637 | 0.9679 | 0.9722 | 0.9613 | 0.9681 | 0.9674 | 0.9667 |
| | Ours | 0.9641 | **0.9688** | 0.9722 | 0.9626 | 0.9683 | 0.9687 | 0.9674 |



(a) Original image (b) Noisy (28.10 dB) (c) AverFilter (33.53 dB) (d) MedianFilter (32.88 dB)



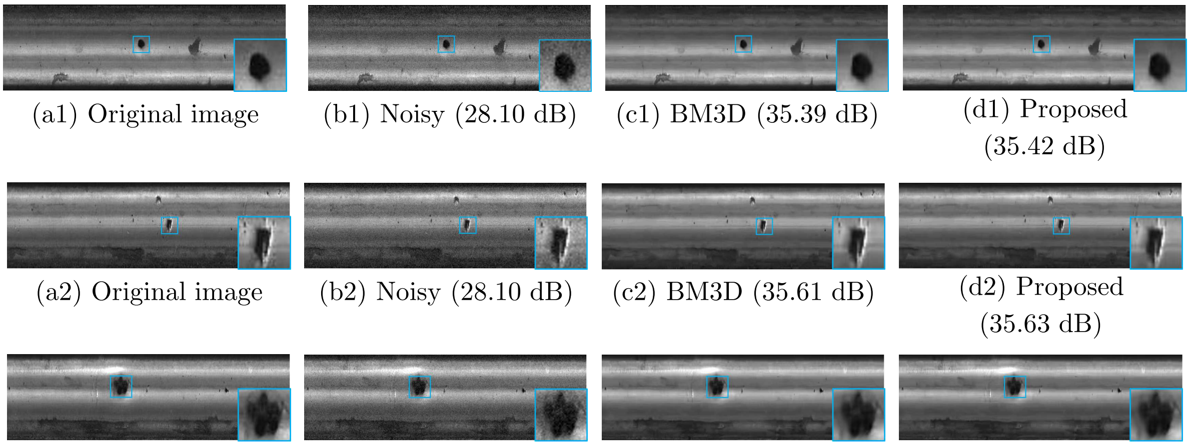
(e) NLM (35.16 dB) (f) DnCNN (38.11 dB) (g) BM3D (38.20 dB) (h) Proposed (38.33 dB)

FIGURE 4. Denoising results on image II-rail-2 by different methods (noise level $\sigma = 10$)

good denoising performance in this study, it did not meet expectations in the second experiment of the RSDDs dataset. The BM3D algorithm consistently performs well, and our algorithm significantly improves the denoising performance compared to the original BM3D algorithm. Among all non-deep learning methods used in this study, the algorithm presented in this paper achieves better PSNR and SSIM values than other image-denoising algorithms, indicating its significant noise reduction and surface defect enhancement capabilities. Moreover, it exhibits greater robustness compared to other denoising methods.

3.3.3. *Experimental analysis of the impact of image denoising algorithms on defect detection.* To test the impact of the denoising algorithm in this paper on the accuracy of defect detection, we performed defect detection on denoised images with standard deviation $\sigma = 10$. The experiment uses the Otsu algorithm to segment the image and analyzes the experimental results from six aspects: Recall rate, Accuracy rate, Precision rate, Dice coefficient, False negative rate and False positive rate.

Since in a low-noise environment, the noise interference in the data is smaller, it is easier for the denoising model to accurately capture features and patterns. Therefore, various denoising models show better results on both subclasses of RSDDs datasets.

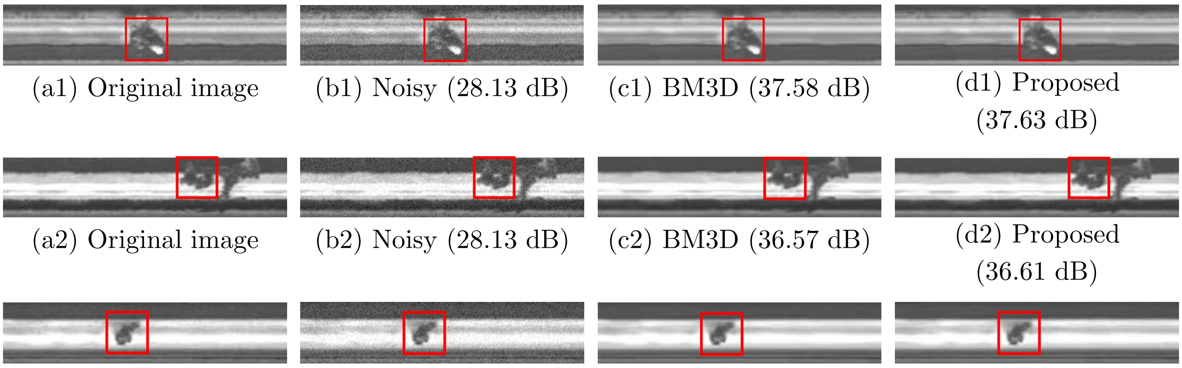


(a1) Original image (b1) Noisy (28.10 dB) (c1) BM3D (35.39 dB) (d1) Proposed (35.42 dB)

(a2) Original image (b2) Noisy (28.10 dB) (c2) BM3D (35.61 dB) (d2) Proposed (35.63 dB)

(a3) Original image (b3) Noisy (28.10 dB) (c3) BM3D (35.26 dB) (d3) Proposed (35.31 dB)

FIGURE 5. Comparison of denoised images and PSNR (dB) between BM3D and the proposed algorithm on the first sub-dataset of RSDDs (noise level $\sigma = 10$)



(a1) Original image (b1) Noisy (28.13 dB) (c1) BM3D (37.58 dB) (d1) Proposed (37.63 dB)

(a2) Original image (b2) Noisy (28.13 dB) (c2) BM3D (36.57 dB) (d2) Proposed (36.61 dB)

(a3) Original image (b3) Noisy (28.13 dB) (c3) BM3D (38.02 dB) (d3) Proposed (38.03 dB)

FIGURE 6. Comparison of denoised images and PSNR (dB) between BM3D and the proposed algorithm on the second sub-dataset of RSDDs (noise level $\sigma = 10$)

In the experimental results of defect detection of denoising images in the first subclass of the RSDDs dataset. DnCNN achieves the best Accuracy rate and False positive rate and is on par with BM3D in evaluating its Dice coefficient. At the same time, our denoising algorithm not only achieved the best Precision rate, Recall rate, Dice coefficient and False negative rate but also performed well in terms of False positive rate and Precision rate. In the comparison with BM3D's defect detection experiment results, we are slightly better than BM3D in terms of Recall rate, Accuracy rate, Precision rate, Dice coefficient, False negative rate and False positive rate.

In the experimental results of defect detection of denoising images in the second subcategory of the RSDDs datasets, the mean filter has achieved better Accuracy rate and False positive rate. We still achieve the best Precision rate, Recall rate, Dice coefficient and False negative rate. In terms of Accuracy rate, although the algorithm in this paper is worse than the mean filter denoising algorithm, it is 0.16% higher than BM3D in

Table 7. Average results of defect detection performance of different methods on the first subclass dataset of RSDDs

| Metric | Accuracy | Recall | Precision | Dice | False positive rate | False negative rate |
|---|---|---|---|---|---|---|
| AverFilter | 0.9808 | 0.9815 | 0.9696 | 0.9758 | 0.0209 | 0.0185 |
| MedianFilter | 0.9764 | 0.9744 | 0.9675 | 0.9711 | 0.0223 | 0.0256 |
| NLM | 0.9779 | 0.9804 | 0.9670 | 0.9735 | 0.0265 | 0.0196 |
| DnCNN | 0.9841 | 0.9869 | **0.9743** | 0.9805 | **0.0187** | 0.0131 |
| BM3D | 0.9838 | 0.9878 | 0.9734 | 0.9805 | 0.0194 | 0.0122 |
| Ours | **0.9842** | **0.9880** | 0.9736 | **0.9807** | 0.0192 | **0.0120** |

Table 8. Average results of defect detection performance of different methods on the second subclass dataset of RSDDs

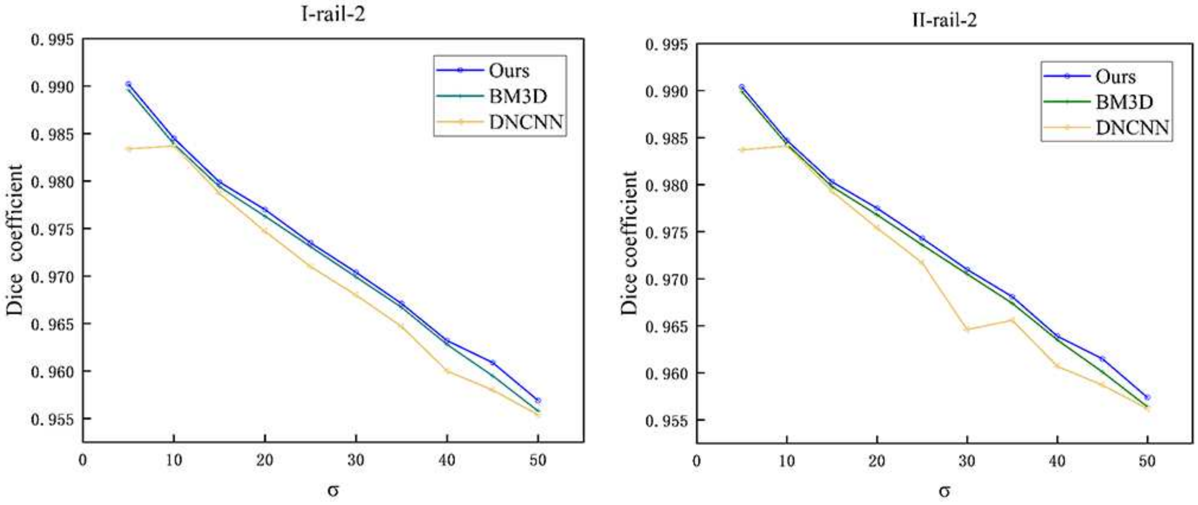
| Metric | Accuracy | Recall | Precision | Dice | False positive rate | False negative rate |
|---|---|---|---|---|---|---|
| AverFilter | 0.9852 | 0.9772 | **0.9892** | 0.9828 | **0.0096** | 0.0228 |
| MedianFilter | 0.9833 | 0.9785 | 0.9839 | 0.9812 | 0.0140 | 0.0215 |
| NLM | 0.9878 | 0.9858 | 0.9867 | 0.9859 | 0.0119 | 0.0142 |
| DnCNN | 0.9782 | 0.9901 | 0.9880 | 0.9890 | 0.0107 | 0.0099 |
| BM3D | 0.9904 | 0.9915 | 0.9870 | 0.9893 | 0.0110 | 0.0085 |
| Ours | **0.9907** | **0.9916** | 0.9886 | **0.9895** | 0.0102 | **0.0084** |

comparison with BM3D. At the same time, the Recall rate, Accuracy rate, Dice coefficient, False negative rate and False positive rate are also better than BM3D.

The experimental results of defect detection based on two subclasses of the RSDD data set show that various denoising algorithms show better data results in low-noise environments, but the algorithm in this paper still maintains strong competitiveness compared with other denoising algorithms. Compared with the BM3D algorithm, the denoising algorithm proposed in this paper can effectively improve the accuracy of defect detection while achieving a more significant denoising effect. In addition, in comparison with the experimental results of defect detection of other common denoising algorithms, our algorithm has excellent performance and has obtained the best experimental results on the two subclass datasets in terms of Accuracy rate, Recall rate, Dice coefficient and False negative rate. These results show that our denoising algorithm has great potential in the field of defect detection and provides important technical support for more accurate and reliable defect detection.

3.3.4. *The impact of denoising algorithms with varied noise intensities on defect detection.* To compare the improvement of our denoising model's accuracy in defect detection under different noise intensities, we selected I-rail-2 and II-rail-2 images corrupted with Gaussian white noise of standard deviations ($\sigma$) ranging from 5 to 50. We compared the denoised images using our method with the traditional BM3D and deep learning method DnCNN. The Otsu algorithm was employed for image segmentation, and the experimental results were analyzed.

Figure 7 presents the Dice coefficient curves for the denoised images of I-rail-2 and II-rail-2 using the three algorithms under different noise environments. It can be observed



(a) (b)

Figure 7. The Dice coefficient comparison among BM3D, DnCNN, and the proposed algorithm for II-rail-1 and II-rail-2

that the DnCNN algorithm performs poorly in terms of Dice coefficient evaluation at lower noise levels, indicating its limited robustness compared to the other two algorithms. BM3D algorithm exhibits relatively stable Dice coefficient curves across different noise environments, and its results are superior to DnCNN. From the curves, it is evident that our proposed algorithm achieves better Dice coefficient curves compared to BM3D and DnCNN under different noise environments. Our denoising algorithm outperforms the original BM3D method in improving the accuracy of defect detection under various noise environments.

4. **Conclusion.** To improve the accuracy of rail surface defect detection, this paper proposes an adaptive search neighborhood BM3D denoising algorithm. This algorithm not only effectively removes a large amount of noise in the rail surface defect image, but also preserves the edge information of the image defect well. The proposed algorithm effectively improves the accuracy of rail surface defect detection. The algorithm has a stable denoising effect under different Gaussian noise environments. In a low-noise environment, the denoising effect is optimal, which meets the actual denoising requirements of rail surface defect images. The algorithm in this paper has superiority compared with traditional methods such as BM3D. Not only did we obtain higher PSNR and SSIM values, but we also successfully improved the accuracy of defect detection. Our study has important practical implications. Through noise reduction processing, we successfully reduced the noise interference in the rail defect image, making the defect area more clearly visible. This will help promptly detect and resolve railway equipment failures and effectively improve the safety and efficiency of railway transportation. In future research, this algorithm can be applied to denoising hyperspectral images, remote sensing images, seismic data, etc. It is believed that the algorithm has certain practical value in the field of image information restoration and improving the accuracy of defect detection.

**Acknowledgment.** This work was supported in part by the National Natural Science Foundation of China under Grant 62071320, the Natural Science Foundation of Shandong Province under Grant ZR2020MF038, Shandong Province Science and Technology SMEs Innovation Capacity Improvement Project under Grant 2022TSGC2195, the Science and

Technology Innovation Development Project of Taian City under Grant 2020GX028 and Grant 2021GX045.

## REFERENCES

[1] G. Kang, S. Gao, L. Yu and D. Zhang, Deep architecture for high-speed railway insulator surface defect detection: Denoising autoencoder with multitask learning, *IEEE Transactions on Instrumentation and Measurement*, vol.68, no.8, pp.2679-2690, 2018.

[2] A. Zhang, X. Hu, M. Jin and X. Wei, Multi-target defect detection of railway track based on image processing, *2020 Chinese Control and Decision Conference (CCDC)*, pp.3377-3382, 2020.

[3] Z. Chen, Q. Wang, Q. He, T. Yu, M. Zhang and P. Wang, CUFuse: Camera and ultrasound data fusion for rail defect detection, *IEEE Transactions on Intelligent Transportation Systems*, vol.23, no.11, pp.21971-21983, 2022.

[4] L. Kou, A review of research on detection and evaluation of the rail surface defects, *Acta Polytech. Hung*, vol.19, no.3, 167, 2022.

[5] A. Sharma and J. Singh, Image denoising using spatial domain filters: A quantitative study, *2013 6th International Congress on Image and Signal Processing (CISP)*, vol.1, pp.293-298, 2013.

[6] R. Oktem, L. Yarovslavsky and K. Egiazarian, Signal and image denoising in transform domain and wavelet shrinkage: A comparative study, *The 9th European Signal Processing Conference (EUSIPCO 1998)*, pp.1-4, 1998.

[7] C. Knaus and M. Zwicker, Dual-domain image denoising, *2013 IEEE International Conference on Image Processing*, pp.440-444, 2013.

[8] Q. Song, L. Ma, J. Cao and X. Han, Image denoising based on mean filter and wavelet transform, *2015 4th International Conference on Advanced Information Technology and Sensor Application (AITS)*, pp.39-42, 2015.

[9] N. Gallagher and G. Wise, A theoretical analysis of the properties of median filters, *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol.29, no.6, pp.1136-1141, 1981.

[10] M. Zhang and B. Gunturk, A new image denoising method based on the bilateral filter, *2008 IEEE International Conference on Acoustics, Speech and Signal Processing*, pp.929-932, 2008.

[11] A. Buades, B. Coll and J.-M. Morel, A non-local algorithm for image denoising, *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, San Diego, CA, USA, vol.2, pp.60-65, 2005.

[12] R. Oktem and N. Ponomarenko, Image filtering based on discrete cosine transform, *Telecommunications and Radio Engineering*, vol.66, no.18, 2007.

[13] L. Zhang, W. Dong, D. Zhang and G. Shi, Two-stage image denoising by principal component analysis with local pixel grouping, *Pattern Recognition*, vol.43, no.4, pp.1531-1549, 2010.

[14] T. Dai, K. Gu, Q. Tang, K.-W. Hung, Y.-B. Zhang, W. Lu and S.-T. Xia, Foveated nonlocal dual denoising, *2017 IEEE International Conference on Image Processing*, pp.1881-1885, 2017.

[15] K. Dabov, A. Foi, V. Katkovnik and K. Egiazarian, Image denoising by sparse 3-D transform-domain collaborative filtering, *IEEE Transactions on Image Processing*, vol.16, no.8, pp.2080-2095, 2007.

[16] K. Dabov, A. Foi, V. Katkovnik and K. Egiazarian, BM3D image denoising with shape-adaptive principal component analysis, *Signal Processing with Adaptive Sparse Structured Representations (SPARS'09)*, 2009.

[17] M. Hasan and M. R. El-Sakka, Improved BM3D image denoising using SSIM-optimized Wiener filter, *EURASIP Journal on Image and Video Processing*, vol.2018, no.1, 25, 2018.

[18] Y. Hou and D. Shen, Image denoising with morphology- and size-adaptive block-matching transform domain filtering, *EURASIP Journal on Image and Video Processing*, vol.2018, no.1, 59, 2018.

[19] M. Wang, K. Li, X. Zhu and Y. Zhao, Detection of surface defects on railway tracks based on deep learning, *IEEE Access*, vol.10, pp.126451-126465, 2022.

[20] K. Zhang, W. Zuo, Y. Chen, D. Meng and L. Zhang, Beyond a Gaussian denoiser: Residual learning of deep CNN for image denoising, *IEEE Transactions on Image Processing*, vol.26, no.7, pp.3142-3155, 2017.

[21] D. Yang and J. Sun, BM3D-Net: A convolutional neural network for transform-domain collaborative filtering, *IEEE Signal Processing Letters*, vol.25, no.1, pp.55-59, 2017.

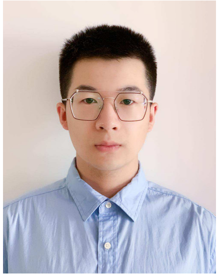
[22] H. Acikgoz and D. Korkmaz, MSRConvNet: Classification of railway track defects using multi-scale residual convolutional neural network, *Engineering Applications of Artificial Intelligence*, vol.121, 105965, 2023.

[23] Y. Xia, S. W. Han and H. J. Kwon, Image generation and recognition for railway surface defect detection, *Sensors*, vol.23, no.10, 4793, 2023.
[24] F. Wu, X. Xie, J. Guo and Q. Li, Internal defects detection method of the railway track based on generalization features cluster under ultrasonic images, *Chinese Journal of Mechanical Engineering*, vol.35, no.1, 59, 2022.
[25] J. Gan, Q. Li, J. Wang and H. Yu, A hierarchical extractor-based visual rail surface inspection system, *IEEE Sensors Journal*, vol.17, no.23, pp.7935-7944, 2017.

## Author Biography

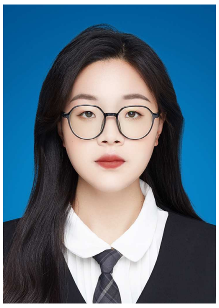
**Tao Lin** received his B.E. degree from the School of Electronic Engineering, Chaohu University, Hefei, China, in 2021. He is pursuing an M.S. degree at the School of Information Science and Engineering, Guilin University of Technology, Guilin, China. His current research interests are in the areas of image processing.

**Yingkun Hou** is an IEEE senior member. He received his Ph.D. degree from the School of Computer Science and Technology, Nanjing University of Science and Technology, in 2012. He is currently a professor at the School of Information Science and Technology, Taishan University, Taian, China. His current research interests include image processing, pattern recognition, and artificial intelligence.



**Hao Hou** received his B.Sc. degree in Information and Computing Science from the School of Mathematical Science, the University of Jinan, in 2019. He received his M.Sc. degree in Biomedical Engineering from the College of Intelligence and Information Engineering, Shandong University of Traditional Chinese Medicine, in 2023. He is currently a Ph.D. candidate at the School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing, China. His current research interests include image processing, medical image processing, and analysis.

**Zekun Lv** received his B.E. degree from the College of Software Engineering, Zhengzhou University of Light Industry, Zhengzhou, China, in 2019. He received his M.Sc. degree at the Guilin University of Technology, Guilin, China. His current research interests are in the areas of image processing. He is currently working at the School of Big Data and Digital Media at Chongqing Telecommunication Polytechnic College.



**Xiaoya Dai** received her B.E. degree from the School of Computer Engineering, Jiangsu University of Technology, Changzhou, China, in 2022. She is pursuing an M.S. degree at the School of Information Science and Engineering, Guilin University of Technology, Guilin, China. Her current research interests are in the areas of image processing.