GOSPODARKA SUROWCAMI MINERALNYMI – MINERAL RESOURCES MANAGEMENT

2024 Volume 40 Issue 4 Pages 107–130

DOI: 10.24425/gsm.2024.152723

HUI LUO[1], GAIPIN CAI[2], HAOXIANG ZOU[3]

Research on image preprocessing algorithm based on mixed denoising and texture weakening of ore images

Introduction

Image preprocessing is a key step in the image processing process, in computer vision tasks such as image segmentation, target detection, image restoration, and super-resolution reconstruction, image preprocessing is used to realize the improvement of image quality, processing efficiency, and data enhancement (Fan et al. 2022). In the actual mining production process, due to the impact of the site environment and mechanical equipment, the image acquisition and transmission caused by a variety of mixed noise makes the visual quality of the image worse, reducing the performance of the subsequent image recognition, image segmentation, target detection and other tasks processing process (Liu 2021), seriously

✉ Corresponding Author: Hui Luo; e-mail: 1982473318@qq.com

[1] School of Mechanical and Electrical Engineering, Jiangxi University of Science and Technology, Ganzhou 341000, Jiangxi, China; ORCID iD: 0009-0007-2026-1725; e-mail: 1982473318@qq.com

[2] School of Mechanical and Electrical Engineering, Jiangxi University of Science and Technology, Ganzhou 341000, Jiangxi, China; Jiangxi Province Engineering Research Center for Mechanical and Electrical of Mining and Metallurgy, Ganzhou 341000, Jiangxi, China; e-mail: 1123615286@qq.com

[3] School of Mechanical and Electrical Engineering, Jiangxi University of Science and Technology, Ganzhou 341000, Jiangxi, China; e-mail: 2591319931@qq.com

affecting the accuracy of the image processing of advanced visual tasks, so denoising and enhancement of the image is often used as a pre-processing step in some image processing tasks. Therefore, image denoising and enhancement are often used as a preprocessing step in some image-processing tasks. Realising noise removal while preserving the image texture and edge structure is the key and difficulty of the image denoising problem (Guo 2024), so it is of great significance to perform fast and efficient image restoration on noisy images.

Traditional image denoising methods such as median filtering (Maheswari and Radha 2010), mean filtering (Uğur 2019) non-local mean filtering (Zhang 2022) and other spatial domain denoising algorithms, and Fourier transform (Yan and Li 2023), wavelet transform filtering (lei et al. 2019), BM3D filtering (Dabov et al. 2007) and other transform domain denoising algorithms are mainly from the perspective of Bayesian algorithmic rules, and the use of the a priori information of the input image to achieve denoising purposes. Although they have achieved results to a certain extent, they may be limited in terms of algorithm optimisation to seek the optimal solution for image denoising and practical application scenarios.

In recent years, denoising algorithms based on convolutional neural networks (CNNs) have achieved significant visual denoising effects due to their ability to learn common a priori information about images by acquiring a large number of clear and noisy images to train the network model and reconstruct the noisy images.

Zhang et al. (2017) proposed a DnCNN network model to achieve multi-task noise removal by optimising the network structure and using batch normalisation and residual learning strategies for the problem of effective removal of specific noise. However, due to the limited sensory field of the algorithm, it cannot remove noise effectively, resulting in poor visualisation. Jingwen He et al. (2019) proposed the ADAFM denoising algorithm, which utilises fractional order differentiation and fuzzy mean filtering to capture the long-range dependencies in the image, better preserving the image details and effectively suppressing the noise, with a better denoising effect. Shi Guo et al. (2019) used the idea of residual learning and blind denoising to propose a flexible denoising algorithm, CBDNet, which learns the noise distribution and denoising pattern of an image by learning the difference between noisy and noise-free images through cross-layer connections to achieve efficient denoising of common noises, such as Gaussian noises, but the model's generalisation is weak and is not suitable for removing mixed and nonlinear noises. To improve the resolution of the denoised image and capture the detailed information of different scales in the image to improve the denoising effect, Yiyun Zhao et al. (2019) constructed a PRIDNet network using a pyramid structure adaptive mechanism to carry out feature extraction and integrate the features at different levels to capture the multi-scale information of the image at different levels, to improve the robustness and accuracy of the denoising model, and to realize the efficient denoising for real images. However, the processing ability for complex noise is limited, resulting in the blurring of some structures and details of the image. Yuzhi Wang et al. (2020) proposed a PMRID denoising model by combining residual learning and multiscale residual fusion strategy, which splits the input image into multiple overlapping image blocks, and performs

multiscale decomposition of each image block to capture the detail information of different scales in the image, and thus reconstructs the entire image after denoising. The model adopts a more complex network structure and inference algorithm, the parameter adjustment requires a lot of experiments and tuning, and the processing efficiency of the model denoising is not high. Zongsheng Yue et al. (2020) proposed a VDNet denoising model combining variational inference and CNN, which removes noise from an image by approximating the true noise-free image distribution and effectively removes the image noise by modeling the noise distribution and maximising the variational lower bound to preserve the detail information in the image.

In summary, it can be seen that the image-denoising algorithm based on the convolutional neural network can effectively remove the noise of a specific scene, but due to the locally limited receptive field of CNN, it cannot effectively capture the long-distance dependency between pixels in the image, which is not conducive to reconstructing the overall structure of the image. In addition, the CNN network model fails to achieve a good balance between the computational efficiency of up- and down-sampling and the detailed information of the image, which can easily lead to distortion of the image and cause edge blurring and texture loss.

In this paper, based on the model structure of U-Net (Ronneberger et al. 2015), the Swin Transformer U-Net network denoising model is constructed to remove the mixed noise. The U-Net residual connection and the self-attention mechanism in the Swin Transformer module (Fan et al. 2022) are used to move the window for feature extraction, to establish the information exchange and dependence between pixels between different windows, and to realize the feature fusion between different layers. The extracted feature module is input to the dual up-sampling module for image resolution reconstruction, which recovers the structure, texture, and edge information of the image and enhances the detail information of the image to improve the denoising performance of the network. Aiming at the complexity of the ore surface texture and the blurring of some edges of the hybrid denoising image, the ore image is processed with granular ore surface texture weakening and edge enhancement to improve the effect of the subsequent ore image segmentation.

1. Hybrid denoising methods

In the process of collecting ore images in the field, due to the complex field environment, poor lighting conditions, mechanical equipment operation of the violent vibration and sound interference, electronic equipment generated by the electronic noise and thermal noise in the environment caused by shot noise, Gaussian noise, pepper noise and image blurring and other composite situations. Existing denoising algorithms for a single noise have achieved very good results, but for the above multiple noise composite situation, the need for denoising algorithms in the elimination of noise and retain the image details of the trade-offs between the composite noise as far as possible to eliminate the impact of the noise compound,

and at the same time to prevent the original image over-filtering, to avoid the decline in the quality of the ore image or loss of feature information.

Figure 1 shows the overall structure of the ore image denoising SUNet network used in this paper, through the inter-module coding-decoding structure for the basic architecture of the network, the coding part of the input noisy image feature extraction, the decoding part of the reconstruction of the image to recover the image, and the use of residual jump connection splicing the feature information extracted by the coding part, which is more helpful for the decoding part of the combination of shallow pixel-level feature information to retain the detail information of the image. The network is mainly composed of three parts: shallow feature extraction, U-Net feature extraction and image reconstruction.

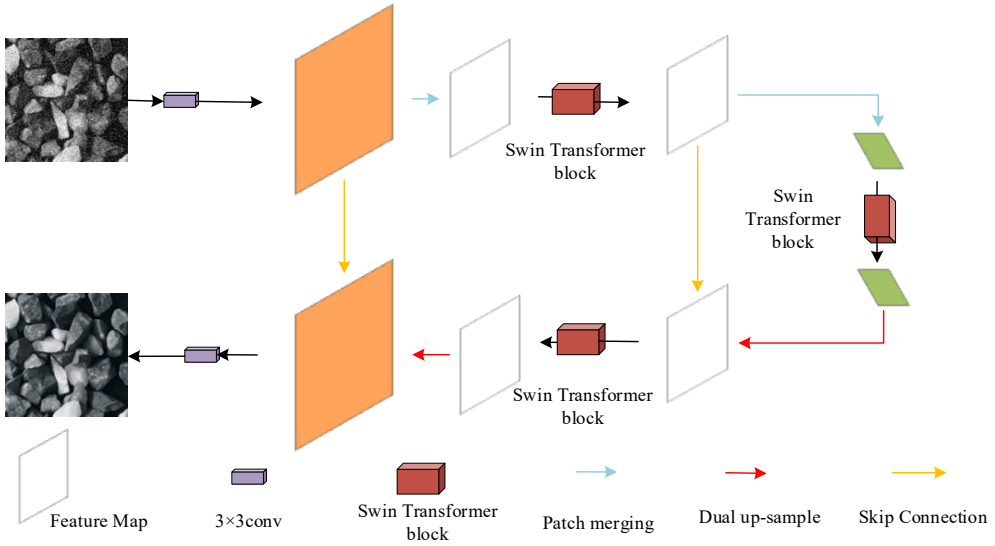

Fig. 1. Structure of SUNet denoising network

Rys. 1. Struktura sieci odszumiającej SUNet

The input noisy image is first passed through the shallow feature extraction part of the 3×3 feature extraction convolutional layer Extracting low-resolution information such as color and texture, this feature extraction process can be expressed as shown in Equation (1).

$$F_{shallow} = M_{SFE}(Y) \tag{1}$$

where C is the number of channels of the shallow feature extraction module, which is set to 96 in the network structure. The extracted shallow features are transmitted to the UNet feature extraction module Extract advanced and multi-scale deep feature information .

$$F_{deep} = M_{UFE}(F_{shallow}) \tag{2}$$

Where is the UNet feature extraction structure of Swin Transformer Block (STB) replacing the conventional convolutional layer. Finally, the feature map is double up-sampled and the convolutional layer is extracted by 3×3 features. The denoised image is reconstructed from the depth feature information.

$$\hat{Y} = M_R(F_{deep}) \tag{3}$$

1.1. U-Net feature extraction module

U-Net feature extraction module is made by Swin Transformer Block to replace the traditional convolutional layer to extract deeper information of the image, where the STB module consists of eight STL layers in series, and every two STL layers are a group as shown in Figures 2 and 3. One STL layer is the Window Multi-head Self-Attention Mechanism (W-MSA), which divides the feature map into multiple non-intersecting regions in the window during the image down-sampling process. Compared with Vision Transformer's direct self-attention processing of the entire input feature map, it can effectively reduce the amount of computation while retaining the characteristics of convolution when dealing with the input shallow feature map. The other STL layer is the Window Transformer Multiple Attention Mechanism (SW-MSA), since the W-SMA module only performs self-attention computation within the divided window, and cannot communicate and transfer feature information between windows, the SW-MSA module is introduced to realize the window offset, and the corresponding query, key, and value are generated within the new window, and the resulting query, key, and value for each feature within the offset window is generated. The SW-MSA module is introduced to realize window shifting and generate the corresponding query, key, and value within the new window, so that the self-attention calculation is performed on each pixel within the shifted window, and the long-distance information exchange between multiple different windows is realised.

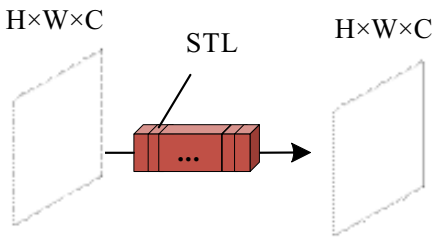

Fig. 2. Swin Transformer Block (STB)

Rys. 2. Blok transformatora STB

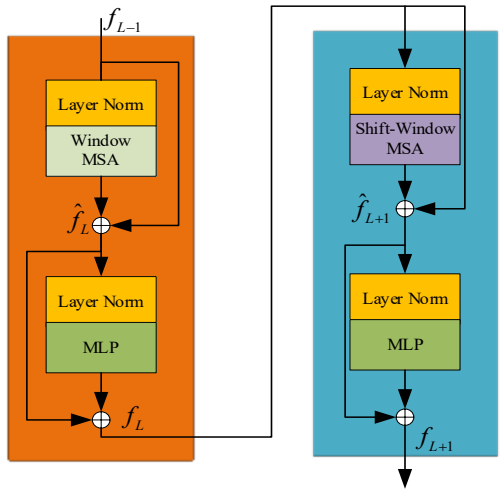

Fig. 3. Swin Transformer Layer (STL)

Rys. 3. Warstwa transformatora typu STL

The feature information f_{L-1} from the shallow feature extraction module is input into the first layer of the structure in the STL module, Layer Normalization regularisation operation is performed, and then Window Multiple Self-Attention computation is carried out, and the feature L is obtained by summing the features f_{L-1} and the output of the computation at the element level through the residual connection.

$$\hat{f}_L = W - MSA\left(LN(f_{L-1})\right) + f_{L-1} \tag{4}$$

The feature L is continued to the regularisation operation, and the new feature f_L is obtained by adding the residual network feature L through the fully connected layer MLP.

$$f_L = MLP\left(LN(\hat{f}_L)\right) + \hat{f}_L \tag{5}$$

Similar to the above steps, the feature f_L is transmitted to the second layer STL structure, but instead of computing the whole feature map of the input, the input feature map is divided into multiple small windows in the self-attention operation (SW-MSA) to obtain the feature $\hat{f}_{L+1}$ within each small window. This establishes the linkage of the local information of the feature map among different windows, and at the same time improves the efficiency of the self-attention computation.

$$\hat{f}_{L+1} = SW - MSA\left(LN(f_L)\right) + f_L \tag{6}$$

Finally, the features are again subjected to regularisation and feed-forward neural network operations, and are summed with the input $\hat{f}_{L+1}$ feature element level to obtain the new feature $\hat{f}_{L+1}$.

$$f_{L+1} = MLP\left(\mathrm{LN}(\hat{f}_{L+1})\right) + f_{L+1} \tag{7}$$

The whole process is divided into four steps: regularisation, self-attention, regularisation, and fully-connected layer, which is carried out on a small window of the feature map by shifting operations and has direct references to the original features at each step by residual connections to improve the feature extraction and learning ability of the model.

1.2. Image reconstruction

Feature maps of different sizes of the input image are extracted through the U-Net feature extraction network, and the size of the feature maps needs to be further adjusted. In the down-sampling Swin Transformer Block process, firstly, the pixel-level image features are divided into multiple non-overlapping small block regions by window, and each small block region is flattened and converted into a fixed-length feature vector by linear transformation through the fully connected layer. The linear mapping matrix splices the adjacent input features and realises the size reshaping and the output features of the specified number of channels, integrates the results of all the windows, so that each window contains the global

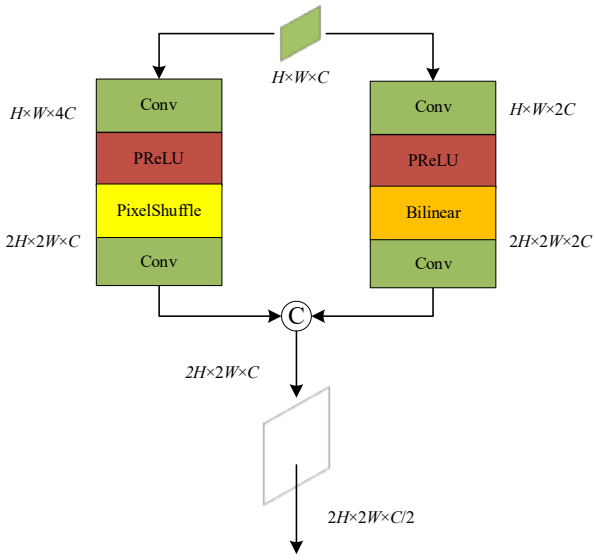

Fig. 4. Dual up-sampling structure

Rys. 4. Struktura podwójnego próbkowania w górę

information of the image, and obtains the feature vector with the same length as the original one which contains richer spatial information, and at the same time reduces the spatial dimensionality of the feature vector, and then carries out the adjustment of the window size, and reduces the size of the feature map to realize the down-sampling.

The patch expanding method, which is equivalent to transposed convolution, is used in the original Swin-unit model sampling module, but it is not possible to up-sample the feature map uniformly and continuously because transposed convolution may produce structures that look like checkerboards or blocks in the reconstructed image when up-sampling.

Therefore, the dual up-sampling operation is carried out by using a parallel combination of bilinear interpolation and sub-pixel convolution. Firstly, the image size is set to a predetermined size by bilinear interpolation to smooth the image content; at the same time, sub-pixel convolution enhances the details of the feature maps to improve the quality of the image, and the results of the two treatments are stacked in the channel dimension to complete the up-sampling operation. The architecture of the up-sampling module is shown in Figure 4.

2. Image enhancement algorithm

The texture details of an image characterize the quality of the image (Wang 2019), and for medical image segmentation, the more detailed information, the richer the image content, and the more detailed information of the segmented image the more accurate the diagnosis of the disease. However, in ore image segmentation, due to the existence of ore stacking and edge sticking in the ore image, coupled with the influence of ore surface texture features, the edge of the ore obtained by segmentation is incomplete, and at the same time, the ore surface texture is also segmented into the result, resulting in the existence of redundant points and lines in the segmented region, which is easy to segment the texture into the edges of ore particles, which seriously affects the segmentation accuracy of the ore image. Therefore, to address the above problems, this paper performs texture weakening and edge enhancement on ore images through image preprocessing operations. The general flowchart of the algorithm in this paper is shown in Figure 5.

After the ore image is processed by wavelet transform and non-local mean algorithm, the low-frequency information part of the ore particle surface retains the smooth part of the original image, while most of the noise and texture in the image is maintained in the high-frequency part, and the low-frequency part of the information is processed by Canny (1986) operator, which can effectively attenuate the influence of the texture on the ore surface on the edge extraction, and at the same time, make the surface of the ore particles smooth. The original image is segmented by the OTSU algorithm, and the binary image obtained by segmentation is contour finding and Huff transforms edge fitting to the incomplete segmented region (Xu et al. 2022). The Canny operator is averaged and weighted with the fitted edge contour to obtain the ore edge contour and superimposed to the texture-weakened

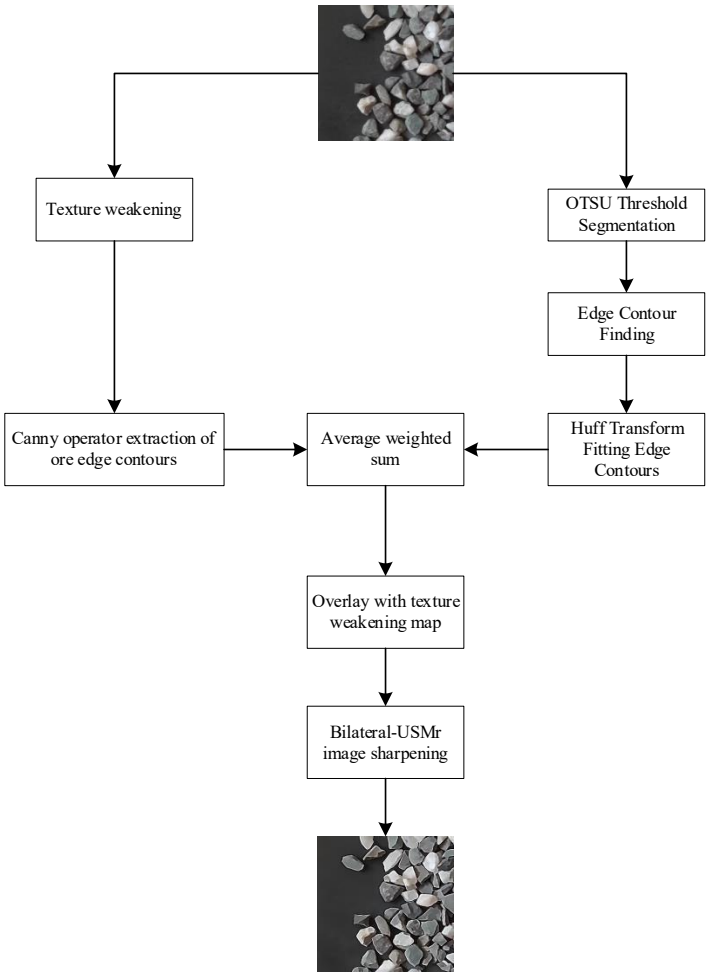

Fig. 5. Algorithm flowchart

Rys. 5. Schemat działania algorytmu

image to enhance the edges of the ore particles, and then the resultant image is sharpened by Bilateral-USMR to obtain the final texture-weakened edge-enhanced ore image.

2.1. Texture feature weakening based on wavelet transform with fast nonlocal averaging

The discrete wavelet transform proposed by Mallat (1989) can be used to obtain the relationship between neighboring scale coefficients by decomposing and reconstructing the image, and the decomposed low-pass approximation and high-pass detail components can be

fast inverse transformed to perform the discrete wavelet transform on the rows and columns of the image, respectively. The redundancy of each pixel in the image with its domain approximation pixels is processed using the Fast Non-Local Means (FNLM) algorithm while preserving the details and structural features of the original image. By setting the size of the search window and calculating the weights of the pixels within the window, the computational complexity is reduced and the speed of the operation is improved while maintaining high denoising and smoothing effects. Figure 6 shows the 2D first layer decomposition of the ore image obtained by discrete wavelet transform with fast nonlocal homogenisation, where LL represents the low-scale approximation information of the image, and LH, HH, and HL represent the image in the horizontal, diagonal, and vertical details, respectively, and it can be seen that the detail information in the low-frequency part of the image is smoother to show most of the non-texture information on the surface of the ore particles.

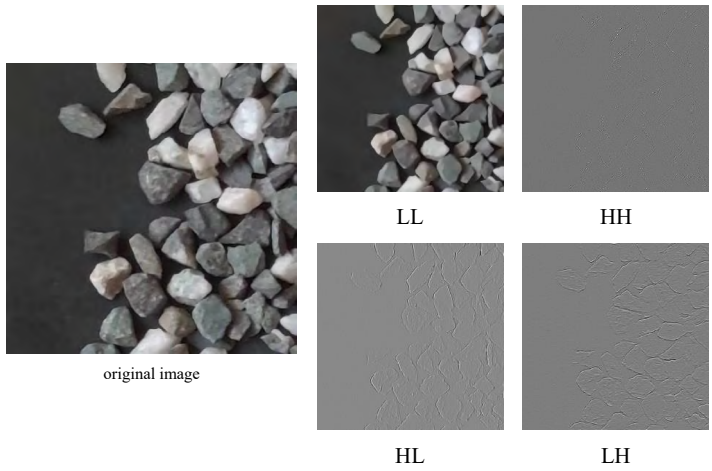

Fig. 6. Schematic diagram of image discrete wavelet transform with nonlocal mean decomposition

Rys. 6. Schemat ideowy dyskretnej transformaty falkowej obrazu z nielokalnym rozkładem średniej

2.2. Edge enhancement

2.2.1. Hough transform edge fitting

In the results obtained by the Canny operator for edge detection of images, the effect of edge detection is not obvious for the weak edges of ore and adhesion in the image, and the edges of some ore particles can't form a closed region so that it is not possible to obtain accurate and continuous edge detection results, the Hough transform algorithm is used to detect the image contour through global characteristics, and the straight lines and curves in the image are transformed from the image space to the Hough space. Through the property

that points and lines can be converted to each other in different spaces, the feature information in the image space can be converted to the parameter space, and the straight line in the image space is mapped to a point in the Hough space, where denotes the angle between the straight line and the -axis connecting the point and the origin, and denotes the distance from the origin to the straight line. This transforms the problem of detecting co-linear points into the problem of finding concurrent curves. The curve equation is shown in Equation (8).

$$y = A\sin(\omega x) \tag{8}$$

- the values of parameters A,
  ω – are in the range of $(0 < A \leq m)$, $(0 < A \leq \pi)$, respectively;
  $m \times n$ – denotes the image pixels;

The image space parameters are mapped to the parameter space expression as shown in Equation (9).

$$A = \frac{y}{\sin(\omega x)} \tag{9}$$

The parameter in the formula takes the value $0 < x < m$, $0 < y < n$; the binary map obtained by OTSU threshold segmentation algorithm is subjected to the Hough edge detection process, and according to the judgment conditions of the edge point, through the relationship between the point line and the equation, if the curve equation is satisfied, then it is judged to be an edge point; if the points in the direction of gradient in the point and the pixel point both satisfy the curve equation, then it is judged to be an edge point in the direction of gradient in the point; from this, we get an accurate and This results in an accurate and continuous edge of the ore.

2.2.2. Bilateral-USMR image sharpening

The traditional linear unsharpened mask algorithm creates a blurred version of the input image by Gaussian blurring the input image; subtracting the blurred image from the input image to obtain a "mask" to obtain an image that contains only the high-frequency details of the image, including the edges of the ore and the protruding edges of the surface; and finally superimposing the "mask" with the input image to achieve image sharpening and edge enhancement. Finally, the "mask" and the input image are superimposed to achieve image-sharpening edge enhancement. The formula is as follows in Equation (10).

$$I_{sharp}(m,n) = I(m,n) + \alpha \cdot z(m,n) \tag{10}$$

- $I(m,n)$ – denotes as the information of the input image;
  $z(m,n)$ – is the output information of the high-pass filter;

α – denotes the gain coefficient, and the magnitude of its value corresponds to the degree of image sharpening;
I_{sharp} – is the output image after sharpening.

Since the USM algorithm acquires high-frequency information images there is a part of the edge information distortion in the image sharpening result due to the problem of over-shooting. The bilateral filter can well retain the edge of the image in the process of image processing, therefore, by combining the bilateral filter (Tomasi and Manduchi 1998) with the USMR algorithm to smooth and sharpen the image, to ensure that the completeness of the edge information in the process of image sharpening. The calculation formula is as follows in Equation (11).

$$I_{sharp} = I + \alpha \cdot (I - I \cdot B) - \rho \cdot R(I - I \cdot B) \tag{11}$$

I – is the input image;
$I \cdot B$ – denotes the process of bilateral filtering of the input image;
$R(I - I \cdot B)$ – is the regularisation function, which is used to measure the intensity and complexity of the high-frequency details;
ρ – is the parameter that regulates the intensity of the regularisation.

The flowchart of the improved algorithm is shown in Figure 7.

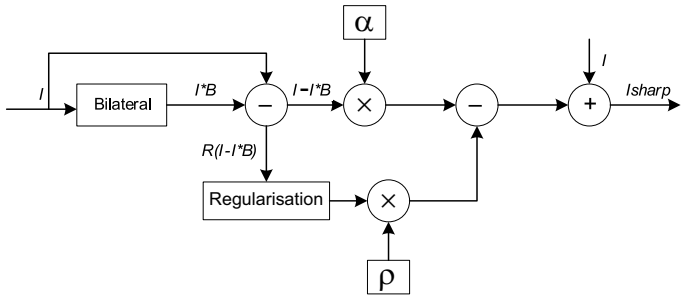

Fig. 7. Flowchart of the improved USMR algorithm

Rys. 7. Schemat działania ulepszonego algorytmu USMR

The input image is subjected to bilateral filtering process $I \cdot B$ to make difference with the input image to get the mask, the high-frequency edge information ΔI of the image, multiply ΔI by the weight and the regularisation process to make difference, respectively, to inhibit the intensity of excessive high-frequency details and prevent the loss of image edges and the amplification of the noise; and finally superimpose the result of the difference to the input image I to get the sharpened image I_{sharp}.

3. Experimental methods and results analysis

3.1 Ore image denoising dataset

The dataset used to train the model in this paper is a simulated collection of ore images on a conveyor belt using an industrial camera in a laboratory environment. The dataset is divided into training, validation, and test sets of 316 images in a ratio of 1:2:7, and the size of each image is 512 × 512. Each image is cropped into three image blocks of 256 × 256 pixels, and the training and validation sets are synthesised in the cropped image with varying degrees of Gaussian noise compounded with pretzel noise, where the training set 2,736 image groups of clean and noisy images and the validation set contains 63 image groups. The ore image part of the dataset for this paper is shown in Figure 8.

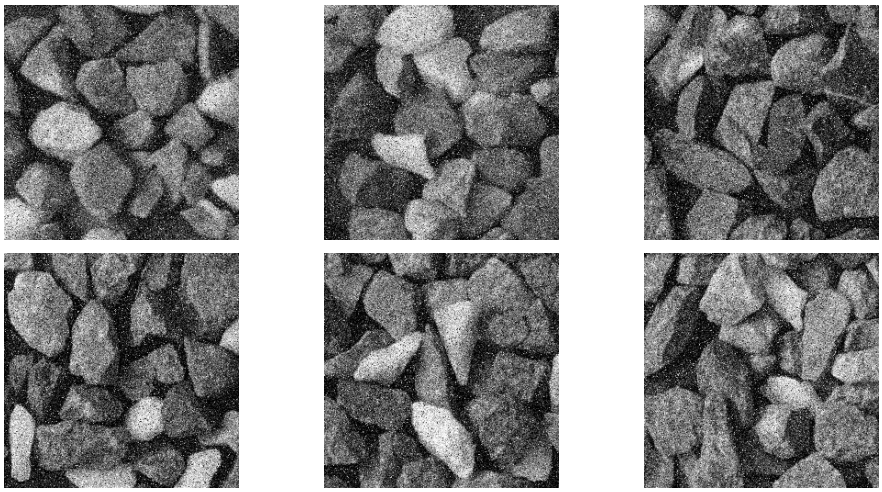

Fig. 8. Part of the dataset

Rys. 8. Część zbioru danych

3.2. Experimental configuration and parameter setting

The experiment in this paper is to build the Pytorch framework to complete the training and performance testing of the denoising model under the Windows 11 system, the processor is 13th Gen Intel(R) Core(TM) i5-13490F 2.50 GHz, the RAM capacity is 16GB, the graphics card is NVIDIA GeForce RTX3070, the PyTorch environment is Pytorch-GPU1.10.0+cu113 and the interpreter used is Python 3.8.

The input image is an RGB color ore image of 256 × 256 pixel size and the model is trained using randomly cropped image blocks for model training. The total number of iter-

ations (epoch) during training was set to 200, the network parameters were optimised using the Adam optimizer, the initial learning rate (learning rate) was set to 2e-4, the batch size was set to 4, and the size of the moving window, WIN_SIZE, was set to 8 × 8.

3.3. Evaluation indicators

3.3.1. Denoising evaluation metrics

Peak signal-to-noise ratio (PSNR, denoted as P_{PSNR} in the computation) and structural similarity (SSIM, denoted as S_{SIM} in the computation) are commonly used in image-denoising tasks to evaluate the effectiveness of image-denoising.

PSNR is a commonly used metric for evaluating the quality of a denoised image, which is defined by the root-mean-square error (MSE, denoted by M_{MSE} in the calculations) between the reference image and the denoised image. PSNR is often expressed in logarithmic form in dB, and the larger the value is, the smaller the distortion of the denoised image is and the closer it is to the reference image. The formulae for calculating PSNR and MSE are as follows in Equation (12).

$$M_{MSE} = \frac{1}{mn}\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}\left[I(i,j) - K(i,j)\right]^2 \tag{12}$$

$$P_{PSNR} = 10 \cdot \log_{10}\left(\frac{MAX_I{}^2}{M_{MSE}}\right) \tag{13}$$

- M_{MSE} – denotes the root mean square error between the denoised image I and the reference image K;
- m and n – denote the height and width of the image, respectively;
- MAX_I^2 – denotes the maximum pixel value in the color point of the image I.

If each color point is represented by 8 bits, then $MAX_I^2 = 255$. PSNR is a criterion for the detection of distortion of denoised images, but the sensitivity of human eyes to spatial frequency and luminance variations has not been taken into account in the process of calculation. Although PSNR is a criterion for detecting the degree of denoising image distortion, the sensitivity of human eyes to the change of spatial frequency and luminance is not considered in the calculation process, so when the PSNR value scores higher, it will be inconsistent with the visual perception of human eyes.

To evaluate the quality of the denoised image more comprehensively and to improve the perception of the image structure information by the visual system of the human eye, the structural similarity is introduced to evaluate the image quality, and SSIM is an index that

measures the similarity between the reference image and the resultant image by measuring the three aspects of brightness, contrast, and structure. For the denoised image I and the reference image K, the SSIM of the two images is calculated as follows in Equation (14).

$$S_{SSIM} = \frac{(2\mu_I\mu_K + c_1)(2\sigma_{IK} + c_2)}{(\mu_I{}^2\mu_K{}^2 + c_1)(\sigma_I{}^2 + \sigma_K{}^2 + c_2)} \tag{14}$$

- μ_I, μ_K – denote the mean value of I, K respectively,
- $\mu_I{}^2$, $\mu_K{}^2$ – denote the variance of I, K respectively,
- σ_{IK} – denotes the covariance of I, K,
- c_1, c_2 – denote the constants to maintain the stability, $c_1 = (k_1L)^2$, $c_2 = (k_2L)^2$, $k_1 = 0.01$, $k_2 = 0.03$,
- L – denotes the range of the pixel values of the image, range of the image pixel values.

The degree of similarity between the two images is reflected by calculating the size of the SSIM value, the value of SSIM is in the range of 0~1, the closer the value is to 1 indicates that the error between the two images is smaller, i.e., the higher the similarity between the denoised image and the reference image.

3.3.2. Texture weakening evaluation indicators

The size of the texture entropy (Jia 2015) value reflects the texture complexity or non-uniformity in the image: if the texture is complex, the corresponding entropy value is larger; conversely, if the image is uniform in the gray level, the entropy value is smaller. Because of the complex texture features of the ore surface, it is necessary to obtain all the texture information, while the first-order local texture entropy can only reflect the statistical distribution law of individual pixel gray level, and cannot capture the global texture information of the image. Each local binary pattern in the histogram of Local Binary Pattern (Zhu 2017) (LBP) corresponds to a specific local texture, which can represent the change of texture, and by combining the histogram of Local Binary Pattern and the second-order texture entropy as the evaluation index of texture weakening, the effect of texture weakening is comprehensively evaluated.

LBP calculates the texture entropy mainly on the probability distribution of the local binary patterns of the image and obtains the texture information of the image by comparing the relative grayscale of the pixels and their domains, in which the probability distribution reflects the distribution of the texture of the image, and each column in the histogram represents the frequency of the occurrence of a particular local binary pattern across the entire image. The formula for the computation of the LBP is as follows in Equation (15) and Equation (16).

$$LBP\{3,R\} = \sum \{p=0\}^2 \cdot s\left(l_p - l_c\right) \cdot 2^p \tag{15}$$

$$s(x) = \begin{cases} 1, x \geq 0 \\ 0, else \end{cases} \tag{16}$$

- 3 – is the order;
- R – is the radius of the neighborhood centered on the circle at pixel point c;
- $p = 0$ – is the index value indicating the first pixel point on the circle;
- l_p – denotes the luminance of the pth pixel point on the circle, l_c and is the pixel point luminance value;
- $s(x)$ – the function is a symbolic function, and the return value of $s(x)$ is 1 if $l_p \geq l_c$, and 0 otherwise.

Second-order texture entropy is calculated from the entropy value of an image using a gray-level Co-occurrence Matrix (GLCM) in the image space. The GLCM starts from a pixel point (x,y) in the image of a certain gray level i, and counts the probability of the simultaneous occurrence of the two points of the pixel point $(x + \Delta x, y + \Delta y)$ of a gray level j at a distance δ from the point, $P(i,j,\delta,\theta)$.

$$P(i,j,\delta,\theta) = \left\{ \left[(x,y), (x+\Delta x, y+\Delta y) \right] \middle| \begin{matrix} f(x,y) = i, f(x+\Delta x, y+\Delta y) = j; \\ x = 0,1,\cdots N_x - 1; y = 0,1,\cdots, N_y - 1 \end{matrix} \right\} \tag{17}$$

- gray value at pixel point $i,j = 0.1, ..., G - 1$,
- G – denotes the gray level of the image,
- N_x, N_y – denotes the number of rows and columns of the image.
- θ – denotes the angular direction between two pixels of the GLCM, and the directions of 0°, 45°, 90° and 135° are used.

By extracting the quadratic statistics for the spatially dependent GLCM, the formula for the second order texture entropy of the image is obtained as follows in Equation (18).

$$\text{Texture ENT2} = -\sum_{i=0}^{G-1} \sum_{j=0}^{G-1} P(i,j,\delta,\theta) \cdot \log_2 P(i,j,\delta,\theta) \tag{18}$$

3.4. Denoising comparison experiment results and analysis

In this paper, the model is trained using a homemade ore image dataset, and the model performance is tested by a test set. To visualize the performance metrics of the denoising

algorithms proposed in this paper, the PSNR and SSIM of different denoising algorithms on the ore image dataset are shown in Table 1, with bold denoting the optimal data. The representative algorithms include five blind denoising models PRIDNet, ADaFM, CBDNet, PMRID, VDNet, and the non-blind denoising model BM3D. The PSNR of SUNet proposed in this paper on ore image dataset can be up to 38.87 dB, which exceeds that of the VDNet algorithm by 8.08 dB, and the test results show that it has superior denoising performance, which outperforms that of other advanced algorithms.

Figure 9 visualises and compares the denoising effect of some algorithms with this paper's algorithm on the ore image dataset. As can be seen from Figure 9, this paper's algorithm can effectively remove the mixed noise in the real scene, and the recovered image obtained does not introduce additional artifacts while ensuring that the detailed information is not lost, which proves the effectiveness of this paper's algorithm. In contrast, other algorithms have problems such as noise residue, texture loss, and edge blurring when dealing with mixed noise.

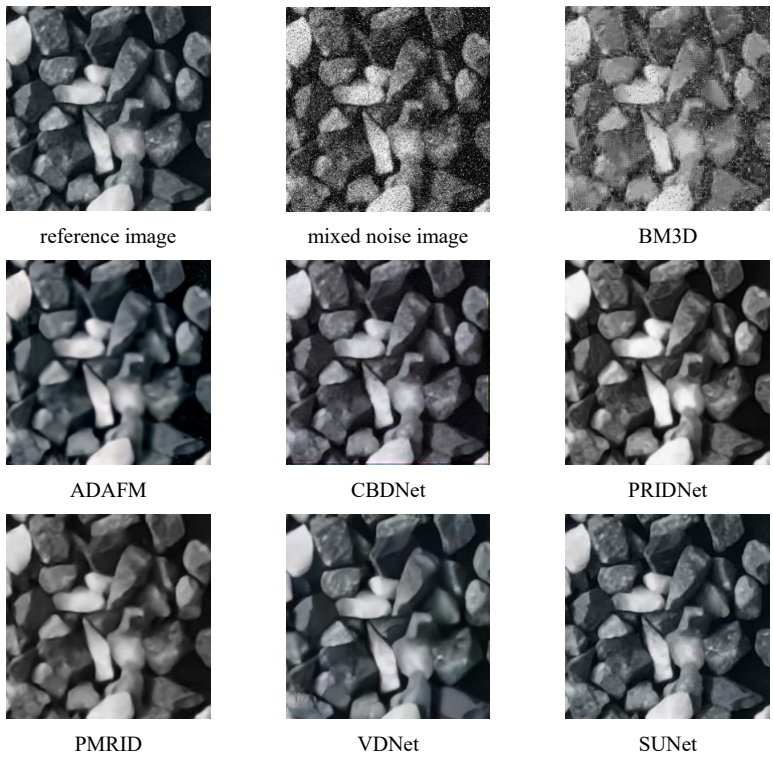

Fig. 9. Denoising visualisation results of different algorithms on ore image dataset

Rys. 9. Odszumianie wyników wizualizacji różnych algorytmów na zbiorze danych obrazu rudy

Table 1. PSNR and SSIM of different models on ore image dataset

Tabela 1. PSNR i SSIM różnych modeli zbioru danych obrazu rudy

Method	PSNR/dB	SSIM
BM3D	17.05	0.4429
PRIDNet	26.13	0.8655
ADaFM	28.18	0.8141
CBDNet	28.47	0.8946
PMRID	29.58	0.8619
VDNet	30.79	0.8964
SUNet	**38.87**	**0.9816**

3.5. Analysis of texture weakening experiment results

To verify the effectiveness of the proposed algorithm, three groups of experiments are conducted to show the effect of ore image after texture weakening respectively, calculate the evaluation index of texture weakening, and compare the effect of segmentation before and after preprocessing. The input ore image and the texture weakened image are shown in Figure 10, the calculation results of LBP histogram and second-order texture entropy of the corresponding experiments are shown in Figure 11, and the segmentation results before and after preprocessing are shown in Figure 12.

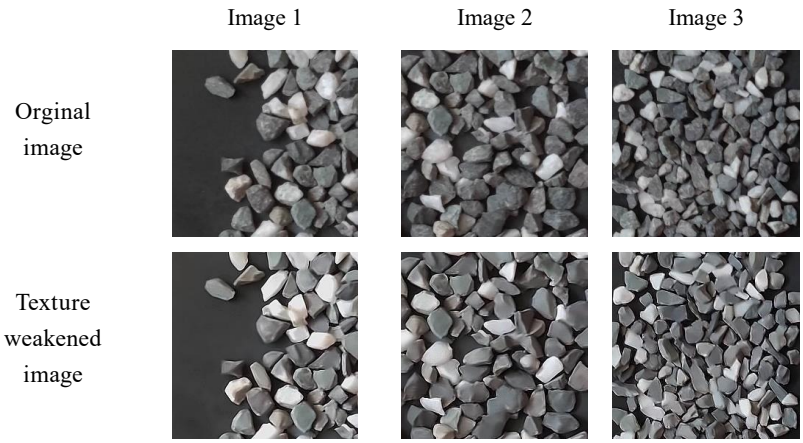

Fig. 10. Experimental original image and texture weakened image

Rys. 10. Eksperymentalny obraz oryginalny i obraz z osłabioną teksturą

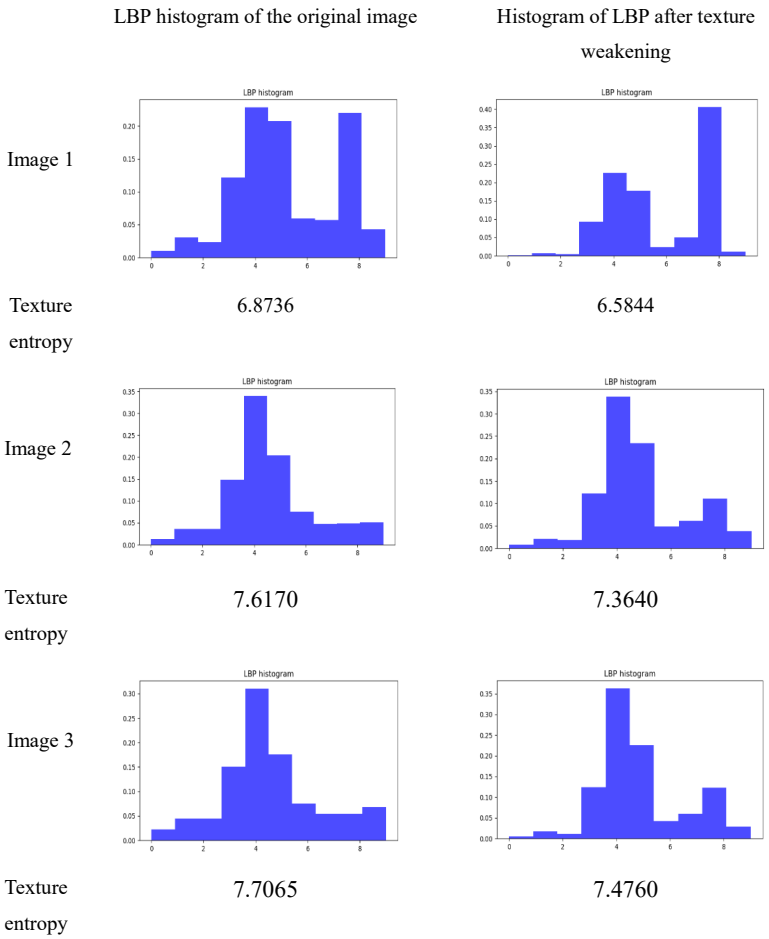

Fig. 11. LBP histogram and texture entropy calculation results

Rys. 11. Histogram LBP i wyniki obliczeń entropii tekstury

From the comparison of the results before and after the texture weakening of the ore image in Figure 10, it can be seen that the texture weakening algorithm proposed in this paper can effectively weaken the texture on the surface of the ore particles and enhance the edges of the ore, and at the same time improve the contrast between the target and the background. Comparing the LBP histogram and texture entropy before and after texture weakening in Figure 11, it can be seen that the decrease in the texture entropy value indicates that the texture information in the image is simpler, more regular, and centralised, which suggests that the texture features on the ore surface are weakened and smoothed, corresponding to the decrease in the number of local low-frequency texture patterns in the LBP histogram. With edge enhancement processing, the detailed information of the ore edges corresponding

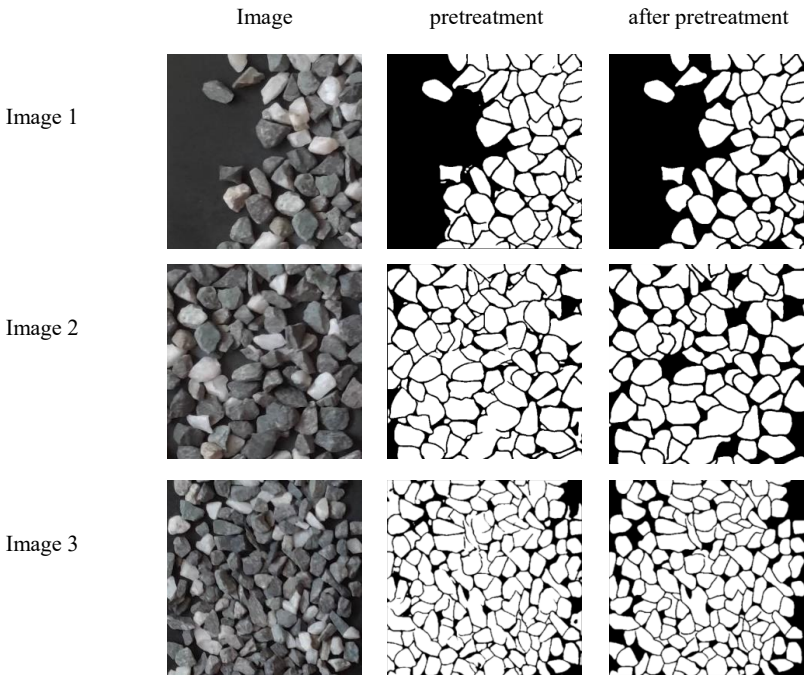

Fig. 12. Ore image segmentation results before and after preprocessing

Rys. 12. Wyniki segmentacji obrazu rudy przed i po obróbce wstępnej

to high-frequency texture patterns increases, and the number of high-frequency texture patterns in the LBP histogram increases.

Comparing the segmentation results before and after preprocessing of the ore image in Figure 11, it can be seen that after the weakening of the ore surface texture and edge enhancement, the interference of the ore surface texture on the segmentation results is effectively removed, and the problem of the existence of redundant points and lines in the segmentation region is solved; at the same time, more accurate segmentation can be realised for the ore particles that are blocked and the edges are sticky, and the incomplete segmentation of the ore edges is solved, which further It also solves the problem of incomplete segmentation of ore edges and further improves the segmentation accuracy.

In this paper, through the study of mixed noise removal algorithm of ore image and the study of texture weakening algorithm of ore surface, the removal of mixed noise is realised, and compared with the current advanced algorithms, the detailed information and the overall structure of the image are better restored; through the weakening of texture features of the ore surface to reduce the impact of local low-frequency information, increase the probability of the appearance of the edge information, and realize the texture weakening and at the same time enhance the edges of the ore particles.

Conclusions

1. This paper proposes a deep learning image denoising method based on U-Net model architecture. The Swin Transformer Block module is introduced in the coding part, which carries out self-attention computation in different area windows through the self-attention mechanism, increases the sensory field of the model, realises the information exchange between pixels of multiple windows, and improves the feature extraction ability and learning ability of the model; in the image reconstruction stage, the dual up-sampling module and the residual jump connection structure are introduced, which integrates the extracted Local and global features are introduced in the image reconstruction stage to combine the extracted local and global features, realize the smooth processing and detail enhancement of the image content, reconstruct the structure, texture, edges and other detail information of the image, and improve the denoising performance of the model. Experiments show that the denoising algorithm in this paper can realize good denoising performance while maintaining the original detail information, and effectively improving the visual quality of images in mixed noise scenes.
2. The proposed texture weakening and edge enhancement method is based on the combination of wavelet transform and non-local averaging. The texture features are weakened by smoothing the surface of the ore, and the edge profile extracted by combining the Haff edge fitting and canny operator is averaged and weighted to obtain the complete edge profile, which is then sharpened by the Bilateral-USMR to emphasize the edge portion of the ore. Experiments show that the local low-frequency information on the ore surface is effectively reduced, and the number of high-frequency texture patterns corresponding to the ore edges is increased, realising the purpose of texture weakening and edge enhancement.

Supported by the National Natural Science Foundation of China under the project "Study on the damage evolution and dissociation mechanism of encapsulated mineral interface under multi-point symmetric ultrasonic loading" (52364025).

The Authors have no conflicts of interest to declare.

REFERENCES

Canny, J. 1986. A Computational Approach to Edge Detection. *IEEE transactions on pattern analysis and machine intelligence* 8(6), pp. 679–698, DOI: 10.1109/TPAMI.1986.4767851.

Dabov et al. 2007 – Dabov, K., Foi, A., Katkovnik, V. and Egiazarian, K. 2007. Image denoising by sparse 3-D transform-domain collaborative filtering. *IEEE transactions on image processing: a publication of the IEEE Signal Processing Society* 16(8), pp. 2080–2095, DOI: 10.1109/TIP.2007.901238.

Fan et al. 2022 – Fan, C.M., Liu, T.J. and Liu, K.H. 2022. SUNet: Swin transformer UNet for image denoising. *IEEE International Symposium on Circuits and Systems (ISCAS)*. pp. 2333–2337, DOI: 10.48550/arXiv.2202.14009.

Guo et al. 2019 – Guo, S., Yan, Z., Zhang, K., Zuo, W., Zhang, L. 2019. Toward convolutional blind denoising of real photographs. *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. pp. 1712–1722, DOI: 10.1109/CVPR.2019.00181.

Guo et al. 2024 – Guo, X., Xu, L., Cheng, C. and Huo, J. 2024. Image denoising algorithm based on weighted kernel norm minimization and improved wavelet threshold function. *Journal of National University of Defense Technalogy* 46(2), pp. 238–246.

He et al. 2019 – He, J.W., Dong, C., Qiao, Yu. 2019. Modulating Image Restoration with Continual Levels via Adaptive Feature Modification Layers. *32nd IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, DOI: 10.1109/CVPR.2019.01131.

Jia, Y. 2015. *Digital Image Processing*. Wuhan University Press, 264.

Lei et al. 2019 – Lei, Z., Su, W. and Hu, Q. 2019. Multimode Decomposition and Wavelet Threshold Denoising of Mold Level Based on Mutual Information Entropy. *Entropy* 21(2), DOI: 10.3390/e21020202.

Liu, J.S. 2021. *Research On Real-World Image Denoising Based On Deep Neural Network*. Xidian University.

Maheswari, D. and Radha, V. 2010. Noise Removal In Compound Image Using Median Filter. *International Journal on Computer Science and Engineering* 2(4), pp. 1359–1362.

Mallat, S. 1989. A theory for multiresolution signal decomposition. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 11(7), pp. 674–693, DOI: 10.1109/34.192463.

Ronneberger et al. 2015 – Ronneberger, O., Fischer, P. and Brox, T. 2015. U-Net: Convolutional Networks for Biomedical Image Segmentation. *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 234–241, DOI: 10.1007/978-3-319-24574-4_28.

Sedaghatjoo et al. 2024 – Sedaghatjoo, Z., Hosseinzadeh, H. and Bigham, B.S. 2024. *Local Binary Pattern(LBP) Optimization for Feature Extraction*, DOI: 10.48550/arXiv.2407.18665.

Tan et al. 2024 – Tan, T., Wang, K. and Zhou, W. 2024. Edge Detection of Plant Root Images Based on Improved Canny Algorithm.*Photonics & Electromagnetics Research Symposium (PIERS)*. Tomasi, C. and Manduchi, R. 1998. Bilateral filtering for gray and color images. *Sixth International Conference on Computer Vision (IEEE Cat. No.98CH36271)*, DOI: 10.1109/ICCV.1998.710815.

Uğur et al. 2019 – Uğur, E., Dang, N.H.T., Le, M.H. and Serdar, E. 2019. An Iterative Mean Filter for Image Denoising. *IEEE Access*. 7(1), pp. 167847–167859, DOI: 10.1109/ACCESS.2019.2953924.

Wang, Y. 2019. *Research on Multi-Feature Fusion for Image Texture Analysis.* Nanjing University of Aeronautics and Astronautics.

Wang et al. 2020 – Wang, Y., Huang, H., Xu, Q., Liu, J., Liu, Y. and Wang, J. 2020. *Practical Deep Raw Image Denoising on Mobile Devices*. [In:] *Computer Vision – ECCV 2020*. 12351, pp. 1–16, DOI: 10.1007/978-3-030-58539-6_1.

Yan, F.J. and Li, B.Z. 2022. Spectral graph fractional Fourier transform for directed graphs and its application. *Signal Processing* 210, DOI: 10.48550/arXiv.2203.10199.

Yue, Z.S., Zhao, Q., Zhang, L. and Meng, D.Y. 2020. Dual Adversarial Network: Toward Real-World Noise Removal and Noise Generation. *Computer Science – Computer Vision and Pattern Recognition,* pp. 41–58, DOI: 10.48550/arXiv.2007.05946.

Zhang et al. 2017 – Zhang, K., Zuo, W., Chen, Y., Meng, D. and Zhang, L. 2017. Beyond a gaussian denoiser: Residual learning of deep cnn for image denoising. *IEEE Transactions on Image Processing* 26(7), pp. 3142–3155, DOI: 10.1109/TIP.2017.2662206.

Zhang, X.B. 2022. Two-step non-local means method for image denoising. *Multidimensional Systems and Signal Processing* 33(6), pp. 341–366, DOI: 10.1007/s11045-021-00802-y.

Zhao et al. 2019 – Zhao, Y., Jiang, Z., Men, A. and Ju, G. 2019. Pyramid Real Image Denoising Network. *IEEE Visual Communications and Image Processing (VCIP)*, DOI: 10.1109/VCIP47243.2019.8965754.

RESEARCH ON IMAGE PREPROCESSING ALGORITHM BASED ON MIXED DENOISING AND TEXTURE WEAKENING OF ORE IMAGES

Keywords

image hybrid denoising, texture feature weakening, edge enhancement, texture entropy

Abstract

Ore particle size information is an important basis for mining enterprises to regulate crushing parameters, due to the complex and harsh environment during the acquisition of ore images on the conveyor belt, resulting in the existence of a variety of composite noise interference in the motion target image, the surface texture characteristics of the ore and the edge of the fuzzy and other problems, thus affecting the effective acquisition of ore particle size information. To address the above issues, an image-denoising network based on global and local feature extraction and an edge enhancement algorithm for texture feature weakening is proposed. The denoising network consists of a shallow local feature extraction module and a Transformer-based U-Net global feature extraction module, which aims to combine the powerful global modeling capability of the Transformer with the local modeling advantage of convolutional network, and reconstructs the image resolution through the dual up-sampling structure, to realize the accurate output of contextual detail information. A texture weakening method based on wavelet transform and fast non-local averaging is proposed to smooth the image and weaken the texture characteristics of the ore surface, and edge sharpening is combined with Bilateral-USMR to enhance the edges of the ore particles to realize the preprocessing of the ore image. The preprocessing results were objectively evaluated and experimentally verified. The results show that the image preprocessing method improves the accuracy of image segmentation and the applicability of the ore particle size measurement technology in complex environments.

BADANIA NAD ALGORYTMEM WSTĘPNEGO PRZETWARZANIA OBRAZU OPARTYM NA MIESZANYM ODSZUMIANIU I OSŁABIANIU TEKSTURY OBRAZÓW RUDY

Słowa kluczowe

hybrydowe odszumianie obrazu, osłabienie cech tekstur, wzmocnienie krawędzi, entropia tekstur

Streszczenie

Informacje o wielkości cząstek rudy stanowią dla przedsiębiorstw wydobywczych ważną podstawę do regulowania parametrów kruszenia ze względu na złożone i trudne środowisko podczas akwizycji obrazów rudy na przenośniku taśmowym, co skutkuje występowaniem różnorodnych zakłóceń szumu złożonego w ruchomym obrazie docelowym, właściwości tekstury powierzchni rudy i krawędzi rozmytej oraz inne problemów, wpływających w ten sposób na skuteczne pozyskiwanie informacji o wielkości cząstek rudy. Aby rozwiązać powyższe kwestie, zaproponowano sieć odszumiającą obraz opartą na globalnej i lokalnej ekstrakcji cech oraz algorytm wzmacniania krawędzi

w celu osłabienia cech tekstury. Sieć odszumiająca składa się z płytkiego modułu ekstrakcji cech lokalnych i modułu globalnej ekstrakcji cech U-Net opartego na transformatorze, którego celem jest połączenie potężnych możliwości globalnego modelowania transformatora z przewagą lokalnego modelowania sieci splotowej i rekonstrukcja rozdzielczości obrazu poprzez podwójną strukturę próbkowania w górę, aby uzyskać dokładne dane wyjściowe szczegółowych informacji kontekstowych. Zaproponowano metodę osłabiania tekstury opartą na transformacji falkowej i szybkim uśrednianiu nielokalnym w celu wygładzenia obrazu i osłabienia właściwości tekstury powierzchni rudy, a wyostrzanie krawędzi wiązało się z dwustronnym-USMR w celu uwydatnienia krawędzi cząstek rudy w celu realizacji wstępnej obróbki obrazu rudy. Jej wyniki zostały obiektywnie ocenione i zweryfikowane eksperymentalnie. Pokazują one, że metoda wstępnego przetwarzania obrazu poprawia dokładność segmentacji obrazu i możliwość zastosowania technologii pomiaru wielkości cząstek rudy w złożonych środowiskach.