GOSPODARKA SUROWCAMI MINERALNYMI – MINERAL RESOURCES MANAGEMENT

2024 Volume 40 Issue 4 Pages 131–146

DOI: 10.24425/gsm.2024.152724

HUI LUO[1], QINGWEN ZENG[2], XIAOYAN LUO[3], ZHEN HU[4]

Improved Pix2PixGAN water-bearing ore reflection image restoration method

Introduction

In recent years, digital image processing technology has been deeply applied to the beneficiation process in the field of mining, with the intention of improving production efficiency and reducing the energy consumption of the production process. In the beneficiation process, the washing step causes water to saturate the surface of the ore and fill the pores formed by the adhesion of the ore. Light reflects off the water-bearing ore particles, resulting in reflective spots in the captured ore image. The dense reflective spots form a reflective area, which makes it impossible to accurately extract the characteristic information of the ore,

✉ Corresponding Author: Xiaoyan Luo; e-mail: lxy9416@163.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, Jiangxi Province, China; ORCID iD: 0009-0002-5693-2977; e-mail: 3089411264@qq.com

[3] Jiangxi Mining and Metallurgical Engineering Research Center, China; School of Mechanical and Electrical Engineering, Jiangxi University of Science and Technology, Ganzhou, Jiangxi Province, China; e-mail: lxy9416@163.com

[4] School of Mechanical and Electrical Engineering, Jiangxi University of Science and Technology, Ganzhou, Jiangxi Province, China; e-mail: 2330830247@qq.com

resulting in erroneous segmentation of the ore image. A clear ore image can extract ore features more accurately, so removing the water and repairing the reflective areas into a clear ore image becomes an important step for subsequent image segmentation.

Currently, the restoration of images usually utilises auxiliary data around or outside the missing region to reason about the information and repair the damaged region (Yong Goo et al. 2020). The repair methods are in two categories: traditional methods and deep learning (Efros and Leung 1999; Bertalmio et al. 2000). The traditional methods mainly treat the reflections in the image as damaged regions and utilize the surrounding pixels to fill in the way, but the processed image leaves obvious differences between the damaged region and the neighboring regions, and the effect is not satisfactory. Deep learning methods such as (Zhang et al. 2018) use the attention mechanism against the network of the raindrop method, the generative network to join the attention mechanism and long and short-term memory network to extract the raindrops attention map, and then use the jump connection network to remove the raindrops image, the discriminative network uses the convolutional layer and the all-connected layer, the effect of raindrops removal is obvious (Wu et al. 2021) proposed neural network de-reflections method by training two networks, i.e. the U-Net and VGG network, to achieve the removal of light spots from the image, which is characterised by a good removal effect (Chen 2022) used an improved pairwise generative adversarial network to remove highlights from the surface of specular objects, using pairwise learning to realize the process of specular reflection to diffuse reflection, thus achieving the effect of removing highlights (Isola et al. 2016). It can be seen that the deep learning algorithm has outstanding results in practical applications.

Therefore, this paper proposes an improved Pix2PixGAN deep learning model to repair the reflective areas of ore images and remove the water to get clear ore images. In this model, a combination of ResNet network and structural similarity loss function is used to extract the features of the watery ore image and select the minimised structural similarity loss value to update the network parameters through the ResNet network. The effectiveness of the algorithm is verified by experiments.

1. Pix2pixGAN network model

Isola et al. (2016) proposed the network architecture of Pix2pixGAN in 2016 to solve the difficulty of interconverting different image domains. Improved on the framework of CGAN's network (Jiaosheng et al. 2021), the input control conditions are replaced by images instead of random noise. Pix2pixGAN needs paired images as training data, supervised completion of the one-to-one image domain mapping relationship, the output result map completes the conversion of the corresponding image domain, and the overall flow of its network is shown in Figure 1.

In Figure 1, X denotes the real image; G denotes the generator network; D denotes the discriminator network; y denotes the input image. When training the generator network,

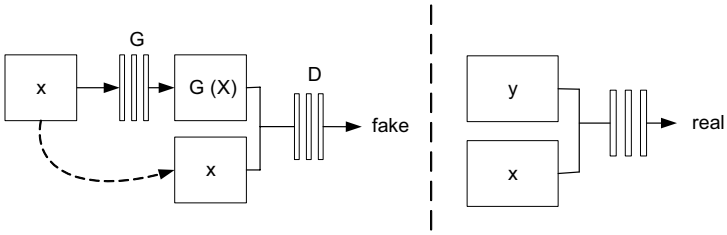

Fig. 1. Pix2pixGAN network structure diagram

Rys. 1. Schemat struktury sieci Pix2pixGAN

the input image and the generator-generated image need to be jointly input into the discriminator after the loss value obtained by the discriminator back-propagation to the network to update the parameters of the generator so that the generated image is infinitely close to the input image. Therefore, the task of the discriminative network consists of two aspects: on the one hand, to determine the authenticity of the generated image, and on the other hand, to determine the degree of similarity between the generated image and the original image. In addition, Pix2pix is improved from the CGAN network, and the two have similar network architectures, so the Pix2pixGAN network loss function and the CGAN network loss function have the same $L_{cGAN}(G,D)$ conditional generation of the adversarial network loss function (Baoqi et al. 2022), as shown in Equation (1):

$$L_{cGAN}(G,D) = E_{x,y}\left[\log D(x,y)\right] + E_{x,z}\left[\log\left(1 - D\left(x, G(x,z)\right)\right)\right] \tag{1}$$

In order to reduce the phenomenon of blurring of the generated image, the $L1$ loss function is introduced in the Pix2pix network (Nan et al. 2020) as $L1$ enhances the clarity of the generated image, and its $L1$ loss function is shown in Equation (2):

$$L_{L1}(G) = E_{x,y,z}\left[\left\|y - G(x,z)\right\|_1\right] \tag{2}$$

Therefore, the total loss function of the Pix2pixGAN network is shown in Equation (3):

$$G^* = \arg \min_G \max_D L_{cGAN}(G,D) + \lambda L_{L1}(G) \tag{3}$$

In Equation (1) to (3), x denotes the real image, y denotes the image associated with the real image, and z denotes the random noise vector.

2. Research on the network of reflective remediation of water-bearing ores

2.1. Improved model generation network design

The residual module is introduced into the U-Net network structure to form a complete generative network. The main body of the U-Net network structure consists of the left side, composed of convolution and downsampling, and the right side, composed of inverse convolution and upsampling. This structure is able to capture the features in the image efficiently (Wenjie and Honglei 2021). Combined with the characteristics of the U-Net network structure, the generative network is designed to improve the Pix2PixGAN model.

As shown in Figure 2, the U-Net network model includes a compression module on the left and a decompression module on the right. The compression module on the left is used to extract the features and contextual information of the image, and the decompression module on the right is used to accurately locate the feature positions. The use of the jump connection can enhance the features of the image, reduce the loss of information due to the decompression process, and by splicing together the up-sampled feature maps and the down-sampled feature maps. The fusion of shallow and deep semantic information can be realised to reduce the loss of image information. Only a few training samples are needed to achieve strong segmentation and denoising effect. However, there is a lack of shallow network structure, which can not effectively extract the deeper feature information of the image.

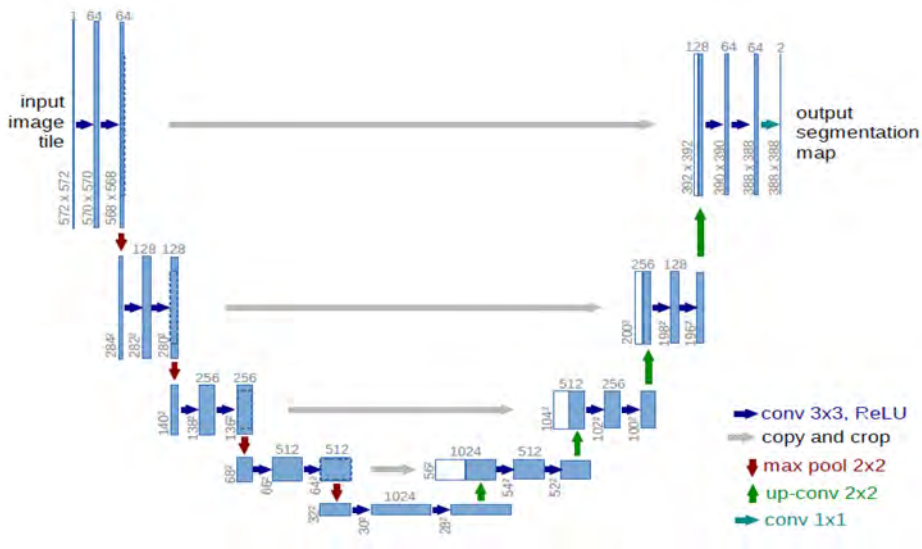

Fig. 2. U-Net network model structure

Rys. 2. Struktura modelu sieci U-Net

Therefore, to improve the network model's ability to extract the image's deep-level feature information, the residual module is introduced into the U-Net network model so that the model can extract deeper features and reduce the blurring phenomenon of the reflective area of the image. The introduction of the residual network effectively solves the problem of gradient disappearance caused by increasing the number of network layers to extract the deep-level feature information of the image, as well as obtaining more comprehensive features of the water-bearing ore image. The structure of the improved generative network is shown in Figure 3.

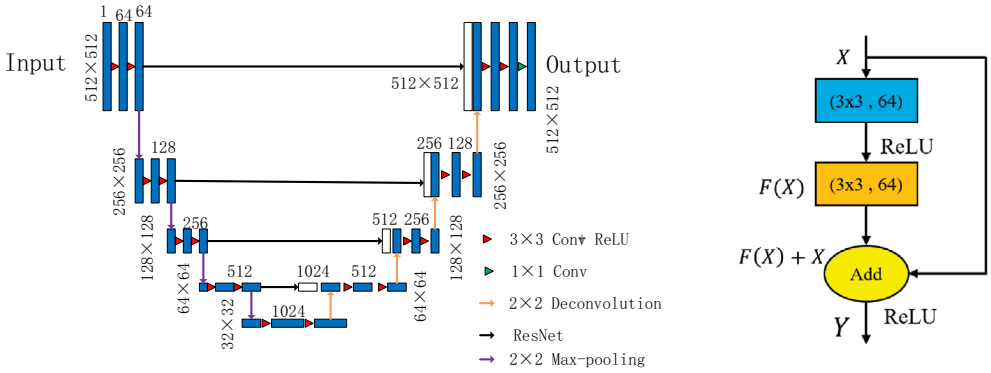

Fig. 3. Improved generative network structure

Rys. 3. Ulepszona struktura sieci generatywnej

X is the output of the upper layer as the input here, $F(X)$ is the result of X as the input operated twice in the convolutional layer, and finally the residual mapping learned by the residual module and the constant mapping of the input form the output Y. Assuming that W_1 and W_2 denote the learned parameters of the first and second convolutional layers, and the ReLU activation function is denoted by $\sigma(W_1X)$, the computation of the residual module is shown in Equation (4):

$$F = W_2\sigma(W_1X)$$
$$Y = F\left(X,\{W_i\}\right) + X \tag{4}$$

2.2. Improved model discriminative network design

The role of the discriminator is to differentiate between the image generated by the generator and the real image by judging their differences to improve the generation quality and

realism of the generator. The specific goal is to establish a reliable boundary between the generator and the real image to facilitate the learning and optimisation of the generator. The GAN network discriminator is to generate a probability value for the image, and its generated image is generally fuzzy. By using the full convolutional neural network PatchGAN architecture, as shown in Figure 4, the generated image and the real image will be jointly input into the discriminator network, through the discriminator network will output a discriminator matrix by controlling the size of the discriminator matrix can be divided into different blocks of the image, the value of a discriminator matrix represents the probability value of the original image block, which can effectively repair the effect of the reflective area on the image, to obtain a clearer and more precise generated image. This can effectively repair the effect of reflective areas on the image, thus obtaining a clearer generated image.

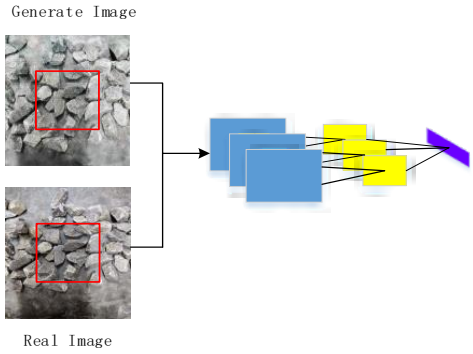

Fig. 4. Discriminator Schematic Diagram

Rys. 4. Schematyczny diagram dyskryminatora

2.3. Introduction of structural similarity loss function

The loss function (Ronneberger 2015) mainly uses the L1 and MSE loss functions at the Pixel level, but the differences in the content structure and the mismatch of the edge texture details that exist in this method in terms of the fidelity of the model-generated images lead to a more ambiguous effect of the images containing reflective regions.

To address this problem, perceptual loss and adversarial loss, which focus more on image appearance effects, are introduced. This new loss function is designed to focus on capturing contextual information, aiming to more comprehensively measure the difference between the generated image and the real image, thus allowing the model to determine whether the image generation is better or not more accurately (Demir and Unal 2018). It takes more account of contextual information and image similarity, which includes the perceptual loss, structural similarity function loss, and adversarial loss of the VGG19 network (Yixiang et al. 2022).

To judge the quality of reconstructed images, the corresponding evaluation index Structural Similarity Index (*SSIM*) loss function is used (Cao et al. 2023). It contains three factors: brightness, contrast, and structure, and its value ranges from [0,1]. The larger the *SSIM*, the smaller the structural difference between the reconstructed image and the real image, and the better the reconstruction quality. The smaller the *SSIM*, the larger the difference between the reconstructed image and the real image structure, and the poor reconstruction quality. Therefore, $SSIM_i$ is defined to represent the structural similarity between the reconstructed image of the ith batch of the generating network and the real image, and is the average value of the structural similarity of the reconstructed images of all batches of the generating network, then the loss function of structural similarity of the generating network can be expressed as shown in Equation (5).

$$L_{SSIM} = 1 - E\left(\sum SSIM_i\right) \tag{5}$$

The calculation of the brightness in the *SSIM* is shown in Equation (6):

$$l(x,y) = \frac{2\mu_x\mu_y + c_1}{\mu_x^2 + \mu_y^2 + c_1} \tag{6}$$

The contrast in *SSIM* is calculated as shown in Equation (7):

$$c(x,y) = \frac{2\sigma_x\sigma_y + c_2}{\sigma_x^2 + \sigma_y^2 + c_2} \tag{7}$$

The structural calculations in *SSIM* are shown in Equation (8):

$$s(x,y) = \frac{\sigma_{xy} + c_3}{\sigma_x\sigma_y + c_3} \tag{8}$$

μ_x represents the x means and μ_y represents the y means; σ_x^2 and σ_y^2 represent the variance of x and y, respectively; σ_{xy} represents the covariance of x and y; and c is the corresponding determinant value to address the effects of contrast and luminance. Thus, *SSIM* is calculated as:

$$SSIM(x,y) = \left[l(x,y)^{\alpha} c(x,y)^{\beta} s(x,y)^{\gamma} \right] \tag{9}$$

In order to enhance the computational efficiency, the values of α, β, γ are fixed as constant 1, and the computational formula in the above Equation is brought in. Then, the computational formula of SSIM is obtained, as shown in Equation (10).

$$SSIM = \frac{\left(2\mu_x\mu_y + c_1\right)\left(2\sigma_x\sigma_y + c_2\right)\left(\sigma_{xy} + c_3\right)}{\left(\mu_x^2 + \mu_y^2 + c_1\right)\left(\sigma_x^2 + \sigma_y^2 + c_2\right)\left(\sigma_x\sigma_y + c_3\right)} \tag{10}$$

The image reconstruction quality is proportional to the structural similarity index (*SSIM*) and inversely proportional to the structural similarity loss function. Then, the network parameters are updated by minimising the structural similarity loss value to make the generated image closer to the real image and reduce the structural differences, especially in terms of reducing the influence of reflective areas on the generated image. The specific steps are as follows:

1. Calculate the Structural Similarity Index (*SSIM*) value for the generated and real images.
2. Calculate the SSIM value out of the loss based on the difference in the *SSIM* value between the real image and the generated image.
3. Combine the structural similarity loss with the *L*1 and *MSE* loss functions by the backpropagation algorithm and then use the gradient descent method to update the parameters of the network.

3. Experimental results and analysis

3.1. Experimental environment construction

Generating adversarial networks itself is difficult to train and requires a large number of datasets, so there are certain requirements for the experimental equipment, the experiment needs to be run on a 64-bit Ubuntu 20.04 system, relying on GPUs. The entire deep-learning experimental environment is shown in Table 1 and Table 2:

Table 1. Experimental hardware environment configuration

Tabela 1. Eksperymentalna konfiguracja środowiska sprzętowego

Hardware Configuration			
CPU	GPU	RAM	Graphics Memory
16G Intel(R) Xeon(R) Platinum 8350C CPU @ 2.60GHz	RTX 3090 * 1	43GB	24GB

In addition to the main configuration described above, the network requires other python libraries, including: matplotlib = 3.5.2, Pillow = 9.1.1, pytorchvision = 0.12.0, etc.

Table 2. Experimental software environment configuration

Tabela 2. Eksperymentalna konfiguracja środowiska programowego

Software Configuration			
Python 3.8	CUDA 11.3	PyTorch 1.10.0	Pycharm Community2019.2

3.2. Dataset production

To study the moisture removal effect of water-bearing tungsten ore particles and the existence of the reflective zone phenomenon, two datasets of water-bearing tungsten ore particle images and dry tungsten ore particle images were used in the experimental process. The real particle size of the ore is in the range of 525 mm, and the image size is 1900 × 2048 pixels. A total of 500 pairs of training datasets and 200 pairs of test datasets were randomly selected.

The relative positions of the ore particles and the industrial camera remain unchanged during the production of the dataset to ensure data consistency under the same conditions, and then Mosaic (Tun et al. 2024) data enhancement is carried out to adjust the image size to 512 × 512 pixels, which improves the processing speed of the dataset while increasing the diversity of the dataset, which is conducive to the model training that can better capture the real ore features without being disturbed by the changes in the position caused by the interference of artifactual features caused by location changes. Figure 5(a) is the image of dry tungsten ore particles, and Figure 5(b) is the image of watery tungsten ore particles.

(a) Image of dried ore particles

(b) Image of aqueous ore particles

Fig. 5. Ore particle image data pairs

Rys. 5. Pary danych obrazu cząstki rudy

3.3. Improved network parameterisation and training

In order to train to get the best water removal and repair reflective area effect image, when training the generative adversarial network, a single hyper-parameter, and optimizer are prone to repeated oscillations in the direction of gradient descent and training dispersion phenomenon when encountering complex data samples, so different hyper-parameters are set, and different optimizers are selected to optimize the training performance of the generator and the discriminator.

In this experiment, the Adam (Adaptive Moment Estimation) optimizer is used to optimize the model parameters and calculate the adaptive learning rate for each parameter, which avoids generating repeated oscillations in the direction of gradient descent. Experiments show that the network model using the Adam optimizer begins to converge after 1,200 rounds of training and is able to maintain convergence all the time to avoid divergence.

Generating Adversarial Network hyperparameters are set, and the network is trained with an initialised learning rate of 0.002, keeping the subsequent learning rate constant. The input water-bearing ore image size is 512 × 512. Experimentally, it is found that the best training results are achieved when the number of samples per training is 1. The model converges when the number of iterations is 1,900 and the SSIM and PSNR values are at their highest.

To reduce the probability of model non-convergence in model training, it is proposed to normalize the input water-bearing ore particles pictures and use training improvement measures such as the ReLU activation function and the Adam optimizer to make the generative adversarial network model reach the optimal state.

As shown in Figures 6 and 7, the loss values of the generator network and the discriminator network increase at 720 and 910 times of training, which is due to the presence of challenging water-bearing reflective ore samples in the dataset, and therefore leads to an

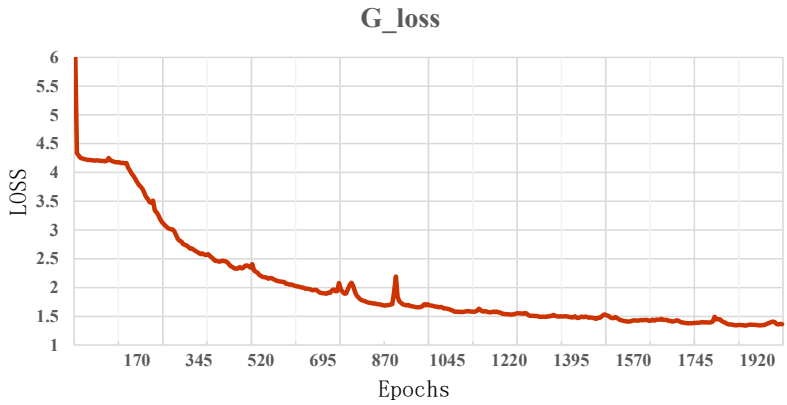

Fig. 6. Generate network loss curves

Rys. 6. Generowanie krzywych strat sieciowych

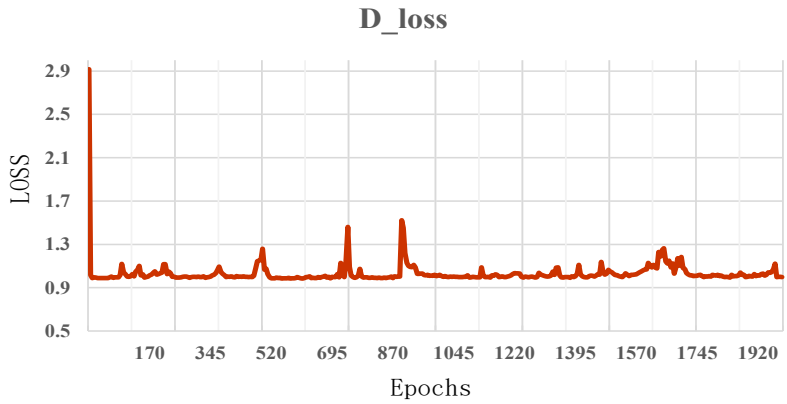

Fig. 7. Discriminant Network Loss Curve

Rys. 7. Krzywa strat sieci dyskryminacyjnej

increase in the loss values. With the increase of training times, the model gradually adapts to the samples, and the loss value gradually decreases and the decreasing trend slows down after 1,200 times of training, and the overall trend of the loss curve value stays stable after 1,900 times, and the overall trend of the discriminator network loss value stays stable and fluctuates around 1, which indicates that the network training reaches a desirable state as a whole, and further proves that the improved algorithm model is more effective. In addition, the statistics during the validation of the trained network show that the processing time of a single image ranges from 32.28 to 37.69 ms, which reveals the characteristics of the trained network, such as short time-consuming and fast speed in processing a single image.

3.4. Evaluation of moisture removal and reflective zone restoration effectiveness

In order to evaluate the effectiveness of the improved Pix2PixGAN algorithm, different algorithm comparison experiments were carried out with the test dataset of water-bearing tungsten ore particle images in this paper. Pix2pixGAN and CycleGAN models are selected for comparative analysis to analyze their differences visually from the human eye on the one hand and to compare their practical effects from objective and subjective evaluation indexes on the other hand.

As can be seen from Figure 8, Figure 8(a) and Figure 8(b) are the watery and dry ore images that need to be processed. The Pix2pixGAN algorithm removes the effect as in Figure 8(d). For the watery ore particles image of the water removal effect of the color distortion, the reflective area is not enough to repair the boundary structure is not obvious and a small number of noise points are distributed so that there is a blurring phenomenon in the image.

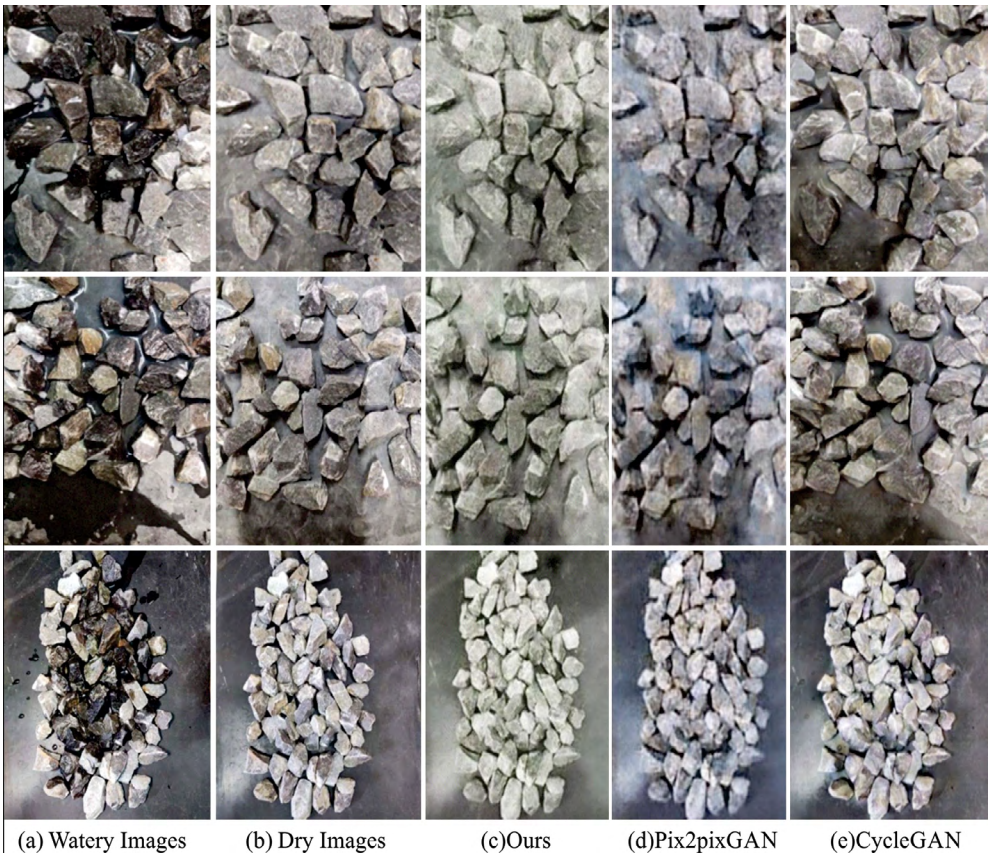

Fig. 8. Comparison of water removal by different methods

Rys. 8. Porównanie usuwania wody różnymi metodami

The effect is shown in Figure 8(e), in the water-containing ore, the ore image containing less water removal and repair of the reflective area still has the phenomenon of fuzzy edges, the repair effect is not good, and the ore image containing more water in the water-containing ore image does not remove the water well, and the reflective area is poorly repaired, so it cannot achieve the application effect. Improved model repair results in Figure 8(c), it can be seen that the overall no obvious fuzzy phenomenon, water removal and repair of reflective area effect is better, and can be very good to maintain the contextual semantic coherence of the image, for the more water-containing ore image also has a good repair effect, is conducive to obtaining a clear ore image for segmentation.

In order to better compare and analyze the related algorithms, an objective evaluation method is used to analyze the repaired tungsten ore images using structural similarity and peak signal-to-noise ratio indexes, and Pix2pix, CycleGAN, and Ours algorithms are selected to carry out *PSNR* and *SSIM* values to react to the restoration effect of the

tungsten ore images, in which the *PSNR* is calculated (Fuzhen et al. 2023) as shown in Equation (11–12):

$$MSE = \frac{1}{mn}\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}\left[I(i,j)-K(i,j)\right]^2 \tag{11}$$

In Equation (11) *MSE* denotes mean square error; *I* denotes the number of pixels; denotes the original image; and *K* denotes the noise image. The *PSNR* value is then obtained by substituting the *MSE* value in Equation (12).

$$PSNR = 10\cdot\log_{10}\left(\frac{MAX_I^2}{MSE}\right) = 20\cdot\log_{10}\left(\frac{MAX_I}{\sqrt{MSE}}\right) \tag{12}$$

In Equation (12) MAX_I denotes the maximum possible pixel value. Where the *PSNR* value is directly proportional to the quality of the restored image. The experimental results are shown in Table 3:

Table 3. PSNR and SSIM of different methods

Tabela 3. PSNR i SSIM różnych metod

	Pix2pix	CycleGAN	Ours
PSNR	23.671	22.997	23.759
SSIM	0.7013	0.6651	0.7141

Through the analysis of Table 3, it is found that the improved algorithm in this paper has improved in both *PSNR* and *SSIM*, in which it has improved 8.8% and 1.28% for Pix2 pixGAN, respectively. Compared with CycleGAN, there is a substantial improvement. Through the above experimental analysis, it can be verified that the improved algorithm is superior and practicable in removing moisture from water-containing tungsten ores.

Conclusion

1. An improved Pix2pixGAN network is proposed to solve the problem of water removal and reflection area repair for water-bearing ore images, and the residual module is added to the Pix2pixGAN network so that the model can extract deeper features and repair the

blurring problem of the reflection area of the image. The structural similarity loss function is introduced into the loss function to more comprehensively measure the difference between the generated and real images and obtain clearer and drier ore images.

2. Comparing the effects of Pix2pixGAN network and CycleGAN network on removing water and repairing reflections of watery ore images, it is found that the color distortion of the de-watering effect of the image of Pix2pixGAN network, the boundary structure of the reflective ore image is not obvious, and there are a small number of noise points, and the image effect is not good. The CycleGAN network has poor image de-watering and repairing reflective effect when the water content is higher. A comparison of the output results of the improved Pix2pixGAN model reveals that it is able to effectively remove water and repair reflective areas in watery ore images.
3. In terms of evaluation indexes, the *PSNR* and *SSIM* of the improved Pix2pixGAN network are improved by 8.8% and 1.28%, respectively, which further verifies the effectiveness of the improved Pix2pixGAN network model for water-bearing ore images in removing water and restoring reflective areas. Combined with the field practice, the improved Pix2pixGAN network model is integrated into the conveyor belt conveying system, which can output clear, dry, and reflection-free tungsten ore images and provides effective methods and ideas for effectively solving the problem of detecting the particle sizes of tungsten, iron, and other kinds of ores in the water-laden environment, and for improving the intelligent level of water-laden ore particle image processing.

This work was supported by the National Natural Science Foundation of China (Grant No. 52364025) and the Key R&D Program of Jiangxi Provincial Department of Science and Technology (Grant No. 20181ACE50034).

The Authors have no conflicts of interest to declare.

REFERENCES

Baoqi et al. 2022 – Baoqi, L., Haining, H., Jiyuan, L., et al. 2022. Research on image enhancement algorithm of turbid water body based on improved CycleGAN. *Journal of Electronics and Information* 44(7), pp. 2504–2511 (*in Chinese*).

Bertalmio et al. 2000 – Bertalmio, M., Sapiro, G., Caselles, V. and Ballester, C. 2000. Image inpainting. *Proceedings of the 27th annual conference on Computer graphics and interactive techniques*, New York, United States, pp. 417–424.

Cao et al. 2023 – Cao, J., Qiang, Z., Lin, H., He, L. and Dai, F. 2023. An Improved BM3D Algorithm Based on Image Depth Feature Map and Structural Similarity Block-Matching. *Sensors* 23(16), DOI: 10.3390/s23167265.

Demir, U. and Unal, G. 2018. Patch-Based Image Inpainting with Generative Adversarial Networks, DOI: 10.48550/arXiv.1803.07422.

Efros, A.A. and Leung, T.K. 1999. Texture synthesis by non-parametric sampling. *IEEE International Conference on Computer Vision (ICCV)*, Kerkyra, Greece, pp. 1033–1038, DOI: 10.1109/ICCV.1999.790383.

Fuzhen et al. 2023 – Fuzhen, Z., Chen, W., Bing, Z., Sun, C. and Qi, C. 2023. An improved generative adversarial networks for remote sensing image super-resolution reconstruction via multi-scale residual block. *The Egyptian Journal of Remote Sensing and Space Sciences* 26(1), pp. 151–160, DOI: 10.1016/j.ejrs.2022.12.008.

Isola et al. 2017 – Isola, P., Zhu, J.Y., Zhou, T. and Efros, A.A. 2017. Image-to-Image Translation with Conditional Adversarial Networks. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, DOI: 10.1109/CVPR.2017.632.

Junchao, C. 2022. A study on highlight removal from specular object surface based on improved pairwise generative adversarial network. Wuhan Textile University, (*in Chinese*).

Li et al. 2021 – Li, J., Li, Y., Li, J., Zhang, Q., Yang, G., CHen, S. and Li, J. 2021. Single Exposure Optical Image Watermarking Using a cGAN Network. *IEEE Photonics Journal* 13(2), DOI: 10.1109/JPHOT.2021.3068299.

Lu et al. 2022 – Lu, Y., Qiu, Y., Gao, Q. and Sun, D. 2022. Infrared and visible image fusion based on tight frame learning via VGG19 network[J]. *Digital Signal Processing* 131(3), DOI: 10.1016/j.dsp.2022.103745.

Nan et al. 2020 – Nan, Z., Xiaoming, X., Xin, X., Wei J., Chu Z., Tian P., Shi, L., Li, C. and Zhou, J.Z. 2020. Robust T-S Fuzzy Model Identification Approach Based on FCRM Algorithm and L1-Norm Loss Function. *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, DOI: 10.1109/ACCESS.2020.2973722.

Ronneberger, O., Fischer, P. and Brox, T. 2015. U-Net: Convolutional Networks for Biomedical Image Segmentation. *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*. [Online:] http://lmb.informatik.uni-freiburg.de/people/ronneber/u-net.

Shin et al. 2020 – Shin, Y.G., Sagong, M.C., Yeo, Y.J., Kim, S.W. and Ko, S.J. 2020, PEPSI++: Fast and Lightweight Network for Image Inpainting. *IEEE transactions on neural networks and learning systems*, pp. 252–265, DOI: 10.48550/arXiv.1905.09010.

Tun et al. 2024 – Tun, Y., Yongcun, G., Shuang, W., et al. 2024. Obstacle recognition of unmanned railroad motor vehicle in coal mine. *Journal of Zhejiang University (Engineering Edition)* 58(1), pp. 29–39 (*in Chinese*).

Wenjie, W. and Honglei, W. 2021. Improvement and implementation of occluded face image restoration based on generative adversarial network. *Computer Application and Software* 38(01), pp. 217–221+249 (*in Chinese*).

Wu et al. 2021 – Wu, Y., He, Q., Xue, T., Garg, R., Chen, J., Veeraraghavan, A. and Barron J.T. 2021. How to Train Neural Networks for Flare Removal. *IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 2219–2227, DOI: 10.1109/ICCV48922.2021.00224.

Xuaner et al. 2018 – Xuaner, Z., Ren, N. and Qifeng, C. 2018. Single image reflection separation with perceptual losses. *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 4786–4794, DOI: 10.1109/CVPR.2018.00503.

IMPROVED PIX2PIXGAN WATER-BEARING ORE REFLECTION IMAGE RESTORATION METHOD

Keywords

loss function, ore particle, water removal, generative adversarial network

Abstract

At the ore crushing site, the crushed ore must be washed to remove sediment, and the washing step puts the detected ore in a watery environment, resulting in the presence of reflective areas in the image of watery ore particles. Aiming at the problem of mis-segmentation of ore images due to the masking of ore feature information by the reflective area, an improved Pix2PixGAN model is proposed to solve the problem of removing water and repairing the reflective area in watery ore images.

The ResNet network with good stability is used to comprehensively extract the features of watery ore images, improve the stability of network training, introduce the structural similarity loss function, and update the network parameters by minimising the structural similarity loss value to reduce the structural differences between the reconstructed image and the real image. The experimental results show that the improved Pix2PixGAN model compares with the Pix2PixGAN and CycleGAN models; the watery ore image removes the water image reflection restoration better and, at the same time, improves the structural edge clarity of the generated dry ore particle image. The PSNR and SSIM evaluation metrics are improved by 8.8 and 1.28%, respectively, further verifying the effectiveness of the improved algorithm. This innovative approach provides a feasible solution for image processing at the ore-crushing site. It is of great significance for the subsequent enhancement of image recognition, segmentation, and reduction of misjudgment.

ULEPSZONA METODA PRZYWRACANIA OBRAZU RUDY ZAWIERAJĄCEJ WODĘ PIX2PIXGAN

Słowa kluczowe

funkcja strat, cząstka rudy, usuwanie wody, generatywna sieć przeciwstawna

Streszczenie

W miejscu kruszenia rudy rozdrobniona ruda musi zostać wypłukana w celu usunięcia osadu, a etap płukania umieszcza wykrytą rudę w środowisku wodnym, co powoduje obecność obszarów odblaskowych na obrazie cząstek rudy wodnej. Mając na celu rozwiązanie problemu błędnej segmentacji obrazów rudy z powodu maskowania informacji o cechach rudy przez obszar odblaskowy, zaproponowano ulepszony model Pix2PixGAN w celu rozwiązania problemu usuwania wody i naprawy obszaru odblaskowego na obrazach rudy wodnej. Sieć ResNet o dobrej stabilności jest używana do kompleksowego wyodrębniania cech obrazów rudy wodnej, poprawy stabilności treningu sieci, wprowadzenia funkcji utraty podobieństwa strukturalnego i aktualizacji parametrów sieci poprzez minimalizację wartości utraty podobieństwa strukturalnego w celu zmniejszenia różnic strukturalnych między zrekonstruowanym obrazem a obrazem rzeczywistym. Wyniki eksperymentów pokazują, że ulepszony model Pix2PixGAN wypada lepiej w porównaniu z modelami Pix2PixGAN i CycleGAN; obraz rudy wodnistej lepiej usuwa odtworzenie odbicia obrazu wody i jednocześnie poprawia wyrazistość krawędzi strukturalnych wygenerowanego obrazu suchych cząstek rudy. Wskaźniki oceny PSNR i SSIM poprawiły się odpowiednio o 8,8 i 1,28%, co dodatkowo potwierdza skuteczność ulepszonego algorytmu. To innowacyjne podejście zapewnia wykonalne rozwiązanie do przetwarzania obrazu w miejscu kruszenia rudy i ma duże znaczenie dla późniejszej poprawy rozpoznawania obrazu, segmentacji i redukcji błędnej oceny.