


# GAN Data Augmentation Methods in Rock Classification

Gaochang Zhao <sup>1</sup>, Zhao Cai <sup>1,\*</sup>, Xin Wang <sup>1</sup> and Xiaohu Dang <sup>2</sup> 

<sup>1</sup> School of Science, Xi'an University of Science and Technology, Xi'an 710054, China; zgc\_xa1965@sina.com (G.Z.)

<sup>2</sup> School of Geology and Environment, Xi'an University of Science and Technology, Xi'an 710054, China

\* Correspondence: cz15619985159@163.com

**Abstract:** In this paper, a data augmentation method Conditional Residual Deep Convolutional Generative Adversarial Network (CRDCGAN) based on Deep Convolutional Generative Adversarial Network (DCGAN) is proposed to address the problem that the accuracy of existing image classification techniques is too low when classifying small-scale rock images. Firstly, Wasserstein distance is introduced to change the loss function, which makes the training of the network more stable; secondly, conditional information is added, and the network has the ability to generate and discriminate image data with label information; finally, the residual module is added to improve the quality of generated images. The results demonstrate that by applying CRDCGAN to the augmented rock image dataset, the accuracy of the classification model trained on this dataset is as high as 96.38%, which is 13.39% higher than that of the classification model trained on the non-augmented dataset, and 8.56% and 6.27% higher than that of the traditional dataset augmented method and the DCGAN dataset augmentation method, respectively. CRDCGAN expands the rock image dataset, which makes the rock classification model accuracy effectively improved. The data augmentation method was found to be able to change the accuracy of the classification model to a greater extent.

**Keywords:** GAN; rock image generation; data augmentation; condition information; rock classification



**Citation:** Zhao, G.; Cai, Z.; Wang, X.; Dang, X. GAN Data Augmentation Methods in Rock Classification. *Appl. Sci.* **2023**, *13*, 5316. <https://doi.org/10.3390/app13095316>

Academic Editors: Rui Yao and Hancheng Zhu

Received: 26 March 2023

Revised: 17 April 2023

Accepted: 21 April 2023

Published: 24 April 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Rock classification plays an important role in geological exploration and belt transportation, etc. [1,2]. Misclassification of rocks can cause safety hazards and economic losses [3]. Currently, deep-learning-based image classification techniques are the mainstream approach for rock classification [4]. Deep learning requires a large amount of data for model training, which in reality is constrained by the conditions of rock image acquisition, making it difficult to obtain a large amount of rock image data and obtaining an uneven proportion of samples [5,6]. The scarcity of data becomes an important reason for the accuracy of rock classification models [7].

The poor lighting conditions, electromagnetic interference, dust, and fog in the imaging environment make it difficult to augment rock image data [8]. In recent years, scholars have proposed some practical methods for rock image data augmentation as a way to improve the accuracy of rock image classification. Hong et al. used the traditional data enhancement technique for coal gangue image data for data enhancement, which effectively improved the classification accuracy of coal gangue; the method is not applicable to the dataset containing multiple rock images, the traditional image data enhancement method generates images with little differentiation from the original image data, and there is a large amount of redundant feature information, which is not conducive to improving the generalization ability of the classification model [9]. Baraboshkin et al. subjected the collected 2000 stratigraphic rock images to data augmentation to 20,000 images to improve the accuracy of the classification task and prevent overfitting, but the method in the paper was not effective in augmenting the image data, and there was a large amount of redundant feature information in the generated images [10]. Cheng et al. proposed

SinGAN, a rock image generation network based on generative adversarial networks, and the experimental results showed that the method can automatically generate rock images with diverse features and effectively retain the image feature detail information, but the generated images cannot effectively improve the accuracy of the rock thin section image classification model [11].

In this paper, we propose a CRDCGAN data augmentation method based on Generative Adversarial Network (GAN) by analyzing the features of rock image data in a rock classification task scenario. CRDCGAN starts by adding conditional information to enable the generative network to generate matching sample label pairs and generates multi-class rock image data based on label information and noisy data. At the same time, the residual module is introduced, and the loss function is changed to solve the problems of poor quality of generated rock images, redundancy of generated rock image feature information, and instability of training of the Generative Adversarial Network. By augmenting the rock image data with CRDCGAN and generating a large amount of various rock image data with labels, the accuracy of the classification model is greatly improved, and better results are also obtained in other evaluation methods.

## 2. Generating Adversarial Networks

Generative Adversarial Networks were proposed by Ian Goodfellow [12]. GAN includes two sub-networks: generative network and discriminative network, where the generative network is responsible for learning the distribution of real samples. The discriminative network is responsible for determining whether the source of the sample is the generative network or the original data. The losses of the generation network and the discriminative network are incorporated. The objective function is as follows:

$$\min_G \max_D V(G, D) = E_{x \sim P_{data}(x)} [\log D(x)] + E_{z \sim P_g(z)} [\log(1 - D(G(z)))]. \quad (1)$$

A network is generated when training stops. The optimization goal becomes

$$\max_D V(G, D) = E_{x \sim P_{data}(x)} [\log(D(x))] + E_{z \sim P_g(x)} [\log(1 - D(x))]. \quad (2)$$

Then, points are converted from the following equation:

$$\max_D V(G, D) = \int_x P_{data}(x) \log(D(x)) dx + \int_x P_g(x) \log(1 - D(x)) dx, \quad (3)$$

where the maximum value of  $V(G, D)$  is obtained. The maximum value of the requested equation is calculated as follows:

$$P_{data}(x) \log(D(x)) + P_g(x) \log(1 - D(x)). \quad (4)$$

Then, log base is set to 2,  $P_{data}(x) = A$ ,  $P_g(x) = B$ , and  $D(x) = t$  to obtain the following:

$$f(t) = A \log_2 t + B \log_2(1 - t). \quad (5)$$

The first-order derivative of Equation (5) is given by the following:

$$f'(t) = \frac{1}{\ln 2} \frac{A - (A + B)t}{t(1 - t)}. \quad (6)$$

For the above equation, make  $f'(t) = 0$ , and obtain  $t = A/(A + B)$ ;  $f(t)$  is continuous at that point and changes in monotonicity. Therefore, the function at that point obtains the

maximum value. The network is discriminated to reach the optimal state,  $V(G, D)$  receives the maximum, at this point:

$$D(x) = \frac{P_{data}(x)}{P_{data}(x) + P_g(x)}. \tag{7}$$

Taking (7) into (3), we obtain the following:

$$\min_G V(G, D) = -2 \log 2 + \min_G \int_x (P_{data}(x) \log \frac{P_{data}(x)}{(P_{data}(x) + P_g(x))/2} + P_g(x) \log(\frac{P_g(x)}{(P_{data}(x) + P_g(x))/2})) dx. \tag{8}$$

The Kullback–Leibler divergence is calculated as

$$D_{KL}(p||q) = \int_x p(x) \log \frac{p(x)}{q(x)} dx. \tag{9}$$

Constructing the KL scatter from Equation (8) [13], we obtain

$$\min_G V(G, D) = -2 \log 2 + \min_G [D_{KL}(P_{data}(x) || \frac{P_{data}(x) + P_g(x)}{2}) + D_{KL}(P_g(x) || \frac{P_{data}(x) + P_g(x)}{2})]. \tag{10}$$

For lack of symmetry in KL dispersion, Jensen–Shannon divergence [14] is introduced as follows:

$$D_{JS}(p||q) = \frac{1}{2} D_{KL}(p || \frac{p+q}{2}) + \frac{1}{2} D_{KL}(q || \frac{p+q}{2}) \tag{11}$$

Constructing the JS divergence from Equation (10) gives

$$\min_G V(G, D) = -2 \log 2 + \min_G [D_{JS}(P_{data}(x) || P_g(x))] \tag{12}$$

From the nature of JS divergence, it follows that  $D_{JS}(p || q) \geq 0$ ,  $D_{JS}(p || q) = 0$ ,  $V(G, D)$  obtains the minimum value  $-2 \log 2$ , The state of the generated network at this point is as follows:

$$P_{data}(x) = P_g(x) \tag{13}$$

A network is generated to learn the distribution consistent with the true distribution. We use Generative Adversarial Networks to reach Nash Equilibrium [15].

### 3. CRDCGAN Algorithm

The original GAN is mainly based on full connectivity to implement the generative and discriminant networks [16]. GAN requires a large number of network parameters to generate images. GAN image generation is not effective. Radford et al. [17] proposed Deep Convolutional Generative Adversarial Networks. DCGAN uses transposed convolutional layers to realize the generative network and convolutional layers to realize the discriminative network. Transposed convolution refers to the up-sampling by expanding the image between the inputs via digital padding and then performing the convolution operation to reach an output width and height greater than the input [18].

In this paper, we propose CRDCGAN on the basis of DCGAN. CRDCGAN follows the DCGAN network structure. CRDCGAN improved loss function. CRDCGAN adds additional information as input to the generative and discriminative networks. Adding a residual module to the generative network of CRDCGAN improves the quality of generated images.

#### 3.1. Loss Function Improvements

From (12), it can be seen that  $D_{JS}(p || q)$  is fixed when the two distributions are different, the gradient does not change, and the network parameters cannot be updated. To prevent the above from happening, JS divergence needs to be replaced. Arjovsky et al. [13] proposed

the Wasserstein distance, which denotes the minimum cost of performing a transformation between distributions [19] is defined as follows:

$$W(p_r, q_g) = \inf_{\gamma \in \Pi(p_r, q_g)} E_{(x,y) \sim \gamma} [\|x - y\|]. \quad (14)$$

It is difficult to calculate  $W(p_r, p_g)$  directly via Equation (14). According to Kantorovich–Rubinstein duality [19],

$$W(p_r, p_g) = \sup_{\|f\|_L \leq 1} (E_{x \sim p_r}[f(x)] - E_{y \sim p_g}[f(y)]) \quad (15)$$

The Wasserstein distance between  $p_r$  and  $p_g$  can be transformed into an upper bound for the distribution of functions  $p_r$  and  $p_g$  expectations satisfying the  $K$ -Lipschitz continuum. The Wasserstein distance between the distribution  $p_r$  and  $p_g$  is as follows:

$$W(p_r, p_g) = \frac{1}{K} \sup_{\|f\|_L \leq K} (E_{x \sim p_r}[f(x)] - E_{y \sim p_g}[f(y)]), \quad (16)$$

where  $f$  satisfies  $K$ -Lipschitz continuity, i.e., satisfies  $|f(x_1) - f(x_2)| \leq K|x_1 - x_2|$ .

To ensure that  $f$  satisfies the  $K$ -Lipschitz continuity, it is necessary to restrict the range of values of the linear operator for each layer. In this paper, the range  $(-0.01, 0.01)$  of the literature [19] is followed.

The loss function is finally improved as follows:

$$L = \arg \min_G \max_f E_{x \sim P_{data}(x)}[f(x)] - E_{x \sim P_g(x)}[f(x)]. \quad (17)$$

### 3.2. Joint Condition Information

To solve the problem, DCGAN can only generate images based on random noise and cannot obtain image labels. CGAN-related methods are introduced to add additional image labeling information to the generative and discriminative networks. CGAN delivers additional information to the network as part of the input [20]. The function of generating samples that resemble the real distribution and meet the corresponding conditions is added to CRDCGAN. First, the generative network inputs Gaussian noise and additional information conditions to form a joint representation, turning unsupervised learning into supervised learning. Second, the generative network outputs pseudo-samples that match the labels. Finally, the generative network optimization objective is changed to generate images that match the labels, and the discriminative network objective is changed to discriminate between pseudo-sample–label pairs from the generative network and sample–label pairs from the original data. For the generated sample–label pairs, they are processed and used for classification model training.

### 3.3. Add Residuals Module

In order to improve the quality and increase the diversity of generated images, CRDCGAN introduces the residual block. The residual module adds jump connections between input and output [21] to extract deep feature information from the feature map. This way, the model can automatically choose whether to complete the feature or characteristic transformation via the convolutional layer or skip the convolutional layer directly [22]. Therefore, adding a residual module to the generative network also increases the stability of the network training [23]. The residual module is added to the generative network, and the output dimension is not changed.

The following is an example of generating a 32-pixel  $\times$  32-pixel RGB image. The construction of CRDCGAN generative networks based on conditional information and the deep residual module is shown in Figure 1.

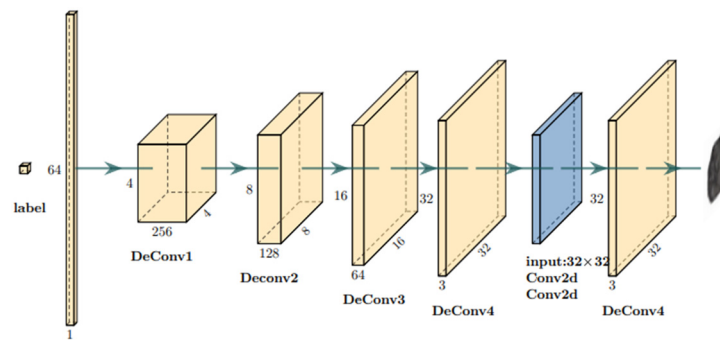


Figure 1. CRDCGAN generative network.

The generating network takes as input a random vector of dimension 64 obeying Gaussian distribution and category labels. It is mapped to 256 4 × 4 feature maps via transposed convolution. A tensor of 32 × 32 × 3 is obtained via 4 transposed convolution operations. The residual module is added to it, and the output is kept unchanged to obtain the pseudo-sample matching the category label and output a 32-pixel × 32-pixel 3-channel image.

The construction of the CRDCGAN discriminative network is shown in Figure 2.

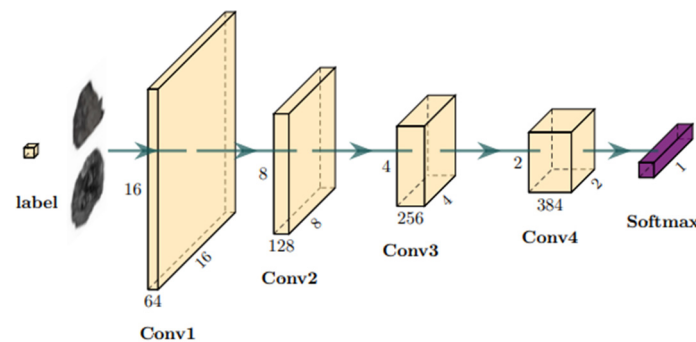


Figure 2. CRDCGAN discriminative network.

Discriminative networks are essentially classifiers for classification tasks. The discriminative network inputs 32 × 32 × 3 image data containing label information and obtains a 2 × 2 × 384 3-dimensional tensors by 4 convolutional layers. The activation functions of the first, second, and third convolutional layers using LeakyRelu. The convolution layer extracts the characteristics, and the fully connected layer transforms the characteristic vector into a 1-dimensional tensor to obtain the probability that the image is a true sample–label pair.

#### 4. Experimental Procedure and Analysis

##### 4.1. Evaluation Indicators

In this paper, we study CRDCGAN to generate rock image data with labels, which are processed and used for image classification, whether the classification performance of the network has improved or not. The key is the accuracy of the classification task. The classification accuracy rate is calculated as follows:

$$acc = \frac{TP + FP}{TP + TN + FP + FN} \tag{18}$$

There is an obvious drawback of accuracy: the categories that account for a larger proportion of the overall have a greater impact on accuracy than the other categories. For this purpose, precision, recall, and  $F_1$  scores are introduced through the confusion matrix for a comprehensive evaluation of the network.

Precision, recall, and  $F_1$  scores are calculated as follows:

$$P = \frac{TP}{TP + FP} \tag{19}$$

$$R = \frac{TP}{TP + FN} \tag{20}$$

$$F_1 = \frac{2PR}{P + R} \tag{21}$$

$TP$  denotes the number of positive samples predicted correctly,  $TN$  denotes the number of negative samples predicted correctly,  $FP$  denotes the number of negative samples predicted incorrectly, the precision rate  $P$  is the proportion of correctly predicted positive samples to all predicted positive samples,  $FN$  is the number of positive samples predicted incorrectly, the recall rate  $R$  is the proportion of correctly predicted positive samples to all positive samples, and the  $F_1$  score is used as a way to indicate the accuracy of the model for the classification task [24].

#### 4.2. Experimental Environment Configuration

The experimental environment in this paper is Windows 10 OS, CPU model is I7-7700HQ, memory is 16 G, Python language is used for implementation, deep learning framework is TensorFlow 2.4.1 +CUDA 11.0.221+ CUDNN8.1.1, and NVIDIA GTX 1050 GPU is used for acceleration.

#### 4.3. Experimental Procedure

The experimental flow is shown in Figure 3.

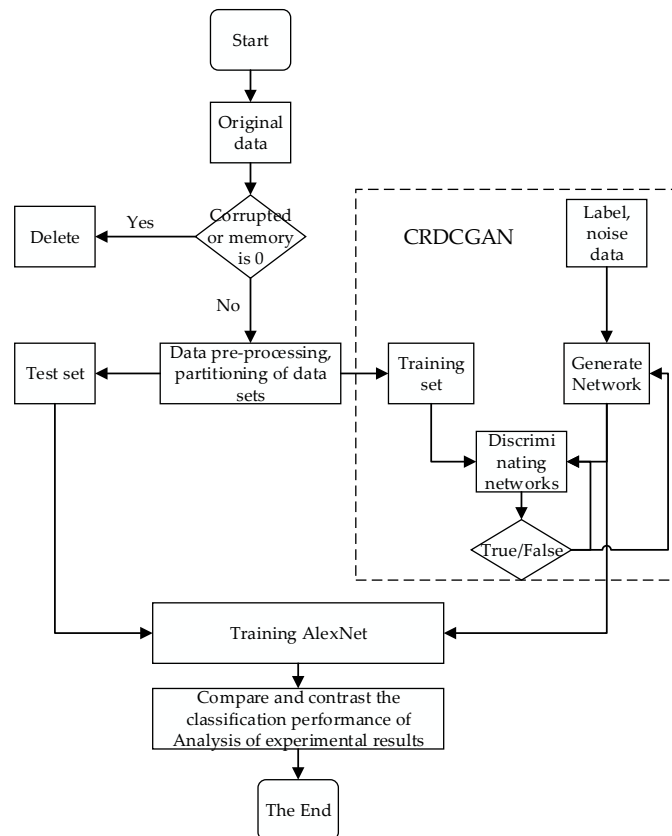


Figure 3. Image data augmentation applied to rock classification experimental process.

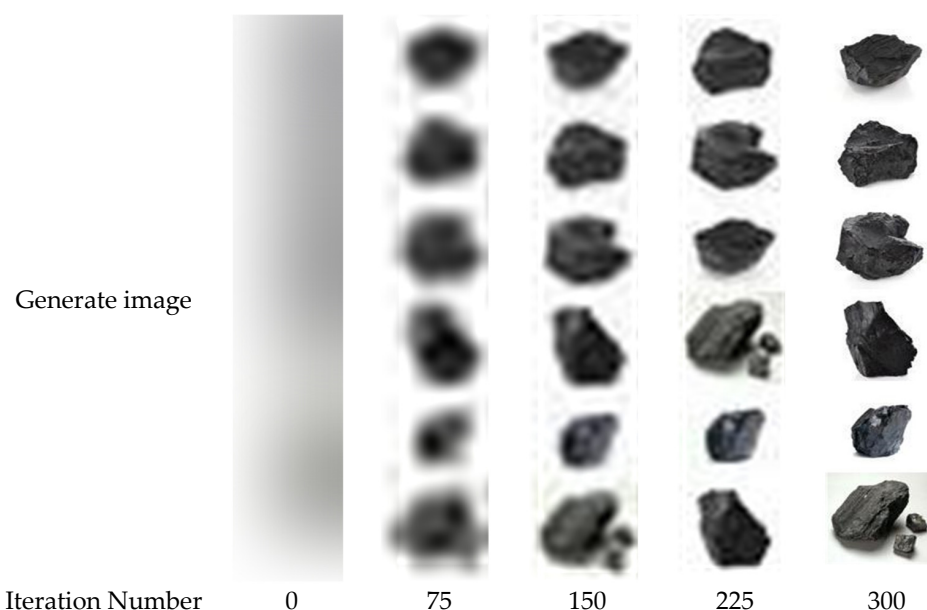
#### 4.4. Experimental Procedure and Analysis of Results

Rock image data is collected on the Internet using crawler technology. The corrupted and memory-occupied image data of the collected rock image data was eliminated. A total of 4173 images of various types of rocks remained, including 172 of basalt, 739 of coal, 203 of granite, 677 of limestone, 775 of marble, 956 of quartzite, and 651 of sandstone. We remove the image watermarks and cropped non-major parts of the images. Image size varies, image sizes larger than 32 pixels are scaled, and image sizes smaller than 32 pixels are expanded. All types of images are read, a table of digital codes is created, all subfolders under the root directory are traversed, the mapping relationship is fixed, and each category corresponds to a digital code. Corresponding codes and image locations are saved in a file, and the file is read to obtain image paths and corresponding labels. The images are read into a (32, 32, 3) tensor form, and the dimensions are added before axis = 0 for tensor stacking to obtain the rock image dataset. The images of each class are divided according to the ratio of 4:1 between the training set and the test set. The uneven distribution of samples in the rock image dataset, the large image noise, and the different number of samples contained in a single image led to the difficulty of the rock image classification task and the difficulty of model training. The rock image dataset (from left to right, basalt, coal, granite, limestone, marble, quartzite, and sandstone) is shown in Figure 4.



**Figure 4.** Rock images.

Training CRDCGAN, the optimizer uses Adam. The learning rate of 0.0001, the batch\_size size of 32, and 300 iterations of training for the generative and discriminative networks were determined through several experiments. The generative network of CRDCGAN is prevented from learning the features of the test set in advance, generating images containing the same features as the test set, and only the rock image training set is used for the training of CRDCGAN. The control generating network input label = 1 generates coal, and the training process generates images presented in a  $6 \times 1$  specification, as shown in Figure 5.



**Figure 5.** CRDCGAN generates a coal image training process.

As observed in Figure 5, at the 75th iteration, CRDCGAN initially has the generation capability, and the generated coal images are blurred with obvious noise; at the 150th iteration, the image generation capability is greatly improved, and the image quality is obviously improved; at the 225th iteration, the network generation capability is basically stable, and the generated images are clear; at the 300th iteration, the image generation capability tends to a steady state, and the generated images are clear in detail, without pattern collapse and rich in diversity.

The control labels generate basalt, granite, limestone, marble, quartzite, and sandstone, as shown in Figure 6.



**Figure 6.** CRDCGAN generates rock images.

The CRDCGAN generation network that has completed training is fed with label information and noisy data to generate rock image data with label information. The original dataset is expanded using the generated images. To verify whether the effect of using CRDCGAN on rock images can be improved, experimental validation is performed using AlexNet. This paper focuses on data enhancement without using data enhancement methods on the data, which is directly used for network training, data enhancement using traditional methods of affine transformation (rotation, cropping, flipping, increasing brightness, and random erasing) methods, and data enhancement using DCGAN methods to expand the dataset. The original image, the image generated via a traditional affine transformation, the image generated via DCGAN, and the image generated via CRDCGAN are shown respectively.

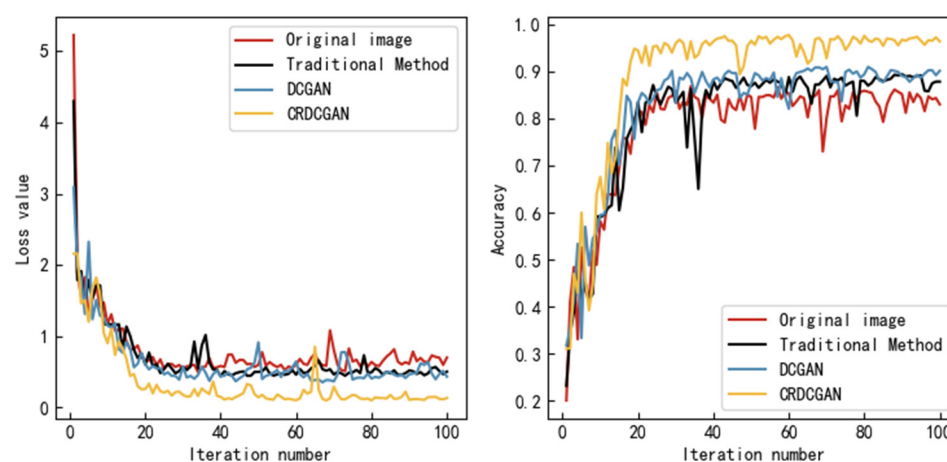
From Figure 7, it can be seen that the images of coal enhanced via the traditional method do not differ much from the original images, and there is redundancy in feature information; the images generated via DCGAN are relatively single and blurred; the images generated via CRDCGAN are clearly visible, rich in details with obvious features. The traditional method, DCGAN method, and CRDCGAN method are used to expand the rock image dataset, respectively, and the classification model AlexNet is trained on the same test



set with 100 iterations. The loss value and accuracy variation curves of the classification model are shown in Figure 8.



**Figure 7.** Different data augmentation methods to generate images of coal.

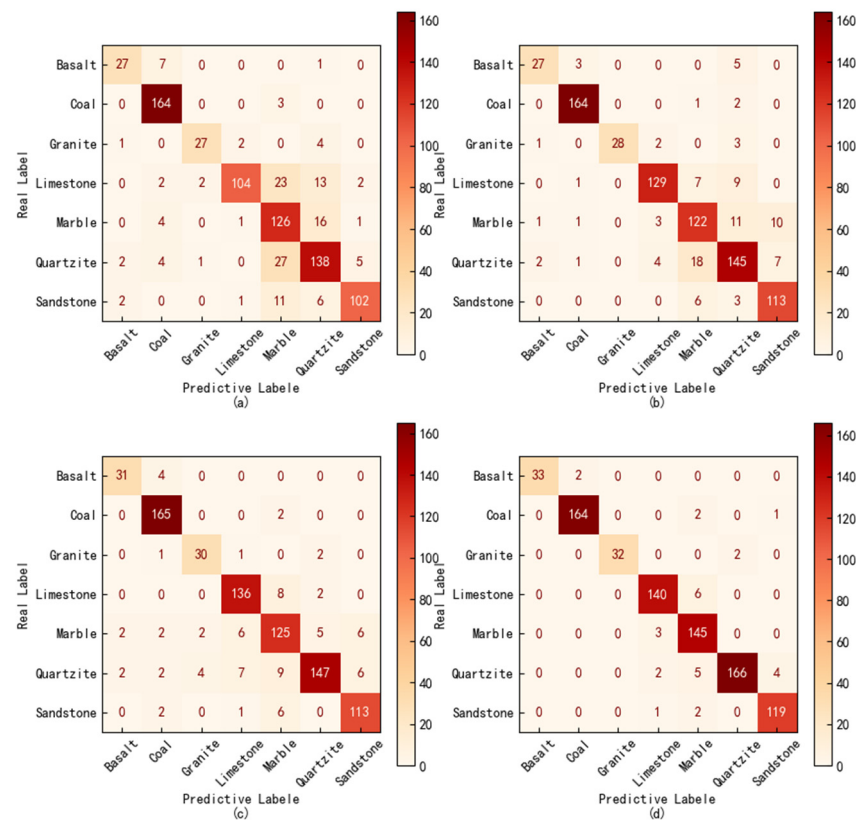


**Figure 8.** Rock image classification model loss value accuracy change curve.

As seen in Figure 8, after 100 iterations, the classification accuracy of the proposed method CRDCGAN reaches 96.38%, which is 13.39%, 8.57%, and 6.28% higher than that of the classification models trained using the non-augmented, traditional data augmentation methods and DCGAN data augmentation methods, respectively, on the same test set. The classification accuracy of the network trained via the CRDCGAN-augmented dataset is less volatile and stable at around 0.13. The loss value of the network trained via CRDCGAN with the expanded dataset fluctuates less and is stable at about 0.13. The CRDCGAN method is more stable than the other methods and achieves higher classification accuracy.

The advantages and disadvantages of classification models are analyzed in depth using confusion matrices [25]. The confusion matrix can visualize the classification of each type of rock. The confusion matrix is shown in Figure 9.

figfig:applsci-2335948-f009 visualizes that the confusion matrix of the classification model trained via CRDCGAN enhancement has significantly larger values at the diagonal line, which also reflects a greater number of correctly classified images. The four classification models were subjected to performance analysis, and the results are shown in Tables 1–4.



**Figure 9.** Rock classification confusion matrix. (a) Confusion matrix without data augmentation methods. (b) Confusion matrix for traditional data augmentation methods. (c) Confusion matrix for DC-GAN data augmentation methods. (d) Confusion matrix for CRDCGAN data augmentation method.

**Table 1.** Original image test set metrics.

Type	Precision	Recall	F <sub>1</sub> Score	Number
Basalt	0.8438	0.7714	0.8060	27
Coal	0.9061	0.9820	0.9425	164
Granite	0.9000	0.7941	0.8437	27
Limestone	0.9630	0.7123	0.8189	104
Marble	0.6632	0.8514	0.7456	126
Quartzite	0.7753	0.7797	0.7775	138
Sandstone	0.9273	0.8361	0.8793	102

**Table 2.** Traditional augmentation method test set metrics.

Type	Precision	Recall	F <sub>1</sub> Score	Number
Basalt	0.8710	0.7714	0.8182	27
Coal	0.9647	0.9820	0.9733	164
Granite	1.0000	0.8235	0.9032	28
Limestone	0.9348	0.8836	0.9085	129
Marble	0.7922	0.8243	0.8079	122
Quartzite	0.8146	0.8192	0.8169	145
Sandstone	0.8692	0.9262	0.8968	113

**Table 3.** DCGAN augmentation test set metrics.

Type	Precision	Recall	$F_1$ Score	Number
Basalt	0.8857	0.8857	0.8857	31
Coal	0.9375	0.9880	0.9621	165
Granite	0.8333	0.8824	0.8571	30
Limestone	0.9007	0.9315	0.9158	136
Marble	0.8333	0.8446	0.8389	125
Quartzite	0.9432	0.8305	0.8833	147
Sandstone	0.9040	0.9262	0.9150	113

**Table 4.** Augmentation test set metrics.

Type	Precision	Recall	$F_1$ Score	Number
Basalt	1.000	0.9429	0.9706	33
Coal	0.9880	0.9820	0.9850	164
Granite	1.000	0.9411	0.9697	32
Limestone	0.9589	0.9589	0.9589	140
Marble	0.9063	0.9797	0.9416	145
Quartzite	0.9881	0.9379	0.9623	166
Sandstone	0.9597	0.9754	0.9675	119

As seen from Tables 1–4, the method CRDCGAN in this paper achieves more obvious results compared with the traditional image data enhancement method and DCGAN image data enhancement method. The number of correct recognitions for quartzite and marble improved significantly. The data from Tables 1–4 can show that the classification performance of the classification network trained by the method in this paper is better, and the classification accuracy recall and  $F_1$  values are higher for each category in the dataset. The experiments and the passing show that the data enhanced by the method of this paper can achieve the classification model with better classification performance.

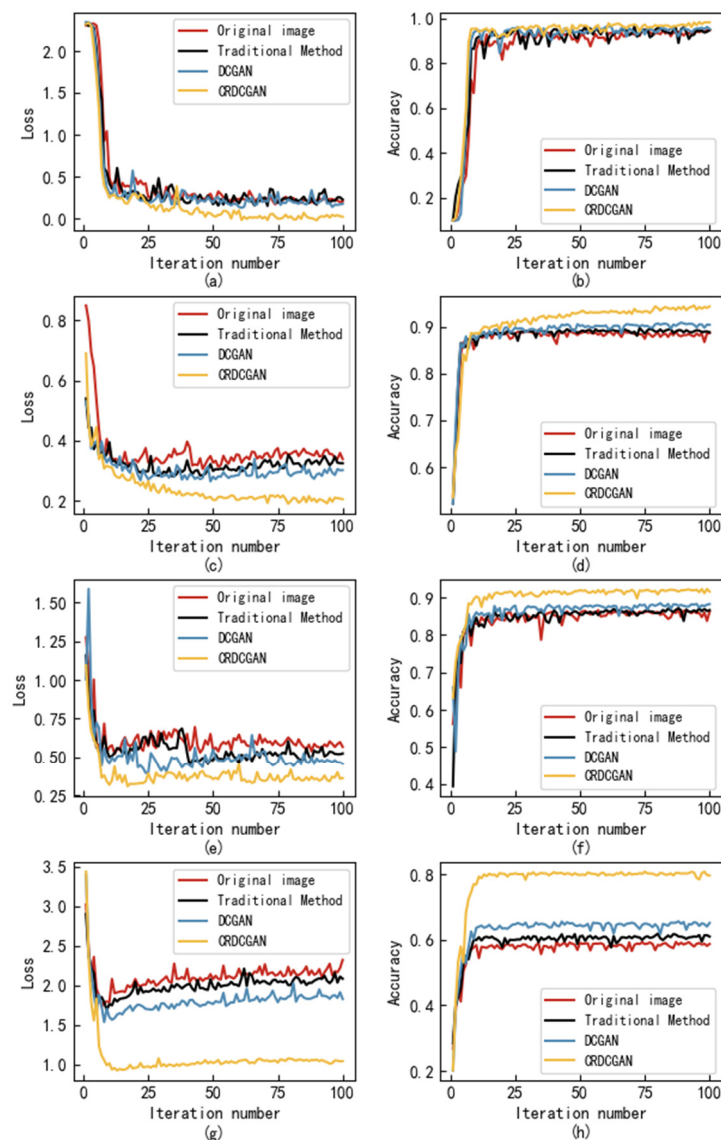
#### 4.5. Experimental Comparison of Different Data Augmentation Methods on Public Datasets

In order to verify the effectiveness of the proposed method in this paper, the method is compared experimentally with no data augmentation, traditional method data augmentation, and DCGAN method data augmentation applied to HWDB10 dataset, FashionMnist dataset, Cifar-10 dataset, and Cifar-100 dataset to complete handwritten Chinese character classification, fashion classification, and physical classification. The control experiments are shown in Figure 10.

Figure 10a,b indicate the loss value and accuracy of the HWDB10 data, respectively. The loss values and accuracy rates on the FashionMnist dataset, Cifar-10 dataset, and Cifar-100 dataset are shown in the following order.

From Figure 10, it can be seen that the CRDCGAN data augmentation method in this paper can positively promote the classification performance of the model and has a large improvement on the Cifar-10 dataset, with a classification accuracy of 79.54%, which is much higher than the other 3 methods of 58.68%, 60.64%, and 65.11%. On the HWDB10 dataset, this paper's method has only a 3.49% improvement over that without the data augmentation method. The reason for the accuracy improvement of only 3.49% is that the AlexNet classification model has 94.68% classification accuracy in this dataset without data augmentation. The performance on the rock sub-image dataset, the HWDB10 dataset, and the Cifar-100 dataset fully illustrate that the classification accuracy of the classification model is improved more significantly by using this paper's method in the difficult classification task. The rock image classification is a 7-classification task, the Cifar-10 dataset is a 10-classification task, and the Cifar-100 dataset is a 100-classification task. Through these three sets of experiments, it can be shown that the method of this paper can

be applied to a variety of multi-classification tasks. The method enhanced the effect on both color and black-and-white images.



**Figure 10.** Comparison of classification model loss value accuracy change curves. (a) Loss value on HWDB10. (b) Accuracy on HWDB10. (c) Loss value on FashionMnist. (d) Accuracy on FashionMnist. (e) Loss value on Cifar-10. (f) Accuracy on Cifar-10. (g) Loss value on Cifar-100. (h) Accuracy on Cifar-100.

The above experiments verify the effectiveness of the proposed method in this paper, which can enhance the data and thus improve the accuracy of the classification model. It also shows that the method in this paper has some general significance and can improve the accuracy of rock image classification.

## 5. Conclusions

In this paper, an image data augmentation method CRDCGAN is given. CRDCGAN is derived from the machine learning algorithm DCGAN and improved on DCGAN. CRDCGAN generates rock image data, obeying the same distribution as the original image data with labels, which can be directly applied to classification network model training. The augmented image dataset was used for rock image classification, and the classification accuracy reached 96.38%. Compared with [9], this method generates clearer images with

less redundant information, which is suitable for datasets containing multiple rock images, and improves the model accuracy while enhancing the model generalization ability. Compared with [11], the rock image data generated by this method can effectively improve the accuracy of the classification model. Compared with [10], this paper uses a relatively weak classification model to achieve higher classification accuracy. The comparison with the above literature proves the effectiveness of our method.

Our next work will continue to investigate the impact of data expansion size on the classification accuracy of the model.

**Author Contributions:** Conceptualization, G.Z.; methodology, Z.C.; software, Z.C. and X.W.; validation, X.W.; investigation, Z.C.; resources, Z.C.; data curation, G.Z.; writing—original draft preparation, X.W.; writing—review and editing, Z.C.; supervision, X.W. and G.Z.; funding acquisition, X.D. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by the National Science Foundation of China grant number [42271309].

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Acknowledgments:** The authors thank the National Science Foundation of China for their support (No. 42271309).

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Murdie, R. Geological Survey of Western Australia: Geological Survey of Western Australia's Accelerated Geoscience Program. *Preview* **2021**, *2021*, 24–25. [[CrossRef](#)]
2. Chen, C.; Chen, L.; Huang, Z. Study on Support Design and Parameter Optimization of Broken Soft Large-Section Roadway at High Altitude. *Min. Res. Dev.* **2022**, *42*, 88–94. [[CrossRef](#)]
3. Wang, W.; Li, Q.; Zhang, D.; Li, H.; Wang, H. A survey of ore image processing based on deep learning. *Chin. J. Eng.* **2023**, *45*, 621–631. [[CrossRef](#)]
4. Xu, S.; Zhou, Y. Artificial intelligence identification of ore minerals under microscope based on deep learning algorithm. *Acta Petrol. Sin.* **2018**, *34*, 3244–3252.
5. Li, C.; Zhang, X.; Zhu, H.; Zhang, M. Research on Dangerous Behavior Identification Method Based on Transfer Learning. *Sci. Technol. Eng.* **2019**, *19*, 187–192.
6. Wu, W.; Qi, Q.; Yu, X. Deep learning-based data privacy protection in software-defined industrial networking. *Comput. Electr. Eng.* **2023**, *106*, 108578. [[CrossRef](#)]
7. Pu, Y.; Apel, D.B.; Szmigiel, A.; Chen, J. Image Recognition of Coal and Coal Gangue Using a Convolutional Neural Network and Transfer Learning. *Energies* **2019**, *12*, 1735. [[CrossRef](#)]
8. Tian, Z.; Wang, M.; Wu, J.; Gui, W.; Wang, W. Mine Image Enhancement Algorithm Based on Dual Domain Decomposition. *Acta Photonica Sin.* **2019**, *48*, 107–119.
9. Hong, H.; Zheng, L.; Zhu, J.; Pan, S.; Zhou, K. Automatic Recognition of Coal and Gangue based on Convolution Neural Network. *Coal Eng.* **2017**, *49*, 30–34.
10. Baraboshkin, E.E.; Ismailova, L.S.; Orlov, D.M.; Zhukovskaya, E.A.; Kalmykov, G.A.; Khotylev, O.V.; Baraboshkin, E.Y.; Koroteev, D.A. Deep convolutions for in-depth automated rock typing. *Comput. Geosci.* **2019**, *135*, 104330. [[CrossRef](#)]
11. Cheng, G.; Zhang, F. Super-resolution Reconstruction of Rock Slice Image Based on SinGAN. *J. Xi'an Shiyou Univ. (Nat. Sci. Ed.)* **2021**, *36*, 116–121.
12. Goodfellow, I.J.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative Adversarial Nets. In Proceedings of the 27th International Conference on Neural Information Processing Systems, Cambridge, MA, USA, 8–13 December 2014; pp. 2672–2680.
13. Liang, H.; Lu, H.; Feng, K.; Liu, Y.; Li, J.; Meng, L. Application of the improved NOFRFs weighted contribution rate based on KL divergence to rotor rub-impact. *Nonlinear Dyn.* **2021**, *104*, 3937–3954. [[CrossRef](#)]
14. Liu, C.; Zhao, J.; Sun, N.; Yang, Q.; Wang, L. IT-SVO: Improved Semi-Direct Monocular Visual Odometry Combined with JS Divergence in Restricted Mobile Devices. *Sensors* **2021**, *21*, 2025. [[CrossRef](#)] [[PubMed](#)]
15. Sun, Y.; Song, D. Research on resource allocation strategy of group robot system. *J. Xi'an Univ. Sci. Technol.* **2022**, *42*, 818–825. [[CrossRef](#)]
16. Zhang, E.; Gu, G.; Zhao, C.; Zhao, Z. Research progress on generative adversarial network. *Appl. Res. Comput.* **2021**, *38*, 968–974. [[CrossRef](#)]

17. Randorf, A.; Metz, L.; Chintala, S. Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks. *arXiv* **2015**. [[CrossRef](#)]
18. Liu, Q.; Yu, B.; Meng, X.; Zhang, X. Pavement Crack Recognition Algorithm Based on Transposed Convolutional Neural Network. *J. South China Univ. Technol. (Nat. Sci. Ed.)* **2021**, *49*, 124–132.
19. Arjovsky, M.; Chintala, B.; Bottou, L. Wasserstein Generative Adversarial Networks. In Proceedings of the 34th International Conference on Machine Learning, International Convention Centre, Sydney, NSW, Australia, 6–11 August 2017.
20. Mirza, M.; Osindero, S. Conditional generative adversarial nets. *arXiv* **2014**. [[CrossRef](#)]
21. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual learning for image recognition. In Proceedings of the 2016 IEEE Conference on Computer Vision and pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 700–778.
22. Zhang, X.; Li, L.; Di, D.; Wang, J.; Chen, G.; Jing, W.; Emam, M. SERNet: Squeeze and Excitation Residual Network for Semantic Segmentation of High-Resolution Remote Sensing Images. *Remote Sens.* **2022**, *14*, 4770. [[CrossRef](#)]
23. Wang, J.; Tang, L.; Wang, C.; Zhu, R.; Dong, R.; Zheng, L.; Sha, W.; Huang, L.; Li, P.; Weng, S. Multi-scale convolution neural network with residual modules for determination of drugs in human hair using surface-enhanced Raman spectroscopy with a gold nanorod film self-assembled by inverted evaporation. *Spectrochim. Acta Part A Mol. Biomol. Spectrosc.* **2022**, *280*, 121463. [[CrossRef](#)]
24. Wang, Z.; Zhang, K.; Zhang, H. Miner's emotion recognition based on deep wise separable convolution neural network miniXception. *J. Xi'an Univ. Sci. Technol.* **2022**, *42*, 562–571.
25. Song, Y.; Wang, X.; Lei, L. Evaluating evidence reliability based on confusion matrix. *Syst. Eng. Electron.* **2015**, *37*, 974–978. [[CrossRef](#)]

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.