



# Article Rock CT Image Super-Resolution Using Residual Dual-Channel Attention Generative Adversarial Network

Liqun Shan <sup>1,2,\*</sup>, Chengqian Liu <sup>1</sup>, Yanchang Liu <sup>1</sup>, Weifang Kong <sup>1</sup> and Xiali Hei <sup>2</sup>

- <sup>1</sup> School of Physical and Electrical Engineering, Northeast Petroleum University, Daqing 163318, China; chengqian\_liu0302@163.com (C.L.); ychang\_liu@nepu.edu.cn (Y.L.); weifang\_kong@163.com (W.K.)
- <sup>2</sup> School of Computing and Informatics, University of Louisiana at Lafayette, Lafayette, LA 70503, USA; xiali.hei@louisiana.edu
- \* Correspondence: liqun.shan.liu@gmail.com

Abstract: Because of its benefits in terms of high speed, non-destructiveness, and three-dimensionality, as well as ease of integration with computer simulation, computed tomography (CT) technology is widely applied in reservoir geology research. However, rock imaging is restricted by the device used as there is not a win-win for both the image receptive field and corresponding resolution. Convolutional neural network-based super-resolution reconstruction has become a hot topic in improving the performance of CT images. With the help of a convolution kernel, it can effectively extract characteristics and ignore disturbance information. The dismal truth is that convolutional neural networks still have numerous issues, particularly unclear texture details. To address these challenges, a generative adversarial network (RDCA-SRGAN) was designed to improve rock CT image resolution using the combination of residual learning and a dual-channel attention mechanism. Specifically, our generator employs residual attention to extract additional features; similarly, the discriminator builds on dual-channel attention and residual learning to distinguish generated contextual information and decrease computational consumption. Quantitative and qualitative analyses demonstrate that the proposed model is superior to earlier advanced frameworks and is capable to constructure visually indistinguishable high-frequency details. The quantitative analysis shows our model contributes the highest value of structural similarity, enriching the more detailed texture information. From the qualitative analysis, in enlarged details of the reconstructed images, the edges of the images generated by the RDCA-SRGAN can be shown to be clearer and sharper. Our model not only performs well in subtle coal cracks but also enriches more dissolved carbonate and carbon minerals. The RDCA-SRGAN has substantially enhanced the reconstructed image resolution and our model has great potential to be used in geomorphological study and exploration.

**Keywords:** rock CT images; super-resolution; convolutional neural networks; residual learning; generative adversarial network; channel attention mechanism

# 1. Introduction

In the field of geological petroleum, limited by the detector hardware in the process of acquiring rock imaging, it is challenging to make a win–win for both the receptive field and its resolution, which means the reservoir characteristics cannot be clearly obtained [1–3]. In this case, X-ray computed tomography (CT) technology is commonly used in reservoir geology studies due to its fast, non-destructive, three-dimensional (3D) characterization and ease of integration with computer simulations [4–6]. High-quality and clear CT images will provide strong support for basic geological research and oil exploration. Therefore, maintaining a higher resolution has always been one of the prospects of CT geological exploration.

Super-resolution (SR) reconstruction is the process of converting images from low resolution (LR) to high resolution (HR). Numerous resolution reconstruction models have



Citation: Shan, L.; Liu, C.; Liu, Y.; Kong, W.; Hei, X. Rock CT Image Super-Resolution Using Residual Dual-Channel Attention Generative Adversarial Network. *Energies* **2022**, *15*, 5115. https://doi.org/10.3390/ en15145115

Academic Editor: Manoj Khandelwal

Received: 9 June 2022 Accepted: 5 July 2022 Published: 13 July 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 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/). been proposed for everything from satellite monitoring images and media images to geological prospecting images [7–10].

Single SR image reconstruction roughly includes three basic types of methods: interpolationbased, reconstruction-based, and learning-based [11–13]. The first two types of methods have a strong dependence on the training samples, and the training models are subject to extrinsic factors. One of the learning-based algorithms, A+ [14], considers the dictionary as a tool for recording the connection between LR and HR image patches ahead of time by training datasets. The A+ method focuses on faster generating SR images from input LR images with the help of the dictionary. Another more powerful learning-based approach, deep learning technology, can greatly help CT images out of the trouble of damaged rock image quality, resulting from the inherent limitations of CT devices. Based on deep learning, CT images have further enriched texture features that provide theoretical guidance and implementation methods for flow simulation and the quantitative characterization of rock micro-pore and micro-throat structure [15].

Convolutional neural network (CNN) reconstruction methods have caught researchers' attention [16–21]. Dong et al. [16] presented a method for generating HR images using CNN, called SRCNN, which firstly made a remarkable breakthrough for image SR reconstruction. Soon afterward, the SRCNN structure was optimized and proposed fast super-resolution using a convolution neural network (FSRCNN) [7]. Kim et al. [22] used deep-recursive convolution to implement SR image reconstruction, which enables the CNN model to be divorced from cumbersome parameters and obtain a larger receptive field. The deeprecursive convolution operations do further improve the network performance, but there is also the problem of how to deal with the efficiency of convolution calculating. Shi et al. [23] proposed an SR model based on an efficient sub-pixel convolution network (ESPCN), which efficiently obtains HR images by directly calculating convolution on LR images, but, at a larger scale factor, the obtained SR images are too smooth and lack realism. There is the concern that deeper networks for SR images are more difficult to train. Fortunately, He et al. [24] proposed a residual network (ResNet), which solves the gradient disappearance problems resulting from too many layers. Inspired by the residual network, a very deep convolutional network for super-resolution (VDSR) was proposed [25] and could further improve the model training efficiency and reduce parameters without damaging the model performance. However, there are still problems with convergence and gradient transmission. In order to further train deeper network layers with as few calculations as possible and still obtain higher accuracy, an enhanced deep residual networks (EDSR) algorithm was proposed [8]; mainly by eliminating unnecessary batch normalization (BN) and rectified linear unit (ReLU), it saves 40% of memory consumption. In this way, it greatly reduces the complexity of the model, as a deeper network structure can be trained under the same computing resources. Although CNNs are extensively used to build SR networks, CNN-based networks are confronted with some challenges, where the generated image is too smooth and the high-frequency details are lost. That means that the accuracy of the super score image does not match people's expected perception.

Compared with CNN-based networks, generative adversarial network (GAN)-based networks focus on detailed texture features, ameliorating adversarial loss and content loss functions, and generating as realistic images as possible [26–33]. GAN consists of a generator and a discriminator. The generator does not try to generate a fake input until the discriminator is enough to identify it as real [34]. Ledig et al. [30] put forward an SR reconstruction algorithm based on generative adversarial networks (SRGAN), which achieved profound results. Stimulated by SRGAN, GAN-based networks have been employed to generate more realistic and highly textured images. Although SRGAN has hugely improved the facticity of texture, there are still disparities between the generated images and prospective images. Wang et al. [35] reorganized enhanced super-resolution generative adversarial networks (ESRGAN); both elevated generator network and better perception loss were proposed, the relative value was also judged rather than the absolute value by the discriminator, leading to restoring more texture details. However, considering SR images

generated by GAN was still impractical, Wang et al. (2018) [13] used conditional generative adversarial network (Conditional GAN) to synthesize HR realistic images from semantic label maps. You et al. (2019) [36] conducted feature extraction and restoration from noisy LR input images by integrating deep CNN, residual learning, and network-in-network techniques (GAN-CIRCLE). GAN-CIRCLE makes an effort to transform the produced features into being real to the human vision, and in addition, the output of the hidden layer is optimized by using parallel  $1 \times 1$  CNN, which compresses the depth of the network to quicken training.

From the aforementioned network, CNN-based models occupy the areas where the SR images are not provided with multiple details, while GAN-based are weakening CNNs' disadvantages mainly through the interdependent generator and discriminator. There are also thorny problems for GAN-based formations, such as too much time-consuming. The generator and discriminator strongly count on each other and need coincident synchronization, resulting in discriminator convergence and generator divergence. Except for used CNNs and GANs models, a new idea is that the residual channel attention block (RCAB) is posted in network structures to enhance feature detail capture. RCAB is replenished in SR reconstruction, which excels at discriminating generated characteristics, as well as being integrated into the end-to-end deep network for training. Tai et al. (2017) [37] proposed a method of local residual connection and recursive learning of residual units (DDRN). A deep residual channel attention network model (RCAN) was put forward to speed up training [38]. To get rid of sophisticated parameters, long and short hop connections are applied to transmit rich low-frequency characteristics, and the RCAN method succeeds in learning more detailed features. Meanwhile, the channel attention (CA) mechanism is capable of adaptively rescaling the features by taking advantage of the inter-channel interdependence. Wang et al. [39] developed enhanced deep super-resolution generative adversarial networks (EDSRGAN) to generate SR images for sandstone, coal, and carbonate samples. It is worth noting that SRCNN restores large-scale edge features, while EDSRGAN better regenerates perceptually indistinguishable high-frequency textures. Shan et al. [40] presented a method named CA-SRResNet combining residual learning and attention mechanism based on convolution networks to capture more texture features of rock CT images. In contrast to other existing algorithm models, despite better the reconstruction effect produced by CA-SRResNet, its network sill can be ameliorated based on GAN in a new perspective. As mentioned above, the attempt of GAN-based networks will be more likely to be coincident between human visual effects and generated CT images. To the authors' best knowledge, both residual learning and dual-channel attention mechanism have not been fused into a generative adversarial network for rock CT images. Based on this, with the help of the above blocks, our model brings first-rate results, making a trade-off made between pixels and calculation.

The remaining sections of this paper are divided into the following sections. Section 2 puts forward a few critical blocks (Residual block, Sub-pixel convolution block, Channel attention mechanism block, and Residual channel attention block) and meanwhile describes the proposed GAN framework with the combination of residual learning and dual-channel attention mechanism (RDCA-SRGAN) in detail. Section 3 will go through how to conduct the experiments, covering the setting up, training procedure, and experimental steps. Further, we compare and evaluate our RDCA-SRGAN network with nine existing approaches. Results indicate that our model outperforms other cutting-edge network infrastructures. Section 4 discusses the research directions in the future and Section 5 summarizes the study's findings.

#### 2. Methodology

#### 2.1. Residual Block

Applying more convolutional layers can broaden the receptive field, but if the layers are piled sequentially, it will result in vanishing or exploding gradients and deterioration. Adding more layers, on the other hand, will fall into the trouble of inappropriate training

precision and more time cost. He et al. [24] presented ResNet, as shown in Figure 1a; a residual learning framework that matches a residual mapping rather than the entire mapping. Residual learning has been extensively added to feature extraction modules to strengthen network training capabilities [8,28,38,41].



**Figure 1.** Original and modified residual blocks: (a) original residual block; (b) modified residual block.

For traditional residual blocks, batch normalization (BN) generally performs direct normalization operations on each batch characteristic of the incoming feature maps and recovers the original input by operations such as stretching, scaling, and transformation. The BN operation not only regularizes parameters but also implements faster convergence, dealing with matters of gradient explosion in the process of model training. However, in the case of image SR reconstruction, the BN layer changes the value of the initial source and limits structural adaptability. The image's hue, sharpness, and saturation will be rescaled after passing through the BN stage, impacting the final reconstruction outcome [8]. In our study, we modified the original residual blocks by properly adjusting the BN layer positions to take advantage of the good side of BN.

As is shown in Figure 1b, the structural diagram of the original residual block is modified, the first BN layer in Figure 1a is removed from the original residual block, but the second BN layer is kept. In the modified residual block (Figure 1b), each convolutional layer owns a  $3 \times 3$  kernel, a  $2 \times 2$  stride, and 64 channels. In particular, the first convolutional layer is followed by an exponential linear unit (ELU) activation function. Leaving out unnecessary blocks, the modified structural diagram of the residual block not only enhances the reconstruction performance but also preserves a large amount of memory and decreases the burden of computation.

## 2.2. Sub-Pixel Convolution Module

Following the deep residual module, an up-sampling operation is required for reconstructing images. Deconvolution, direct up-sampling, and bilinear interpolation are the most popular approaches for amplifying the size for SR reconstruction based on deep learning. In this study, sub-pixel convolution is adopted, which can better suit the connection between pixels [21]. Sub-pixel convolution, known as pixel shuffle, is an effective approach to expanding image feature mapping. Figure 2 depicts the realization principle.



Figure 2. Sub-pixel convolution procedure.

The subpixel convolution module, as shown in Figure 3, increases the extracted features with an additional number of channels which is equal to the square of the enlargement size. When different pixel features are convolved in a specified way, a comprehensive feature map is created, which is called pixel cleaning.



Figure 3. Sub-pixel convolution module.

### 2.3. Channel Attention Mechanism

CNN-based methods treat channel characteristics in the same way, resulting in a lack of flexibility for different types of information. Researchers have worked hard to get out of this thorny problem [38,42]. A soft attention mechanism is introduced to address this issue. It is usually divided into channel attention and spatial attention. Hu et al. [43] established the concept of channel attention to emphasize the correlations between the model's distinct feature channels. Models based on channel attention can automatically acquire characteristics for each feature channel. Each feature channel is allocated various weight factors in terms of improving extracted feature details while suppressing unessential image features.

Each input color image is represented by three channels (R, G, B) in CNN. After traveling through various convolution kernels, each channel can produce new information. One feature of information is divided into numerous channel feature components. The attention operation adds each calculated weight value to the related channel information. Figure 4 depicts the channel attention mechanism in action when the number of input channels is four. Just as in Figure 4, W and H denote the image's width and height, respectively; C stands for the channel, representing the number of convolution kernels; r means the reduction ratio.



Figure 4. Schematic diagram of channel attention mechanism.

With the help of channel attention, the network promotes discriminative learning capacity and concentrates more on valuable channels. The channel attention mechanism is made up of three main processes: squeeze, excitation, and attention. First, with the squeeze process, the summary statistics of each channel are obtained by global pooling, eigenvalues added, and averaging. Then, the excitation part learns the relationship between the channels through the  $1 \times 1$  convolution and rectified linear unit (ReLU) activation function. Ultimately, the attention part uses the new correlation to weight the different input feature channels. Global pooling, as in Equation (1), is used to generate summary statistics for each channel and obtain the global receptive field on each channel.

$$a_{c} = F_{c}(x_{c}) = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{W} x_{c}(i, j)$$
(1)

where  $a_c$  is the element value of the cth channel,  $x_c$  denotes the position of the *c*th channel in the input feature,  $F_c(\cdot)$  is the global average pooling, and W and H represent the input image's width and height, respectively.

The excitation part conducts down-sampling and up-sampling through two  $1 \times 1$  convolutions. Down-sampling reduces the number of feature map channels to C/r, and passes features to an activation function layer. Up-sampling makes the feature map back to C. The weight calculation is expressed as Equation (2).

$$S = f(F_1(\delta(F_2(a))))$$
(2)

where *S* is the channel attention weight,  $f(\cdot)$  denotes the activation function,  $\delta(\cdot)$  is ReLU function,  $F_1(\cdot)$  and  $F_2(\cdot)$  is up-sampling and down-sampling through a 1 × 1 convolution, respectively.

The attention mechanism can multiply the weight coefficients with different input feature values and performs calibration to obtain new features. The calculation is expressed as Equation (3).

$$X_c = S_C \cdot x_c \tag{3}$$

where  $X_c$  denotes the new feature of the output. The output features have the same dimensions as the input features. The channel attention mechanism can assign different weights to different channel features, highlight information-rich features, and suppress non-relevant features.

## 2.4. Residual Channel Attention Block

The adoption of the attention mechanism can effectively increase the network's ability to recover image details and textures as much as possible. Nevertheless, stiffly incorporating the attention block into the core network would significantly impair the network's extraction characteristics, which is not conducive to capturing the image texture and details. The new feature value output is contributed by the attention unit, a weighted product of the input feature value and the attention weight, resulting in the final output being between 0 and 1. As the depth of the network increases and multiple attention units are added, the ultimate feature values will be repeatedly weighted, and eventually, all the features will be at risk of disappearing. Therefore, residual learning is used to sum the input and the weighted features of the attention mechanism as the output of the network [40]. The channel attention block is integrated into the residual block, deploying a new residual dual-channel attention block (RDCAB), as depicted in Figure 5.



Figure 5. Residual dual channel attention block.

Among RDCAB, the residual block is made up of a convolution kernel and a ReLU function, and the first BN layer is deleted. The residual attention unit is accessed in the form of skip connections (as described in Section 2.1, Figure 1) to propagate the input information through the identity mapping, which eliminates the feature disappearance problem brought about by the addition of the attention mechanism. For the sake of effectively calculating the channel attention, the spatial dimension of the input feature needs to be compressed.

For aggregating the spatial features, on the one hand, the popular method is global average pooling (GAP), on the other hand, global maximum pooling (GMP) also can infer unique object characteristics attention on finer channels. Taking this into account, the global average and maximum pooling are used at the same time in our residual dual-channel attention block. The residual dual channel attention block is shown in Figure 5.

# 2.5. Residual Dual Channel Attention Generative Adversarial Network (RDCA-SRGAN)

The GAN consists of two opposing neural networks, namely a generative model and a discriminant model. Just like their names, one generates images, and another discriminates false or real, a trade-off approach for the authenticity of import images and export images. The discriminator and generator are trained together to distinguish the genuine images from the synthetic images. With the discriminator and generator are trained over and over, the discriminator becomes better at identifying fake data, and the generator gets better at producing data. Ultimately, the generated false data is too veritable to be classified as true. For our RDCA-SRGAN, the recombinant model, the residual attention module is used as the generator [40], and the channel attention block is used as a part of the discriminator to construct a dual-channel attention GAN model [44].

# 2.5.1. Generator of GAN

The generator is designed by considering the fundamentals of the residual channel attention mechanism, as shown in Figure 6. The generator network consists of 16 RDCAB modules. Due to the limited number of core CT image samples, a common convolution is added to the input and output for data augmentation. Multiple residual blocks are used for feature extraction. Two subpixel convolutions are used for size upscaling. Apart from that, regular skip connections are used to connect the input and output to ensure the stability of the network.



Figure 6. Architecture of the generator.

# 2.5.2. Discriminator of GAN

The dual residual channel attention units are added to the discriminator, as shown in Figure 7. The discriminator consists of one convolution unit, seven cascaded convolution units, two RDCAB, and two fully connected layers. Before the output of the discriminator, a Sigmoid function is used to determine the output probability value. The first convolution unit is composed of a  $3 \times 3$  convolution with a stride of 1 and the leaky rectified linear unit (Leaky ReLU) activation function. Except for the initial convolution unit, each subsequent convolution unit is broken down into two  $3 \times 3$  convolutional layer using a stride of one or two, one BN layer, and one Leaky ReLU function layer. With a stride of two, the number of feature channels in the convolution kernel is doubled, and a stride of one lessens the size of the feature space by a factor of two. After the fifth and seventh cascaded convolution units, RDCABs with global average pooling are added to obtain non-local feature information. Two fully connected layers, a Leaky ReLU activation function and a Sigmoid activation function are added at the end. Adding two channel attention units to the middle and back



ends of the discriminator not only obtain more context information but also reduce the consumption of computing resources.

Figure 7. Architecture of discriminator.

## 2.5.3. Loss Function

When training the RDCA-SRGAN network, the discriminator will eventually output a value between 1 and 0 for each image, indicating whether the image is real or fake. These probabilities are appropriate when image discretization is classified. Mark the SR image as  $y_{SR} = 0$  and HR image as  $y_{HR} = 1$ . The binary cross-entropy loss function of the discriminator is:

$$BXE_{\rm HR} = -\frac{1}{N} \sum_{i=1}^{N} y_{\rm HR} \log(p(\rm HR))$$
(4)

$$BXE_{SR} = -\frac{1}{N} \sum_{i=1}^{N} y_{SR} \log(p(SR))$$
(5)

where  $BXE_{HR}$  and  $BXE_{SR}$  denote the loss value of the HR image and SR image, respectively. p(HR) represents the probability of identifying the HR image as true. p(SR) refers to the likelihood of judging the SR image as real.

The Adam optimizer is used to optimize the discriminator network with a learning rate of  $1 \times 10^{-5}$ . The classification loss of the real image and the generated image is defined as:

$$L_D = BXE_{\rm HR} + BXE_{\rm SR} \tag{6}$$

where  $L_D$  denotes loss function of the discriminator.

For the generator, most of deep learning SR algorithms use the mean square error (MSE) as the loss function. MSE will result in the generated image to be overly seamless and missing reality in the perceptions. To avoid the problems mentioned above, the losses used by our generator are perceptual loss, content loss, and confrontation loss. The perceptual loss can be obtained by calculating the feature difference between the original HR and the generated SR image using the VGG19 model. The perceptual loss corresponding to the two images  $\phi_{\text{HR}i}$  and  $\phi_{\text{SR}i}$  is expressed as:

$$L_{\rm VGG19} = \sum_{i=0}^{\rm H} \sum_{j=0}^{\rm W} \frac{(\phi_{\rm SRi} - \phi_{\rm HRi})^2}{\rm H \times W}$$
(7)

where  $L_{VGG19}$  denotes loss function of VGG19.

The content loss is defined by the  $L_C$ , and the mathematical expression is shown as Equation (8):

$$L_{\rm C} = \sum_{i=0}^{\rm H} \sum_{j=0}^{\rm W} \frac{|{\rm SR}_i - {\rm HR}_i|}{{\rm H} \times {\rm W}}$$

$$\tag{8}$$

where  $SR_i$  is a SR image and  $HR_i$  represents a HR image.

The generator and the discriminator are coupled together by passing the discriminator output p(SR) as part of the loss function of the generator. The mathematical expression of the adversarial loss  $L_A$  is expressed as:

$$L_A = -\frac{1}{N} \sum_{i=1}^{N} \log(p(\text{SR}))$$
(9)

Therefore, the total loss of the generator  $L_G$  is expressed as:

$$L_G = L_C + \alpha L_{\text{VGG19}} + \beta L_A \tag{10}$$

where  $\alpha$  and  $\beta$  are coefficients that balance different loss items.  $\beta$  represents the weight decay. Recursion is simpler to converge at the beginning of training, so the initial value of  $\alpha$  is made higher to make the training process steady. As the training progresses,  $\alpha$  steadily is consumed to optimize the ultimate output's performance. Comparing different settings in experiments, we find that only when  $\alpha = 1 \times 10^{-3}$  and  $\beta = 5 \times 10^{-5}$  the effect is the best.

# 3. Experiments

The experimental hardware environment in this study is: Intel Core i9-9900KF@3.60 GHz CPU, equipped with two NVIDIA GeForce GTX 1060 GPUs, configured to 32 GB. The software environment is: CUDA Tookit 10.2, using the Pytorch framework, and 64-bit Windows 10 operating system. The SRCNN, A+, FSRCNN, VDSR, DRRN, SRGAN, EDSR, RCAN, ES-RGAN, and our RDCA-SRGAN model are developed and trained with varying parameters.

# 3.1. Experimental Datasets

We implemented the experiments and evaluated nine existing models and the proposed model on the dataset provided by [45], which is composed of 12,000 various rock's CT images of carbonate, sandstone, and coal, with image resolution from 2.7 to 25  $\mu$ m, 500 × 500 unsegmented slices. For three different groups, each set is averagely allocated by 4000 images, which are shuffled and divided into 80% training, 10% validating, and 10% testing sets. The training set and testing set of rock CT images is correspondingly shown in Figures 8 and 9.



Figure 9. Testing set.

The training set consists of three different types of rock CT images, including sandstone, coal, and carbon, with 3200 HR images each, for a total of 9600 images. The remaining 2400 rock images were used for validation and testing. The corresponding size is  $500 \times 500$ . There are no overlapping rock CT images between the training and test sets. To speed up the computation, the images in the training and test sets are randomly cropped to  $96 \times 96$ . The original images are downgraded to the corresponding target images before generating labels on the input photos. The image before down-sampling is called the HR image, and the image after down-sampling is called the LR image. For training SR models, the LR image and the HR image are created as a valid pair. To obtain the desired labels, this work uses bicubic interpolation, striving to generate noisy and blurred LR images with scale factors of 2, 4, and 8. As a result, the corresponding LR images are  $48 \times 48$ ,  $24 \times 24$ , and  $12 \times 12$  pixel size, respectively.

## 3.2. Training Setting

During the experiments, the used parameters are shown in Table 1. As indicated in the table, to train the networks, it is configured as follows: the initial learning rate is set to  $1 \times 10^{-4}$ , the optimizer is Adam, and the exponential decay rate correspond  $\beta_1 = 0.9$  and  $\beta_2 = 0.999$ . When the number of iterations is reduced to half, the learning rate is adjusted in half, obtaining the best reconstruction performance.

Table 1. DCA-GAN network parameters.

<b>Dataset Parameters</b>	Crop Size	96 × 96
Model parameters	Kernel size in generator	$3 \times 3$
	Kernel size in discriminator	$3 \times 3$
	Activation function in generator	ReLU
	Activation function in discriminator	Leaky ReLU
	Number of middle channels	64
	Number of RDCAB units	16
Learning parameters	Training epochs	80
	Initial learning rate	$10^{-4}$

The convolution kernel dimension of the attention unit in the generator and discriminator adopts  $1 \times 1$ , and the kernel size of the other blocks are  $3 \times 3$ . In the generator, sixteen residual channel attention blocks are cascaded together, which minimizes parameters and the complexity of the model and addresses the problem of poor reconstruction effect due to limited model feature expression.

#### 3.3. Experimental Results

#### 3.3.1. Quantitative Results

With regard to the evaluation, the peak signal-to-noise ratio (PSNR) and structural similarity index (SSIM) are employed in this study. PSNR is a full-reference picture assessment index that is extensively used to quantify the quality of image restoration. It is defined by the mean square error as the following:

$$MSE = \sum_{i=0}^{H} \sum_{j=0}^{W} \frac{\left[ \stackrel{\wedge}{f}(i,j) - f(i,j) \right]^2}{H \times W}$$
(11)

$$PSNR = 10\log_{10}\left(\frac{MAX_i^2}{MSE}\right)$$
(12)

where *MSE* is the average error between the original HR image and the reconstructed SR image, *MAX* is the maximum gray scale, and the unit of *PSNR* is dB. The higher the value of PSNR, the better the reconstructed CT image, the more successful it is.

PSNR takes into account pixel inaccuracy but ignores human vision's visual features. Another measure index, SSIM can characterize the structure information of the image from many aspects such as mean ( $\mu$ ), variance and covariance ( $\sigma$ ), dynamic range of the input image (LR), and stabilizing variable (C). SSIM is expressed as:

$$SSIM(x,y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_x\sigma_y + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}$$
(13)

*SSIM* has values range of [0, 1]. When the SSIM is closer to 1, SR image is closer to the real HR sample.

The comparisons of our RDCA-SRGAN model with SRCNN, A+, FSRCNN, VDSR, DRRN, SRGAN, EDSR, RCAN, and ESRGAN are carried out on sandstone, coal, and carbon images, each of which contains CT images with the solution of  $500 \times 500$ . The average PSNR and SSIM metrics evaluated for each subgroup are shown in Table 2. The results in Table 2 show that SRGAN, ESRGAN, and our RDCA-SRGAN model have a considerable increase in structural similarity, while SRCNN, FSRCNN, VDSR, and DRRN perform rather poorly, indicating that the texture of SR images created by CNN is not pixel-matched. The best results in each scale factor are shown in bold. According to all the above experiments, our model provides the optimal SSIM.

**Table 2.** Comparisons of PSNR and SSIM from SRCNN, A+, FSRCNN, VDSR, DRRN, SRGAN, EDSR, RCAN, ESRGAN, and RDCA-SRGAN.

Algorithm	Scale	Average PSNR	Average SSIM
SRCNN (2014)	2	32.90	0.791
	4	31.27	0.790
	8	28.32	0.730
A+ (2014)	2	32.49	0.827
	4	31.45	0.809
	8	28.74	0.797
FSRCNN (2016)	2	33.25	0.800
	4	31.50	0.792
	8	28.40	0.732
VDSR (2016)	2	33.72	0.828
	4	32.13	0.814
	8	28.71	0.748
DRRN (2017)	2	33.82	0.831
	4	32.33	0.822
	8	28.91	0.750
SRGAN (2017)	2	34.11	0.895
	4	32.42	0.862
	8	29.25	0.804
EDSR (2017)	2	34.47	0.863
	4	32.98	0.849
	8	29.57	0.766
RCAN (2018)	2	34.46	0.888
	4	33.11	0.854
	8	29.68	0.775
ESRGAN (2018)	2	34.46	0.900
	4	32.50	0.879
	8	29.32	0.829
RDCA-SRGAN (Ours)	2	34.30	0.914
	4	32.94	0.906
	8	29.43	0.866

The performance curve for a scale factor equal to 4 is shown in Figure 10. We can see from this figure that the SSIM performance of RDCA-SRGAN is better than the others.



Notably, the RDCA-SRGAN combines residual learning and dual-channel attention to achieve better performance in capturing texture details of rock CT images.

**Figure 10.** Comparison of PSNR generated SRCNN, FSRCNN, VDSR, DRRN, SRGAN, EDSR, RCAN, ESRGAN, and RDCA-SRGAN.

Figures 10 and 11 show the PSNR and SSIM produced by the nine models when the scale factor is equal to 4. PSNR is as predicted in Figure 10, SRGAN and ESRGAN images have lower accuracy on PSNR. The presence of small crack structures that cannot be reconstructed by SRGAN and ESRGAN results in lower PSNR values. The results in Figure 11 show that our RDCA-SRGAN beats all other popular models. Conversely, however, SRGAN and ESRGAN perform much better on structural similarity compared to EDSR and RDCA, which is exactly what was mentioned in the previous section, and GANs generate superior visuals.



Figure 11. Comparison of SSIM generated SRCNN, FSRCNN, VDSR, DRRN, SRGAN, EDSR, RCAN, ESRGAN, and RDCA-SRGAN.

## 3.3.2. Qualitative Results

To the human eye, the difference between the SR and HR images is not visible when the scale is set to 2. If the scale is set to 8, the resulting image will be blurry. Therefore, we chose a scale of 4 to present qualitative results. Figures 12–14 show the comparison between HR and SR images generated by SRCNN, A+, FSRCNN, VDSR, DRRN, SRGAN,



EDSR, RCAN, ESRGAN and our RDCA-SRGAN model. In the original image, the part of the image that will undergo super-resolution reconstruction is marked with a red box.

Figure 12. Comparison between HR images on carbon and SR images generated using nine models.



Figure 13. Comparison between HR images on coal and SR images generated using nine models.

As shown in Figures 12–14, SRCANN and FSRCNN produce distortion and blur, and the reconstruction result is the most distorted. After optimizing the settings of VDSR, DRRN, EDSR and RCAN, the reconstruction effect of rock CT images is enhanced in turn. A+ produced SR images with much fewer artifacts than RCAN. Similarly, the images formed by SRGAN and ESRGAN are sharper, but the edges of the images are less smooth when the detail texture is increased. Based on the efficient loss function assistance of SRGAN, it is worth noting that RDCA-SRGAN achieves less content loss and richer texture features.

 Original
 SRCN
 A+
 FSRCN

 Original
 SRCN
 A+
 FSRCN

 VDSR
 DRN
 EDSR
 RCN

 VDSR
 DRN
 EDSR
 RCN

 SRGAN
 ESRGAN
 BCA-SRGAN
 SCAN

Figure 14. Comparison between HR images on sandstone and SR images generated using nine models.

Details inspection of Figures 12–14 indicates that there is a better extension from CNNbased models to GAN-based. The margins of coal cracks and grain features in sandstone and carbon are difficult to reconstruct using CNN-based algorithms, such as SRCNN, A+, FSRCNN, VDSR, and DRRN. The EDSR and RCAN models fail in extracting intragranular information, but they are advantageous when the intricacies inside are relevant. SRGAN and ESRGAN methods have contributed to pixel-for-pixel match to the original images. The proposed RDCA-SRGAN improves on the edge details in the sample images, resulting in a better match with the strongly defined fractured characteristics of the coal, sandstone, and carbon images, seeming the same as their original counterparts.

Meanwhile, Figures 12–14 show that in the above nine training models, the SRGAN, EDSR, RCAN, and ESRGAN algorithms can better analyze the texture and high-frequency information of reconstructed CT images, resulting in clearer and higher-quality SR images. However, in the image's magnified details, the edges are not crisp enough, and the details are not realistic enough. When compared to the SRGAN, EDSR, RCAN, and ESRGAN algorithms, RDCA-SRGAN can enrich more detailed texture information and improve the visual effect. The edges of the SR pictures created by RDCA-SRGAN are crisper and sharper in the image with enhanced features, which improves the image reconstruction performance.

### 4. Discussion and Future Works

LR and HR images share most of the same features, not to mention that residual learning allows deeper network layers and avoids vanishing and degenerating gradients. It is reasonable to explicitly plug residual learning into the framework. Channel attention has been widely used in previous studies to increase the percentage of valuable information obtained. Dual-channel attention can rescale each feature channel, enabling the network to focus on more valuable channels and improving discriminative learning. With the help of dual-channel attention and residual learning, the generated CT rock SR images provide more high-frequency features. A recent trend is to use GANs to improve picture realism, so based on GANs, the above experiments show that the proposed RDCA-SRGAN outperforms earlier techniques.

Our current research focused on two-dimensional (2D) pictures, however, 3D rock images provide more structural information, allowing geologists to better examine geological sciences. Our further research will focus on more refined 3D images of SR rock reconstructions. Furthermore, the advantages of multi-scale image fusion have led to great interest in resolution improvement. Inspired by multi-scale reconstruction [2,46], we will investigate multi-scale fusion instead of reconstructing SR with single-scale images. The reconstructed image properties will be evaluated on the real rock images. The results will be used to determine whether the reconstructed CT image has additional accurate features and is sufficient to describe the real rock. Specifically, for the input, the CT images will be down-sampled using  $\times 2$ ,  $\times 4$ , and  $\times 8$ , and then the up-sampling and features will be fused at different scales. The features derived from different resolutions will be more plentiful. Images from multi-scale fusion reconstruction will be closer to human perception.

Additionally, in this study, the used low-resolution images are synthesized by hands. Considering practical perspective, for LR and SR images from natural rock, LR and SR data sets are required to be meticulously registered. However, in some exceptional circumstances, aligned sample pairs are tough to accomplish, thus, to drive scholars forward new research directions, such as unpaired networks, image registration, and domain shift [47–52]. CycleGAN [48] and SRCycleGAN [49], which are used to convert images from the source domain to the target domain, attempt mapping between LR and HR on unregistered images without using paired examples. But loss function availability and texture details from the above two models were far from the expectations. Wang et al. [53] used CycleGAN for unpaired images and SRGAN for rock CT image reconstruction. Unfortunately, geologically unpaired data are not well taken into account in their evaluation, failing to complete missions that require geometric adjustments. Using the idea of [47–53] for reference, we will explore the solution for addressing rock CT super-resolution reconstruction under a practical scenario in the next work.

## 5. Conclusions

In this paper, a residual two-channel attention mechanism is introduced into a GAN network to improve the resolution of digital rock CT images limited by hardware devices. Evaluations of the CT images of sandstone, coal, and carbon rock samples show that the proposed RDCA-SRGAN excels in capturing image edge details and recovering high-frequency features. Comparing RDCA-SRGAN with SRCNN, A+, FSRCNN, VDSR, DRRN, SRGAN, EDSR, RCAN, and ESRGAN, SRCNN and FSRCNN still contain a lot of noise, and the reconstructed SR images appear blurry, jagged and distorted; VDSR and DRRN reconstructed images are clearer, but the perception of SR images is still relatively blurry; SRGAN improves the pixel accuracy and reconstructs clear images, but the reconstructed SR images are a bit blurry, the edges of the SR images are not sharp enough, and the details of the SR images are not natural enough. EDSR, RCAN, and ESRGAN significantly outperform their five preceding models, with higher values for PSRN and SSIM. However, our proposed RDCA-SRGAN framework outperforms other algorithms. The texture similarity between the SR image are clearer, and the unresolved crack features are clearly resolved.

**Author Contributions:** Conceptualization, L.S.; Data curation, C.L.; Investigation, Y.L.; Methodology, Y.L.; Project administration, L.S.; Validation, C.L. and X.H.; Visualization, W.K.; Writing—original draft, C.L.; Writing—review & editing, L.S. and X.H. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the Natural Science Foundation of Hebei Province grant number E2021107005. This research was funded by Northeast Petroleum University Foundation grant number 2018GP2D-04 and No. 2018QNQ-06.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

**Data Availability Statement:** The data presented in this study are openly available in DeepRock-SR at 10.17612/s3m9-e024, ref. [45].

**Acknowledgments:** We gratefully acknowledge the helpful comments of the editor and anonymous reviewers.

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

# References

- Coenen, J.; Tchouparova, E.; Jing, X. Measurement parameters and resolution aspects of micro X-ray tomography for advanced core analysis. In Proceedings of the 2004 International Symposium of the Society of Core Analysts, Abu Dhabi, United Arab Emirates, 5–9 October 2004.
- Jackson, S.J.; Niu, Y.; Manoorkar, S.; Mostaghimi, P.; Armstrong, R.T. Deep learning of multi-resolution X-ray micro-CT images for multi-scale modelling. *arXiv* 2021, arXiv:2111.01270.
- 3. Zhan, Q.M.; Zhuang, M.W.; Liu, Q.H. A Compact Upwind Flux with More Physical Insight for Wave Propagation in 3-D Poroelastic Media. *IEEE Trans. Geosci. Remote Sens.* 2018, *56*, 5794–5801. [CrossRef]
- 4. Andriamihaja, S.; Padmanabhan, E.; Ben-Awuah, J. Characterization of pore systems in carbonate using 3D X-ray computed tomography. *Pet. Coal* **2016**, *58*, 507–516.
- 5. Cnudde, V.; Boone, M.N. High-resolution X-ray computed tomography in geosciences: A review of the current technology and applications. *Earth-Sci. Rev.* 2013, 123, 1–17. [CrossRef]
- 6. Wang, Y.; Rahman, S.S.; Arns, C.H. Super resolution reconstruction of μ-CT image of rock sample using neighbour embedding algorithm. *Phys. A Stat. Mech. Its Appl.* **2018**, 493, 177–188. [CrossRef]
- Dong, C.; Loy, C.C.; He, K.; Tang, X. Accelerating the super-resolution convolutional neural network. In Proceedings of the 2016 European Conference on Computer Vision, Amsterdam, The Netherlands, 11–14 October 2016; pp. 391–407.
- 8. Lim, B.; Son, S.; Kim, H.; Nah, S.; Mu Lee, K. Enhanced deep residual networks for single image super-resolution. In Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 136–144.
- 9. Song, H.; Jin, Y.; Cheng, Y. Learning interlaced sparse Sinkhorn matching network for video super-resolution. *Pattern Recognit*. **2022**, 124, 108475. [CrossRef]
- 10. Wang, Z.; Zhou, Y.; Xu, R. Seeing the unseen: AIE luminogens for super-resolution imaging. *Coord. Chem. Rev.* 2022, 451, 214279. [CrossRef]
- 11. Wu, W.; Zheng, C. Single image super-resolution using self-similarity and generalized nonlocal mean. In Proceedings of the 2013 IEEE International Conference of IEEE Region, Xi'an, China, 10 October 2013; pp. 1–4.
- Dosovitskiy, A.; Brox, T. Generating images with perceptual similarity metrics based on deep networks. In Proceedings of the Advances in Neural Information Processing Systems 29: Annual Conference on Neural Information Processing Systems 2016, Barcelona, Spain, 5–10 December 2016; pp. 658–666.
- Wang, T.C.; Liu, M.Y.; Zhu, J.Y.; Tao, A.; Kautz, J.; Catanzaro, B. High-resolution image synthesis and semantic manipulation with conditional gans. In Proceedings of the 2018 IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 8798–8807.
- 14. Timofte, R.; De Smet, V.; Van Gool, L. A+: Adjusted anchored neighborhood regression for fast super-resolution. In Proceedings of the 2014 Asian Conference on Computer Vision, Singapore, 1–5 November 2014; pp. 111–126.
- 15. Zhang, Y.; Li, K.; Li, K.; Wang, L.; Zhong, B.; Fu, Y. Image super-resolution using very deep residual channel attention networks. In Proceedings of the 2018 European Conference on Computer Vision, Munich, Germany, 8–14 September 2018; pp. 286–301.
- 16. Dong, C.; Loy, C.C.; He, K.; Tang, X. Learning a deep convolutional network for image super-resolution. In Proceedings of the 2014 European Conference on Computer Vision, Zurich, Switzerland, 6–12 September 2014; pp. 184–199.
- Guo, Y.; Chen, J.; Wang, J.; Chen, Q.; Cao, J.; Deng, Z.; Xu, Y.; Tan, M. Closed-loop matters: Dual regression networks for single image super-resolution. In Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 13–19 June 2020; pp. 5407–5416.
- Maeda, S. Unpaired image super-resolution using pseudo-supervision. In Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Virtual, 14–19 June 2020; pp. 291–300.
- Rad, M.S.; Bozorgtabar, B.; Marti, U.V.; Basler, M.; Ekenel, H.K.; Thiran, J.P. Srobb: Targeted perceptual loss for single image super-resolution. In Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision, Seoul, Korea, 27 October–2 November 2019; pp. 2710–2719.
- 20. Wang, Y.; Armstrong, R.T.; Mostaghimi, P. Enhancing resolution of digital rock images with super resolution convolutional neural networks. *Pet. Sci. Eng.* **2019**, *182*, 106261. [CrossRef]
- Zhou, R.; Susstrunk, S. Kernel modeling super-resolution on real low-resolution images. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Seoul, Korean, 20–26 October 2019; pp. 2433–2443.
- 22. Kim, J.; Lee, J.K.; Lee, K.M. Deeply-recursive convolutional network for image super-resolution. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 1637–1645.

- Shi, W.; Caballero, J.; Huszár, F.; Totz, J.; Aitken, A.P.; Bishop, R.; Rueckert, D.; Wang, Z. Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 1874–1883.
- 24. 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, Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.
- 25. Kim, J.; Lee, J.K.; Lee, K.M. Accurate image super-resolution using very deep convolutional networks. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 1646–1654.
- Demiray, B.Z.; Sit, M.; Demir, I. D-SRGAN: Dem super-resolution with generative adversarial networks. SN Comput. Sci. 2021, 2, 1–11. [CrossRef]
- 27. Gu, Y.; Zeng, Z.; Chen, H.; Wei, J.; Zhang, Y.; Chen, B.; Li, Y.; Qin, Y.; Xie, Q.; Jiang, Z. Medsrgan: Medical images super-resolution using generative adversarial networks. *Multimed. Tools Appl.* **2020**, *79*, 29–30. [CrossRef]
- Gupta, R.; Sharma, A.; Kumar, A. Super-resolution using gans for medical imaging. *Procedia Comput. Sci.* 2020, 173, 28–35. [CrossRef]
- He, X.; Lei, Y.; Fu, Y.; Mao, H.; Curran, W.J.; Liu, T.; Yang, X. Super-resolution magnetic resonance imaging reconstruction using deep attention networks. In Proceedings of the SPIE—The International Society for Optical Engineering, Houston, TX, USA, 10 March 2020; Volume 90.
- Ledig, C.; Theis, L.; Huszár, F.; Caballero, J.; Cunningham, A.; Acosta, A.; Aitken, A.; Tejani, A.; Totz, J.; Wang, Z.; et al. Photo-realistic single image super-resolution using a generative adversarial network. In Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 4681–4690.
- Li, J.C.; Pei, Z.H.; Zeng, T.Y. From beginner to master: A survey for deep learning-based single-image super-resolution. *arXiv* 2021, arXiv:2109.14335.
- Shahsavari, A.; Ranjbari, S.; Khatibi, T. Proposing a novel cascade ensemble super resolution generative adversarial network (CESR-GAN) method for the reconstruction of super-resolution skin lesion images. *Inform. Med. Unlocked* 2021, 24, 100628. [CrossRef]
- Xu, M.; Wang, Z.; Zhu, J.; Jia, X.; Jia, S. Multi-Attention Generative Adversarial Network for Remote Sensing Image Super-Resolution. arXiv 2021, arXiv:2107.06536.
- 34. Goodfellow, I.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative adversarial nets. In Proceedings of the Neural Information Processing Systems, Montreal, QC, Canada, 8–11 December 2014.
- Wang, X.; Yu, K.; Wu, S. ESRGAN: Enhanced Super-Resolution Generative Adversarial Networks. In Proceedings of the 2018 European Conference on Computer Vision, Munich, Germany, 8–14 September 2018; pp. 63–79.
- You, C.; Li, G.; Zhang, Y.; Zhang, X.; Shan, H.; Li, M.; Ju, S.; Zhao, Z.; Zhang, Z.; Cong, W.; et al. CT super-resolution GAN constrained by the identical, residual, and cycle learning ensemble (GAN-CIRCLE). *IEEE Trans. Med. Imaging* 2018, 39, 188–203. [CrossRef]
- Tai, Y.; Yang, J.; Liu, X.M. Image super-resolution via deep recursive residual network. In Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, USA, 21–26 July 2017; pp. 2790–2798.
- Bultreys, T.; Van Hoorebeke, L.; Cnudde, V. Multi-scale, micro-computed tomography-based pore network models to simulate drainage in heterogeneous rocks. *Adv. Water Resour.* 2015, 78, 36–49. [CrossRef]
- 39. Wang, Y.D.; Armstrong, R.T.; Mostaghimi, P. Boosting resolution and recovering texture of 2D and 3D micro-CT images with deep learning. *Water Resour. Res.* 2020, *56*, e2019WR026052. [CrossRef]
- 40. Shan, L.; Bai, X.; Liu, C.; Feng, Y.; Liu, Y.; Qi, Y. Super-resolution reconstruction of digital rock CT images based on residual attention mechanism. *Adv. Geo-Energy Res.* 2022, *6*, 157–168. [CrossRef]
- Qiu, Y.; Wang, R.; Tao, D.; Cheng, J. Embedded block residual network: A recursive restoration model for single-image superresolution. In Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision, Seoul, Kerean, 27 October–2 November 2019; pp. 4180–4189.
- Niu, B.; Wen, W.; Ren, W.; Zhang, X.; Yang, L.; Wang, S.; Zhang, K.; Cao, X.; Shen, H. Single image super-resolution via a holistic attention network. In Proceedings of the 2020 European Conference on Computer Vision, Glasgow, UK, 23–28 August 2020.
- 43. Hu, J.; Shen, L.; Sun, G. Squeeze-and-excitation networks. In Proceedings of the 2018 IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–22 June 2018; pp. 7132–7141.
- 44. Fu, J.; Liu, J.; Tian, H. Dual attention network for scene segmentation. In Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Los Angeles, CA, USA, 15–21 June 2019; pp. 3146–3154.
- 45. Wang, Y.D.; Armstrong, R.T.; Mostaghimi, P.A. Diverse Super Resolution Dataset of Digital Rocks (DeepRock-SR): Sandstone, Carbonate, and Coal; National Science Foundation: Alexandria, VA, USA, 2019.
- Li, J.; Fang, F.; Mei, K.; Zhang, G. Multi-scale residual network for image super-resolution. In Proceedings of the 2018 European Conference on Computer Vision, Munich, Germany, 8–14 September 2018; pp. 517–532.
- Zhang, K.; Gu, S.; Timofte, R. Ntire 2020 challenge on perceptual extreme super-resolution: Methods and results. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, Seattle, WA, USA, 14–19 June 2020; pp. 492–493.
- Zhu, J.Y.; Park, T.; Isola, P.; Efros, A.A. Unpaired image-to-image translation using cycle-consistent adversarial networks. In Proceedings of the 2017 IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 2223–2232.

- 49. Chen, H.; He, X.; Teng, Q.; Sheriff, R.E.; Feng, J.; Xiong, S. Super-resolution of real-world rock microcomputed tomography images using cycle-consistent generative adversarial networks. *Phys. Rev. E* 2020, 101, 023305. [CrossRef]
- 50. Wang, Y.D.; Shabaninejad, M.; Armstrong, R.T.; Mostaghimi, P. Physical accuracy of deep neural networks for 2d and 3d multi-mineral segmentation of rock micro-CT images. *arXiv* 2020, arXiv:2002.05322. [CrossRef]
- Wang, Y.D.; Chung, T.; Armstrong, R.T.; McClure, J.; Ramstad, T.; Mostaghimi, P. Accelerated computation of relative permeability by coupled morphological and direct multiphase flow simulation. *J. Comput. Phys.* 2020, 401, 108966. [CrossRef]
- 52. Wang, Y.D.; Chung, T.; Armstrong, R.T.; McClure, J.E.; Mostaghimi, P. Computations of permeability of large rock images by dual grid domain decomposition. *Adv. Water Resour.* **2019**, *126*, 1–14. [CrossRef]
- 53. Wang, Y.D.; Blunt, M.J.; Armstrong, R.T.; Mostaghimi, P. Deep learning in pore scale imaging and modeling. *Earth-Sci. Rev.* 2021, 215, 103555. [CrossRef]