Super-resolution reconstruction of hydrate-bearing sediment computed tomography images for microscopic detection of pore structure

liang chen[1], Wangquan Ye[2], Chengfeng Li[3], Bin Gui[4], Guohua guo Hou[3], Qiang Chen[5], Gao Wei Hu[3], Jianye jian Sun[3], and Ronger Zheng[2]

[1]Ocean Univisity of China, OUC
[2]Optics and Optoelectronics Laboratory, Ocean University of China
[3]Qingdao Institute of Marine Geology
[4]College of Physics and Optoelectronic Engineering, Ocean University of China
[5]Qingdao Institute of Marine Geology, China Geological Survey

January 3, 2023

Abstract

The pore structure of marine sediments varies with the distribution of gas-hydrate, hence affecting the gas-water permeability. CT image is a conventional approach to view the internal structure, while for hydrate-bearing sediment investigation, rather poor resolution of obtained image has limited the accuracy of the analysis. Recently, super-resolution (SR) reconstruction techniques have been used to enhance the spatial resolution of CT images with varying degrees of improvement. Typical Image Pairs-Based SR (PSR) methods require higher resolution matching images for training, which is challenging for hydrate samples in dynamic temperature and pressure conditions. Here, we introduced a self-supervised learning (SLSR) method that only relies on a single input image to complete the process of training and reconstruction. We conducted a complete training to establish an end-to-end network consisting of two sub-networks, an SR network and a downscaling network. Self-built datasets from three hydrate samples with different sediment grains were trained and tested. Compared with the typical method, the SR results show that our method provides higher resolution while improving clarity. Moreover, in the subsequent calculation of porosity parameters, it has the highest consistency with the liquid saturation method. This study contributes to investigating the water seepage and energy transfer in the gas hydrate bearing sediments, which is particularly important for the exploration and development of marine natural gas hydrate resources. The image super-resolution method established by us has also a broad application prospect in the field of CT imaging.

Hosted file

951539_0_art_file_10514663_rmyy72.docx available at https://authorea.com/users/565008/articles/612024-super-resolution-reconstruction-of-hydrate-bearing-sediment-computed-tomography-images-for-microscopic-detection-of-pore-structure

Super-resolution reconstruction of hydrate-bearing sediment computed tomography images for microscopic detection of pore structure

Liang Chen[1], Wangquan Ye[1*], Chengfeng Li[1,3*], Bin Gui[1], Guohua Hou[2], Qiang Chen[3], Gaowei Hu[3], Jianye Sun[3], Ronger Zheng[1]

[1] College of Physics and Optoelectronic Engineering, Ocean University of China, Qingdao 266100, China.

[2] Zhoushan Field Scientific Observation and Research Station for Marine Geo-hazards, China Geological Survey, Qingdao 266237, China.

[3] Key Laboratory of Gas Hydrate, Ministry of Natural Resources, Qingdao Institute of Marine Geology, Qingdao 266237, China.

Correspondence to:

W. Ye and C. Li

yewangquan@ouc.edu.cn;

lchengfeng@mail.cgs.gov.cn

Key Points:

- Further breakthrough the CT imaging limits, presenting more clear internal structure of sediment pores.
- Enables super-resolution image reconstruction technique for fine structure detection in sediment pores.
- The pore structure of sediments is more clearly resolved and the porosity parameters are calculated more accurately.

Abstract

The pore structure of marine sediments varies with the distribution of gas-hydrate, hence affecting the gas-water permeability. CT image is a conventional approach to view the internal structure, while for hydrate-bearing sediment investigation, rather poor resolution of obtained image has limited the accuracy of the analysis. Recently, super-resolution (SR) reconstruction techniques have been used to enhance the spatial resolution of CT images with varying degrees of improvement. Typical Image Pairs-Based SR (PSR) methods require higher resolution matching images for training, which is challenging for hydrate samples in dynamic temperature and pressure conditions. Here, we introduced a self-supervised learning (SLSR) method that only relies on a single input image to complete the process of training and reconstruction. We conducted a complete training to establish an end-to-end network consisting of two sub-networks, an SR network and a downscaling network. Self-built datasets from three hydrate samples with different sediment grains were trained and tested. Compared with the typical method, the SR results show that our method provides higher resolution while improving clarity. Moreover, in the subsequent calculation of porosity parameters, it has the highest consistency with the liquid saturation method. This study contributes to investigating the water seepage and energy transfer in the gas hydrate bearing sediments, which is particularly important for the exploration and development of marine natural gas hydrate resources. The image super-resolution method established by us has also a broad application prospect in the field of CT imaging.

Plain Language Summary

When trying to break through the hardware limitation of X-ray images by some machine learning methods, it always requires higher resolution images for a training process. That is impossible to operate with gas hydrate samples, for it is hard to keep them stable without a low-temperature and high-pressure environment. We introduced an optimized process that uses only the original images without paired higher-resolution images. We trained and tested this algorithm on actual X-ray images taken from homemade and field hydrate samples. The processed images were presented with higher resolution and higher image quality, which can give more accurate microstructure information hidden in the images. The results show that this method has broad application prospects in marine sediment microscopic detection.

1 Introduction

The morphology of gas hydrate in Marine sediment pores has a significant influence on the physical characteristics such as acoustic velocity, resistivity, and permeability, which largely determines the accuracy of hydrate geophysical exploration and resource evaluation (PRIEST *et al*., 2005; REN *et al*., 2010; ZHANG *et al*., 2020). So far, the natural gas hydrate in the pores of marine sediments is still distributed in a dispersed manner, which is invisible to the naked eye. For example, gas hydrate is mainly filled in the pores of muddy or sandy sediments of Shenhu area, South China Sea (LIU *et al*., 2017). X-ray micro-computed tomography (micro-CT) has been widely leveraged to explore microscopic hydrate-bearing sediments since it can visually present the microstructure characteristics and phase changes of different components without destroying samples. The researchers leverage advances in micro-CT image acquisition and analysis techniques to create 3D digital images of gas hydrate samples, which are used for computational modeling and simulations to calculate physical property parameters of interest, such as saturation, porosity, and permeability (Wang *et al*., 2018). However, the accuracy of calculated parameters is crucially dependent on the quality of digital images, which is currently limited by the resolution of the micro-CT scanning technology. High-resolution data, however, results in a small field of view (FOV), and thus a trade-off between image FOV and image resolution is made (Wildenschild & Sheppard, 2013), which leads to non-representative results (Li *et al*., 2017). Recent developments in Image Super-Resolution (SR) methods allow images of low resolution (LR) to have fine details compared to high-resolution images (HR) using deep learning, which may be an effective means to circumvent the trade-off between high resolution and FOV to assist in more accurate physical analysis.

Image Super-Resolution techniques, an ill-posed and indeterminate inverse problem with an infinite solution space, reconstruct higher resolution output from the LR observation to obtain images with a resolution beyond the limit of hardware. In recent years, SR methods based on deep learning especially convolutional neural networks (CNN), have increasingly become a robust way to improve the performance of Single Image Super-Resolution (SISR) (Dong *et al*., 2016; Kim *et al*., 2016; Tai *et al*., 2017; Ledig *et al*., 2017; Wang *et al*., 2019), which have been used in digital rock micro-CT images. Wang et al. (2019) compared SR-Resnet, Enhanced Deep SR (EDSR), and Wide-Activation Deep SR (WDSR) methods on the performance of super-resolving micro-CT images of sandstone and carbonate rocks, which were trained on paired synthesized LR-HR images where the LR images were bicubically downscaled from original HR images. Hou et al. (2021) proposed a generative adversary network of an image segmentation network as a discriminator constrained by perspective information and prior information (SCPGAN) to enhance micro-CT digital rock images resolution which shows GAN based model with prior information has excellent anti-noise capacity. Janssens et al. (2020) used a generative adversarial network (GAN) to improve the CT image resolution of the reservoir and some physical parameters of the reservoir such as pore network properties and single-phase, unsaturated, and two-phase flow were compared after super-resolution. The results showed relevant small pores and pore surfaces are better resolved thus providing better estimates of unsaturated and two-phase flow.

Note that, the SR methods mentioned above focus more on the super-resolution of synthetic images whose LR images are down-scaled from corresponding HR images, which may cause some problems in practical applications. Firstly, the mapping between downscaled images and original HR images may deviate from the realistic model, which makes state-of-the-art SR methods trained on LR-HR image pairs produced with the assumption suffer from significant performance degradation. Secondly, there is a need to super-resolve the highest resolution images with the best FOV the instrument can achieve, whether it is feasible to apply the model trained based on this hypothesis to super-resolution HR images is still lack of sufficient evidence.

To overcome these challenges, Real-World Image Pairs-Based methods, Domain Translation-Based methods, and Self-Supervised Learning-Based methods have been introduced. Real-World Image Pairs-Based methods directly collect the images of the same scenario with different resolutions to model the direct mapping of realistic LR-HR image pairs (Chen *et al*., 2019; Zhang *et al*., 2019; Cai *et al*., 2019). However, it is difficult to get completely matched LR-HR image pairs in the real world while misalignment may cause blur artifacts. And in the field of gas-bearing hydrate, whose formation and decomposition process may cause a great challenge to collect the realistic LR-HR image pairs. As it is hard to obtain datasets with well-aligned LR-HR image pairs, the Domain Translation-Based methods translate texture from a high-resolution domain to a low-resolution domain without one-to-one correspondence between LR and HR images (Yuan *et al*., 2018; Kim *et al*., 2020; You *et al*., 2020). The Domain Translation-Based methods use Cycle-Consistent Adversarial Networks (Cycle-GAN), based on

Unpaired image-to-image Translation. (Zhu *et al*., 2017). Niu et al (2021). used paired and unpaired micro-CT images of a carbonate rock sample with complicated micro-porous textures to train a convolutional neural network (CNN) and Cycle-GAN respectively, whose quantitative results show that the unpaired GAN approach can reconstruct super-resolution images as precise as paired CNN method. Chen et al. (2020) proposed a cycle-consistent generative adversarial network (Cycle-GAN)-based SR approach for real-world rock micro-CT images super-resolution, which is trained on a set of unpaired rock images at different resolutions. The experimental results showed great consistency with the targets in terms of both the visual quality and the statistical parameters such as the porosity, the lineal-path function, and the pore size distribution. Niu et al. applied a cycle-in-cycle generative adversarial network (CinCGAN) using unpaired training images to improve the resolution of 3-D micro-CT data, which results demonstrated that CinCGAN provides physically accurate images with an order of magnitude larger field of view when compared to other typical methods (Yuan *et al*., 2018; Niu *et al*., 2020).

As most existing SISR methods use external datasets such as paired or unpaired training data to train SR models, Self-Supervise Learning-Based methods were proposed to exploit the internal information of the single specific LR input. KernelGAN (Bell-Kligler *et al*., 2019) estimates a downscaling kernel for the blind SR based on internal learning, which can be plugged into the reconstruction module ZSSR (Shocher *et al*., 2018) to enhance performance. DBPI (Kim *et al*., 2020) and DualSR (Emad *et al*., 2021) assumed that the SR network not only depends on the estimated kernel but also can improve downscaling kernel estimation, which trained the downscaling kernel along with kernel estimation network and SR network. As Self-Supervise Learning-Based methods train on a single input image, which only utilizes the internal information of LR input while a great deal of external information is neglected because online training with time-consuming. The application of self-supervised learning for super-resolving the micro-CT images in the field of the digital core has not been reported yet.

In this work, the DRRN, SRDenseNet, DualSR, DBPI and the improved DBPI methods were used to enhance the resolution of gas hydrate micro-CT images which were collected in the laboratory. The DRNN and SRDenseNet methods were trained on paired synthesized LR-HR images where the LR images were bicubically downscaled from original HR images which is the same as those approaches used in this field before. The DualSR and DBPI methods are based on self-supervised learning which reconstruct SR images from a single input image. And the improved DBPI method combines the advances of self-supervised learning and Image Pairs-Based method. The SR results show that compared with Image Pairs-Based methods, self-supervised learning methods can obtain CT images with higher resolution and contrast, however, it is greatly affected by image noise, leading to a large change degree of gray value near the center of sand component in CT image before and after reconstruction. In order to overcome the above defects, we improved the DBPI method that performs best previously, which further improved the sharpness and contrast of the image while alleviating the above symptoms. After that, the watershed image segmentation algorithm was used to segment the pores and skeletons in the original CT image, the bicubic interpolation image, the results obtained by DBPI and the results obtained by the improved DBPI. Finally, the porosity parameter was calculated based on the segmentation results, and the mean porosity of the result of the improved DBPI closest to the that measured by the saturated liquid weighing method. The results show that the improved DBPI method can help to distinguish the porosity and skeleton better, so as to calculate more accurate porosity parameters.

2 Materials and Methods

The basic model of SISR follows equation 1, which assumes that the low-resolution input image I_{LR} is the result of downscaling a high-resolution image I_{HR} by a scaling factor s using some kernel k_s and with an additional blur factor n :

$$I_{LR} = (I_{HR} * k_s) \downarrow_s + n, \quad (1)$$

where * and $\downarrow_s$ represent the convolution and the subsampling with a scale factor of s respectively. Image Pairs-Based SR (PSR) methods and Self-Supervised Learning-Based SR (SLSR) methods were used to enhance the resolution of micro-CT images for a comparison.

2.1 SISR by PSR

PSR methods assume that k_s is a Gaussian Kernel which usually is a bicubic downscaling kernel with antialiasing, and learn the direct mapping of paired LR-HR images to reconstruct high-resolution images from low-resolution images which were bicubically downscaled from HR images.

2.1.1 DRRN-PSR

The DRRN-PSR structure (Figure 1) used in this study is inspired by Tai et al. (2017). DRRN-PSR has two substructures: the residual unit and the recursive block. The residual unit contains two convolutional layers of kernel size 3 with rectified linear units (ReLU) activation function and each convolutional layer applies 64 filters. The recursive block contains a convolutional layer in the beginning, and then several residual units are stacked at the behind of the first layer. The recursive block number B and the residual unit number U in each recursive block are the two key parameters in DRRN-PSR. The network structure of DRRN-PSR with B=2 and U=9 was designed to super-resolve the micro-CT images in our implementation, the whole structure of our implementation of DRRN-PSR is shown in Figure 1.

Given a training set $\{\mathrm{X}^{(i)}, \widetilde{\mathrm{X}}^{(i)}\}_{i=1}^{N}$, where N is the number of training patches and $\widetilde{\mathrm{X}}^{(i)}$ is the ground truth HR patch of the LR patch $\mathrm{X}^{(i)}$, the loss function with parameter set Θ of DRRN-PSR is

$$\mathcal{L}(\Theta) = \frac{1}{2N}(\widetilde{\mathrm{X}}^{(i)} - \mathcal{D}(\mathrm{X}^{(i)}))^2. \quad (1)$$

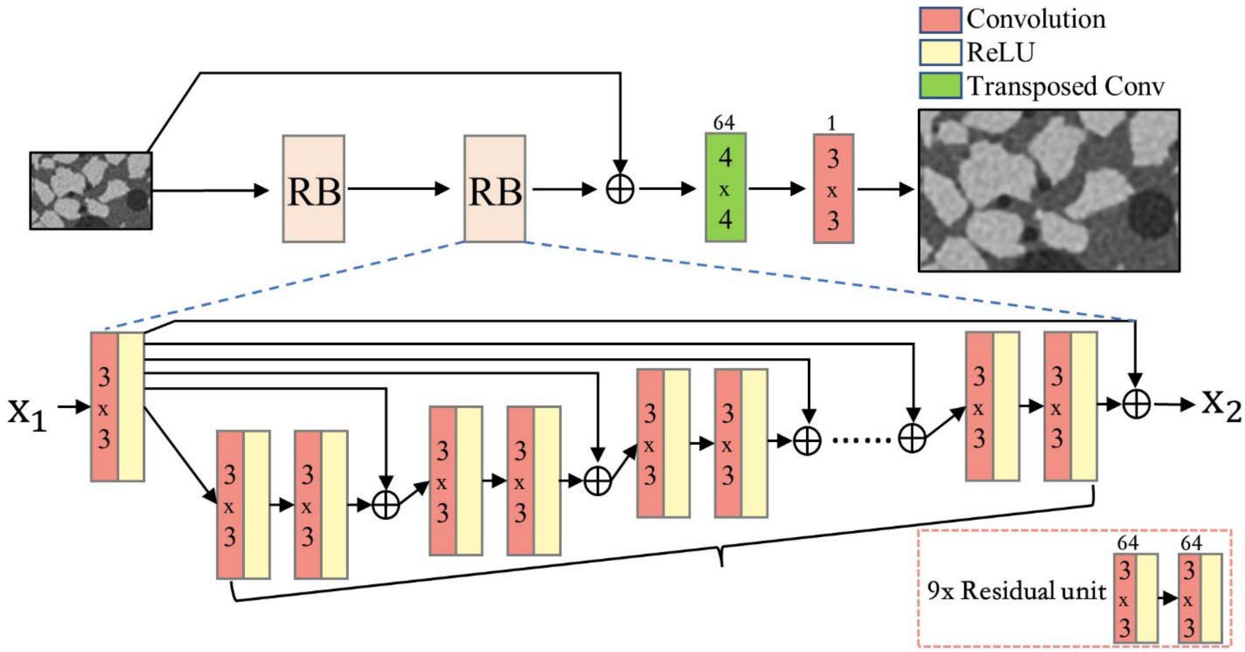

Figure 1. The whole structure of our implementation of DRRN-PSR with B=2 and U=9, where B is the recursive block number and U is the residual unit number in each recursive block. "$k \times k$" in convolution layers represents filter size. The number of filters for each convolution layer is presented above each layer.

2.1.2 SRDenseNet-PSR

SRDenseNet-PSR connects each layer to every other layer in a feed-forward fashion, was proposed by Huang et al. (2018) to strengthen feature propagation, encourage feature reuse, etc. The main body of SRDenseNet-PSR contains dense blocks consisting of several dense layers. Each dense layer are convolutional layers with activation functions such as ReLU. The i^{th} layer in each dense block receives the feature maps of all preceding layers as input:

$$X_i = \mathcal{F}_i\big([X_0, X_1, \dots, X_{i-1}]\big), \tag{2}$$

where $[X_0, X_1, \dots, X_{i-1}]$ refers to the concatenation of the feature maps produced in layers $0, \dots, i-1$ of each dense block, $\mathcal{F}$ is a composite function of convolution and ReLU operations. There has a skip connection after each dense block in our implementation, as shown in Figure 2, and the loss function used is the same as equation 1 in DRRN-PSR.

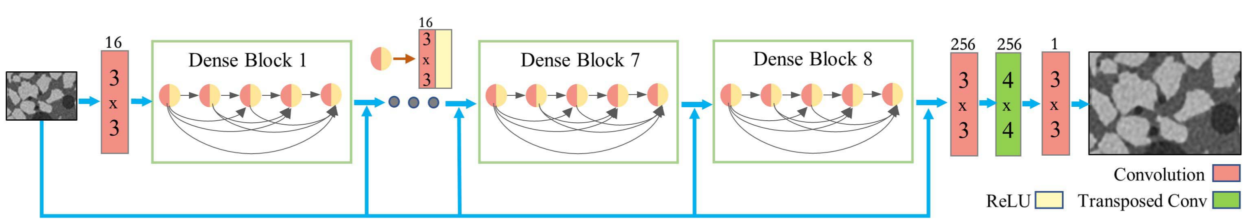

Figure 2. The whole structure of our implementation of SRDenseNet-PSR with eight dense blocks and each dense block has five dense layers (two color colors in the dense block).

2.2 SISR by SLSR

SLSR methods train an image-specific network that learns low-to-high resolution mapping using only patches of the input test image compared to PSR methods which need plenty of paired images for training. A two-stage optimization problem is modeled in this approach, which conducts downscaling kernel estimation followed by SR network training with the estimated kernel.

2.2.1 DualSR-SLSR

The whole network architecture of DualSR-SLSR is shown in Figure 3(a), where the G_{UP} is SR network that trains to super-resolve the input LR image, the G_{DN} is the downscaling network that estimates the downscaling kernel, and the D_{DN} is the discriminator that learns to distinguish between real (patches of the input LR image) and fake (output patches generated by G_{DN}). The downscaling network, SR network, and discriminator are shown in Figure 3 (b), Figure 3 (c) and Figure 3 (d), respectively.

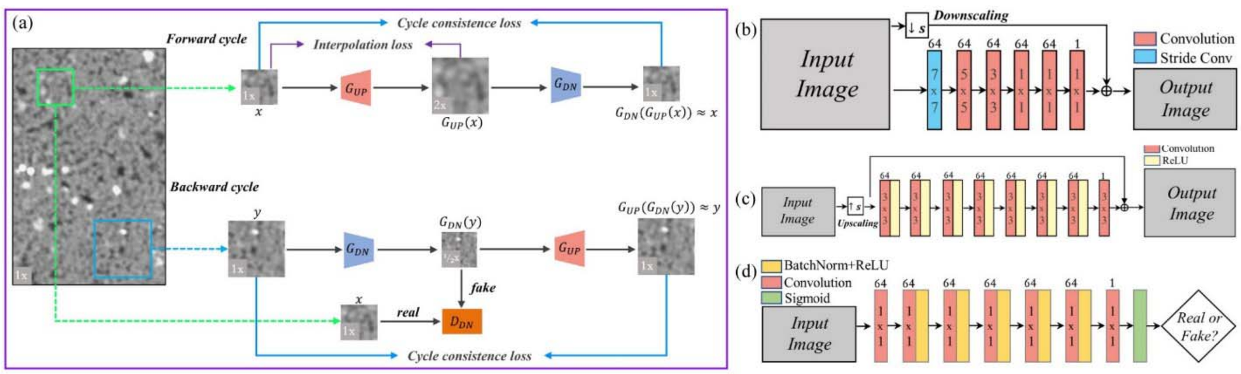

Figure 3. The whole network architecture of DualSR-SLSR (a) contains of the downscaling network (b), the upscaling network (c), and the discriminator (d).

Figure 3(a) demonstrates the forward and backward cycles process with cycle-consistence-loss which are similar to CycleGAN. In the forward cycle, the G_{UP} generates a 2x upscaled image patch, and then G_{DN} is applied and converts the upscaled image patch back to 1x. Similarly, a $^1/_2$x downscaled version of the image patch is generated by G_{DN} and then the G_{UP} upscales the image patch back to the original scale in the backward cycle. Denoting the input image patch as x in the forward cycle and as y in the backward cycle respectively, the cycle-consistence-loss is

$$\mathcal{L}_{cycle} = \mathbb{E}_x \left\| G_{DN}\big(G_{UP}(x)\big) - x \right\|_1 + \mathbb{E}_y \left\| G_{UP}\big(G_{DN}(y)\big) - y \right\|_1. \quad (3)$$

In order to estimate the degradation model accurately, a GAN is used to preserve the distribution of patches across scales of the input image that the output $G_{DN}(y)$ is indistinguishable by the discriminator D_{DN} from input image patches. The adversarial loss for the generator is

$$\mathcal{L}_{GAN} = \mathbb{E}_y [D_{DN}\big(G_{DN}(y)\big) - 1]^2. \quad (4)$$

Then, the final loss function for training the G_{DN}, G_{UP}, and D_{DN} is

$$\mathcal{L}_{total} = \mathcal{L}_{GAN} + \mathcal{L}_{cycle}. \quad (5)$$

2.2.2 DBPI-SLSR

Similar to DualSR-SLSR, the DBPI-SLSR network has a downscaling network G_{DN}, an SR network G_{UP}, the Up-Down process (forward cycle), and the Down-Up process (backward cycle) are shown in Figure 4. In the down-up side, a patch of the input image is first downscaled by G_{DN}, and then the G_{UP} upscales the downscaled one to the original size. Then a $\mathcal{L}_1$ loss is applied to reduce the difference between the input patch x and the output of the Down-Up process, as shown in equation 6:

$$\mathcal{L}_U = \frac{1}{\mathrm{m}\times\mathrm{n}} \left\| G_{UP}\left(G_{DN}(x)\right) - x \right\|_1, \tag{6}$$

where m and n are the height and width of input patch x, respectively. In parallel, a patch of input image is first upscaled by G_{UP} and then G_{DN} downscales the upscaled one to generate an up-down image patch. Then, in the same manner, a $\mathcal{L}_1$ loss is applied to reduce the difference between the input patch x and the output of the Up-Down process, as shown in equation 7:

$$\mathcal{L}_D = \frac{1}{\mathrm{m}\times\mathrm{n}} \| G_{DN}(G_{UP}(x)) - x \|_1. \tag{7}$$

In the training process, $\mathcal{L}_U$ is used to train G_{UP} and $\mathcal{L}_D$ is used to train G_{DN}. The $\mathcal{L}_U$ and $\mathcal{L}_D$ loss consist of the dual back-projection loss, which is

$$\mathcal{L}_{DBP} = \mathcal{L}_U + \mathcal{L}_D = \frac{1}{\mathrm{m}\times\mathrm{n}} \left(\left\| G_{UP}\left(G_{DN}(x)\right) - x \right\|_1 + \left\| G_{DN}\left(G_{UP}(x)\right) - x \right\|_1 \right). \tag{8}$$

The blur kernel of the input image is estimated by the dual back-projection loss implicitly that is different from DualSR-SLSR which uses a GAN to estimate the blur kernel. The network of G_{DN} and G_{UP} of DBPI-SLSR are the same as the network shown in Figure3(b) and Figure 3(c) used in DualSR-SLSR.

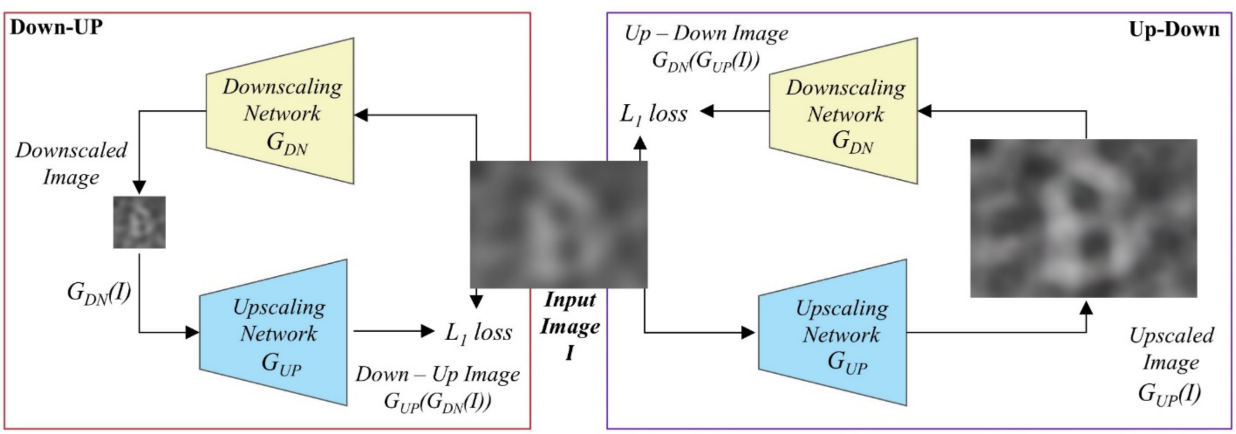

Figure 4. The overall framework of DBPI-SLSR. The downscaling network G_{DN} and the SR network G_{UP} of DBPI are the same as the network shown in Figure3(b) and Figure 3(c).

2.2.3 Our improved SLSR based on DBPI (I-DBPI)

The whole framework of I-DBPI is the same as DBPI-SLSR, however, they differ in the training strategy. More intuitively, I-DBPI learns low-to-high resolution mapping using images of the whole dataset, since the feature information of CT slice images of the same sample is highly coincident. It can make up for the lack of thin feature information when Self-Supervised Learning methods use single image for training, and improve the robustness.

2.3 Performance Evaluation Criteria

2.3.1 Image Similarity Metrics

Peak Singal-to-Noise Ration (PSNR) and Structure Similarity Index (SSIM) are two objective evaluation metrics to measure the difference between the ground truth image and the super-resolved image, which need corresponding reference images. For a better description, let $\mathrm{X} \in \mathbb{R}^{H\times W\times C}$ and $\widehat{\mathrm{X}} \in \mathbb{R}^{H\times W\times C}$ denote the ground truth image and the super-resolved image where H, W, and C are width, height, and channel numbers of the image respectively.

(1) PSNR. PSNR is the most widely used full-reference objective quality assessment metric for image super-resolution, which is more concerned with the proximity between X and $\widehat{\mathrm{X}}$. Given X and $\widehat{\mathrm{X}}$, the PSNR can be calculated by equation 9:

$$PSNR = 10 \times log_{10}\left(\frac{L^2-1}{MSE}\right), \tag{9}$$

where L denotes the maximum pixel value (*i.e.*, 255 for 8-bit images) and MSE is the mean square error between X and $\widehat{\mathrm{X}}$.

(2) SSIM. SSIM is a full-reference objective quality assessment metric that measures structural similarity between X and $\widehat{\mathrm{X}}$. More specifically, SSIM compares the luminance, contrast, and structure between X and $\widehat{\mathrm{X}}$. SSIM is defined as

$$SSIM = [l(\mathrm{X}, \widehat{\mathrm{X}})]^{\alpha}[c(\mathrm{X}, \widehat{\mathrm{X}})]^{\beta}[s(\mathrm{X}, \widehat{\mathrm{X}})]^{\gamma}. \tag{10}$$

Further, equation 11 can be simplified when $\alpha = \beta = \gamma = 1$ and $C_3 = \frac{C_2}{2}$ as

$$SSIM = \frac{(2\mu_{\mathrm{X}}\mu_{\widehat{\mathrm{X}}}+C_1)(2\sigma_{\mathrm{X}}\sigma_{\widehat{\mathrm{X}}}+C_2)}{({\mu_{\mathrm{X}}}^2+{\mu_{\widehat{\mathrm{X}}}}^2+C_1)({\sigma_{\mathrm{X}}}^2+{\sigma_{\widehat{\mathrm{X}}}}^2+C_2)}. \tag{11}$$

2.3.2 Image Clarity Metrics

In this paper, we need to evaluate the image quality when super-resolving the original micro-CT images which have no reference images to calculate the PNSR and SSIM indicators. Since image clarity is an important indicator in the quality evaluation of non-reference images which corresponds better with the subjective feelings of people's eyes, the SMD (Sum of Mean Modulus Difference) is used to quantitatively evaluate the images super-resolved on the original images. The SMD is the sum of the absolute value of the gray difference of the adjacent pixels, which is defined as:

$$SMD = \sum_i \sum_j (|I(i,j) - I(i,j-1)| + |I(i,j) - fI(i+1,j)|), \tag{12}$$

where i and j is the width and height index of the input image, respectively, and $I(i,j)$ denotes the pixel value.

2.4 Datasets and Training Details

2.4.1 Datasets

Three self-built datasets from three hydrate samples were shuffled and split 8:1:1 into training, validation, and testing sets in this study. The three datasets are S1, S2, and S3, as the porosity decreases progressively. They were described in detail as follows:

S1. Quartz sand_Hyd. This dataset contains 800 HR micro-CT images of gas hydrate-bearing quartz sand. The sample is prepared manually in the laboratory and contains four components, methane gas, hydrate, water and quartz sand. The grain size of quartz sand ranges from 500-700 μm. Each image has a spatial resolution of 18 μm and a size of 510×510.

S2. Berea Sandstone_Hyd. This dataset contains 520 HR micro-CT images of gas hydrate-bearing Berea sandstone. The sample is also prepared manually in the laboratory and contains four components, methane gas, hydrate, water and Berea sandstone. The grain size range of Berea sandstone is 150-240um. Each image has a spatial resolution of 16.5 μm and a size of 450×450.

S3. South China Sea sediment_Hyd. This dataset contains 250 HR micro-CT images of gas hydrate-bearing sediment from the South China Sea. There are four components in the image, gas, hydrate, water and sediment. The South China Sea sediment contains foraminiferal shells with coarse particles and clay with fine particles, so its particle size range is large, ranging from 0.02-2000 μm. In addition to argillaceous matrix, there are a small amount of sand and foraminifera shells in the sediments. Each image has a spatial resolution of 18 μm and a size of 450×450.

Figure 5 presents some images of the three datasets mentioned above. From a perspective of pore morphology and image characteristics, the samples differ significantly.

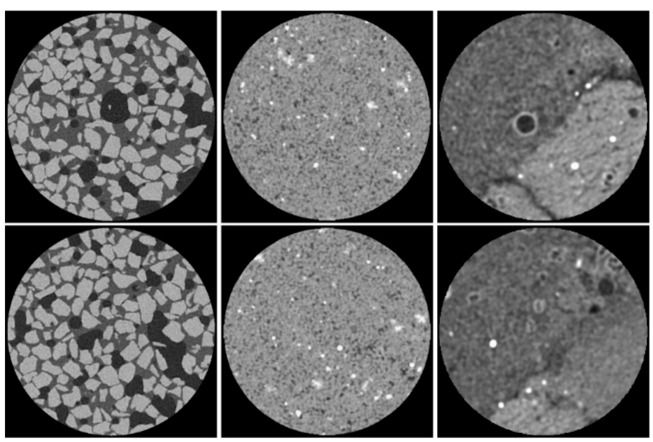

Figure 5. Visualization of images in the proposed datasets *S1 : Quartz sand_Hyd (left), S2 : South China Sea sediment_Hyd (middle),* and *S3 : Berea Sandstone_Hyd (right).*

2.4.2 Training Details

While the DRRN-PSR and SRDenseNet-PSR need paired LR-HR images for training, however, the proposed datasets only contain HR images. Thus, the LR datasets were prepared using images bicubically downscaled from the corresponding HR images. For data augmentation, we crop images to 64×64 size and 32×32 size for the HR and LR training datasets, respectively. The DRRN-PSR and SRDenseNet-PSR were trained using ADAM optimizer with initial learning rate of 0.0001 for 150 iterations, and the batch size is set to 128.

By contrast, the DualSR-SLSR and DBPI-SLSR were trained for every single input image. For the former, the generators and the discriminator were trained successively, a batch of 64×64 and 128×128 batches (patches x and y in Figure 3) were sampled from the input image for each iteration. The networks G_{DN}, G_{UP}, D_{DN} were trained for 3000 iterations with ADAM optimizer. The initial learning rate is 0.001 for G_{DN} and 0.0002 for G_{UP} and D_{DN}, and the learning rate was divided by 10 for every 750 iterations. For the latter, a patch of size 64×64 was sampled from the given image to train the downscaling and upscaling networks successively for each iteration. The networks G_{DN} and G_{UP} were trained with ADAM optimizer for 3000 iterations. The initial learning rate is 0.0001 for both G_{DN} and D_{UP}, and the learning rate was divided by 10 for every 750 iterations. The final super-resolved image is obtained by running the trained upscaling network on the input image.

For I-DBPI, a patch of size 64×64 was sampled from all images of each dataset to train the downscaling and upscaling networks successively for each iteration. The networks G_{DN} and G_{UP} were trained with ADAM optimizer for around 60000 iterations that larger than 3000 iterations in DBPI. The initial learning rate is 0.0001 for both G_{DN} and D_{UP}, and the learning rate was divided by 10 for every 750 iterations.

We use PyTorch for training and testing on an NVIDIA TITAN RTX GPU for these SR methods mentioned above.

3 Results

3.1 Experiments on Synthesized LR Micro-CT Images for 2× SR

The present study was designed to determine the SR methods' effectiveness on micro-CT images through conducting super-resolution experiment on synthesized LR micro-CT images that were bicubically downscaled from the proposed three datasets for 2× super-resolution. Both quantitative and qualitative evaluation between the SR micro-CT images and the corresponding ground truth images can be performed in this experiment, as each synthesized LR micro-CT image has its HR counterpart.

Tabel 1. Quantitative results (PSNR / SSIM / SMD) for 2× SR of different methods on synthesized LR micro-CT images of the proposed three datasets.

Method	***S1***	***S2***	***S3***
Bicubic	33.10 / 0.9739 / 13.07	32.92 / 0.9789 / 11.40	33.66 / 0.9683 / 6.47
DRRN-PSR	35.94 / 0.9877 / 13.90	35.37 / 0.9849 / 10.94	34.92 / 0.9838 / 6.65
SRDenseNet-PSR	36.77 / 0.9913 / 13.94	36.77 / 0.9897 / 11.83	35.31 / 0.9862 / 6.50
DualSR-SLSR	22.12 / 0.8210 / 27.07	26.80 / 0.8453 / 22.55	30.67 / 0.9186 / 10.16
DBPI-SLSR	24.99 / 0.8000 / *27.62	23.64 / 0.7493 / *29.87	25.75 / 0.7687 / *14.76

* represents the best performance.

Table 1 summarizes the quantitative results (PSNR / SSIM / SMD) for 2× SR on the synthesized LR micro-CT images of the proposed three datasets *S1, S2, and S3*. According to Table 1, the PSR methods achieve much better PSNR / SSIM than the SLSR methods, and SRDenseNet-PSR performs the best. By contrast, the SLSR methods achieve much better SMD, among which DBPI-SLSR achieves the best. A possible explanation is that the PSNR and SSIM are used to quantify the similarity between the SR images and the ground truth images, and the PSR methods were trained with multiple paired LR-HR images, thus results in better PSNR and SSIM. In comparison, the SLSR methods consider the information from the image itself, thus results in much sharper images with better SMD.

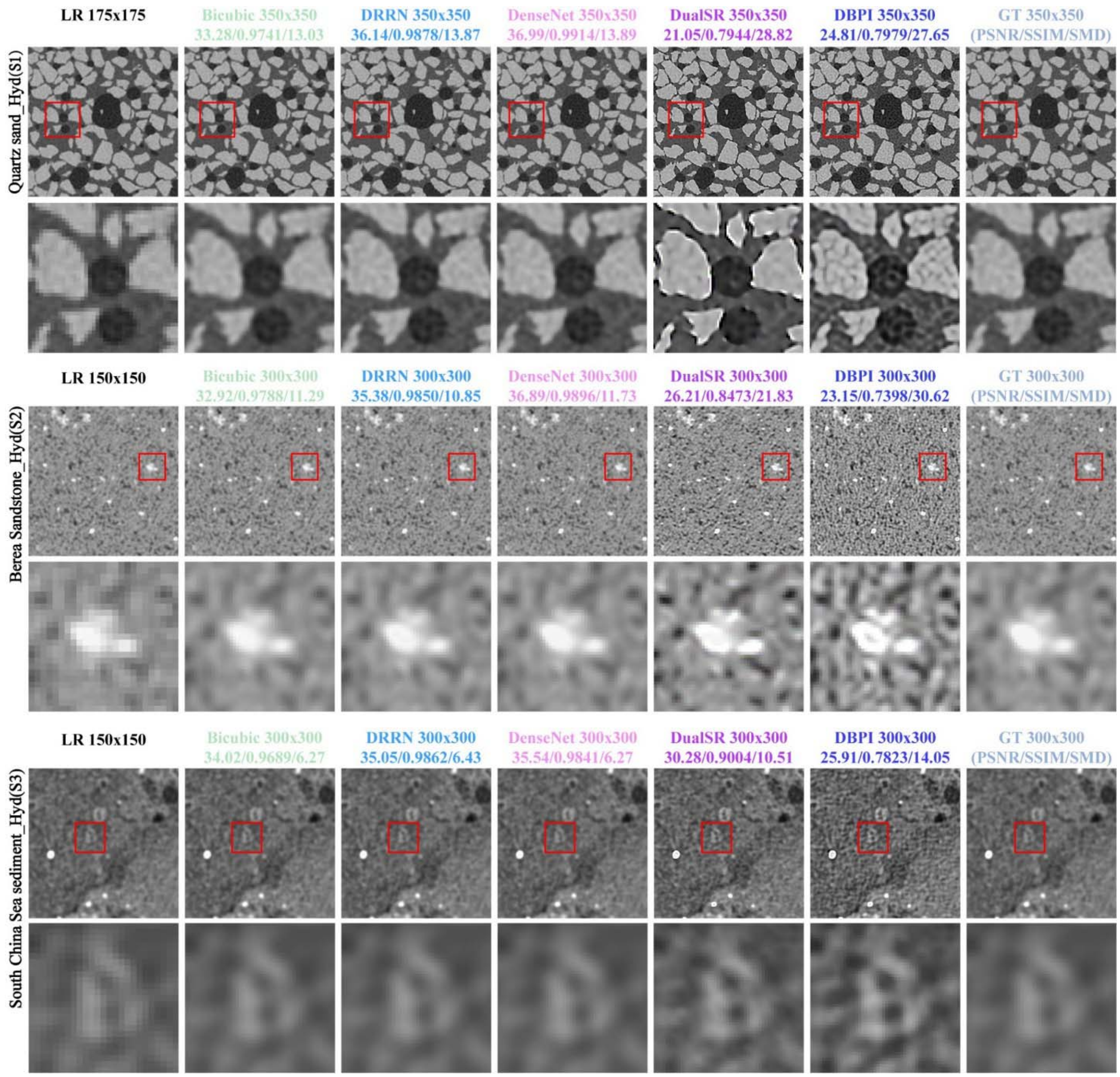

Figure 6. Zoom-in visual comparison for 2× SR on synthesized LR micro-CT images (2× downscaling from the GT) by different SR methods. The images in every other rows are an enlargement of the red box of the previous row. The numbers in the right of different SR methods are the resolution of images and the below are PSNR, SSIM and SMD, respectively. (GT: Ground Truth Image)

The qualitative comparison of different SR methods is shown in Figure 6. It can be seen that the ground truth images are visually blurry, although the ground truth images in our proposed datasets have the best resolution which consider the balance between FOV and resolution. Although the PSR methods tend to achieve comparable visual quality with the ground truth images, however, the performance are limited by the quality of the ground truth images because of the direct mapping learning between the LR-HR images. By contrast, the SLSR methods produce much clearer and sharper results which are not limited to the ground truth images. Furthermore, the SLSR methods enhance the resolution using only the input information of the image itself, no additional fake information would be introduced by contrast to GAN. This benefits us to find more realistic details from the super-resolved micro-CT images and will be helpful to the construction of super-resolution digital core, the accurate image segmentation, and the accurate calculation of physical property parameters.

3.2 Experiments on HR Micro-CT Images for 2× SR

The results above have demonstrated the effectiveness of the SLSR methods on micro-CT images super-resolution. Since our goal is to get higher resolution images beyond the HR images (HR-SR), we need to apply super-resolution on the actually obtained HR micro-CT images, these SR methods mentioned above were also carried out on the HR micro-CT images. The model parameters of DRRN-PSR and SRDenseNet-PSR used in this experiment are the same as those in section 3.1.

The quantitative and qualitative comparison of different SR methods for 2× super-resolution on the actually obtained HR micro-CT images in our proposed three datasets are shown in Figure 7 and Figure 8, respectively. As shown in Figure 7, only the SMD is used in that the SR results have no counterpart higher resolution images to measure the performance of both PSNR and SSIM. In Figure 7, an intuitive comparison of the SMD performance of different SR methods on test sets of the three datasets are illustrated in each graph. In all graphics, the much higher SMD values were achieved by SLSR methods and among them the DBPI-SLSR did the best.

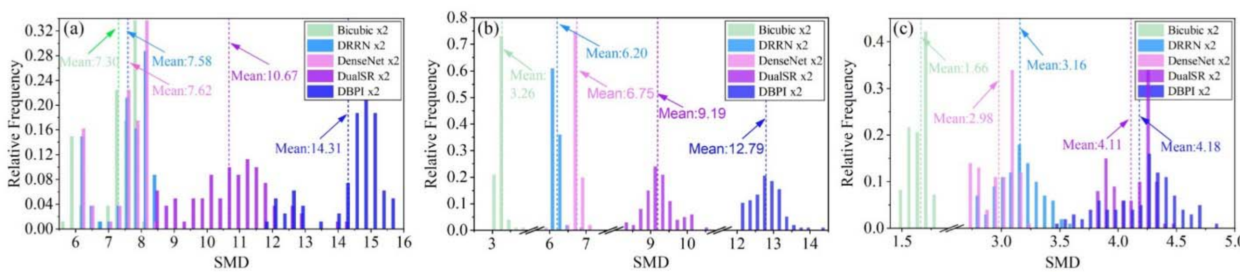

Figure 7. SMD comparison for 2× SR on the actually obtained HR micro-CT images in *Quartz sand_Hyd(S1) (a), Berea Sandstone_Hyd(S2) (b), South China Sea sediment_Hyd(S3) (c)*. by different SR methods.

Looking at Figure 8, it is apparent that the SLSR methods created much clearer and sharper images compared to not only the results of both bicubic interpolation and PSR methods, but also the actually obtained HR images. Internal feature information was extracted by self-supervised learning to make up the missing details for the magnified image, in that improves image contrast while enhancing the resolution. However, one unexpected problem that emerged from the results was that the noise of CT images was enlarged as well, especially in the hydrate samples with lower grain size. One issue is the gray values near the center of sands are much lower than the edge region as shown in the enlargement images in the last two rows of Figure 8. A possible explanation for this might be that the instrumental noise that resulted in the lower gray values near the center than the edge region. Limited by image resolution, it is difficult to find this phenomenon, however, it becomes apparent as image resolution increases. This phenomenon can be more easily found in the results of bicubic interpolation and PSR methods, and was more obvious in the results of SLSR methods, which shows self-supervised learning can explore more image details in comparison. In accordance with the present results, although self-supervised learning amplifiers noise, previous studies have demonstrated that it can produce clearer and sharper results even on the actually obtained HR micro-CT images. However, additional uncertainty arises from the results that the great gray scale difference of sands components may affects the accuracy of image segmentation and physical property parameter calculation. Therefore, in order to attenuate the noise effects in further study, we do efforts to make some improvements based on DBPI-SLSR methods to make better results.

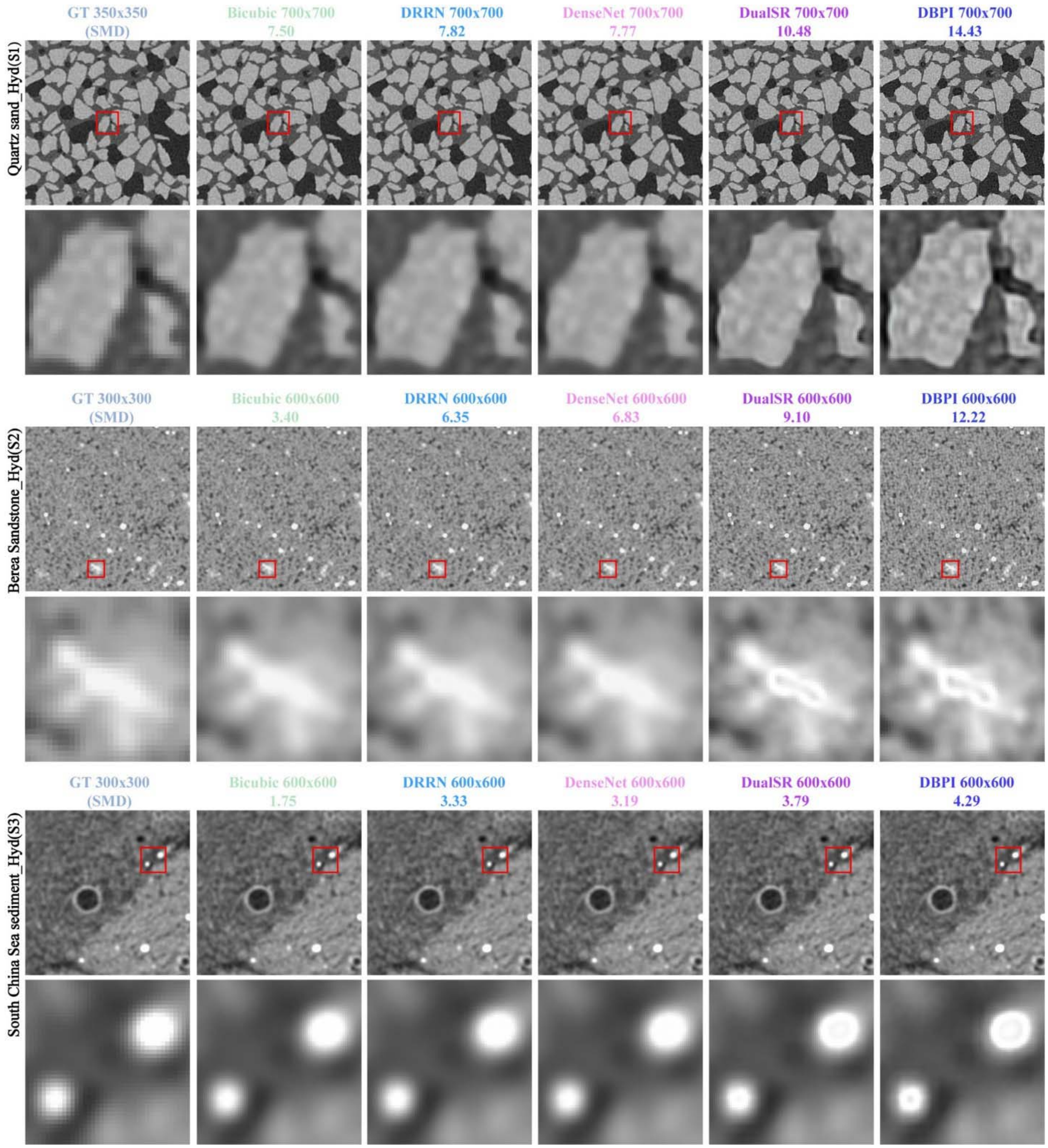

Figure 8. Zoom in visual comparison for 2× SR on the actually obtained HR micro-CT images. The images in every other rows are an enlargement of the red box of the previous row. The numbers in the right of different SR methods are the resolution of images and the below is SMD values.

3.3 Results of I-DBPI for 2× SR on HR Micro-CT Images

The results obtained from the previous experiments show that the self-supervised learning has ability to produce high quality images while enhancing the resolution by training on a single input image compared to the PSR methods, however, affects by image noise. In order to create better results, some improvements were made based on the DBPI-SLSR that performed the best in previous experiments. The yields from the I-DBPI were compared with the DBPI-SLSR on *S2* and *S3* in the aspect of quantitative and qualitative which shown in Figure 9 and Figure 10, respectively.

Figure 10 compares the image quality which is our primary concern. As shown in the enlargement images of the red box in Figure 10, the gray values of sand components in the result of the I-DBPI change smoothly and the gray range is almost the same as the original image comparing with the unsatisfactory results of the DBPI-SLSR. In addition, the contrast and sharpness of images are further enhanced by the I-DBPI, which can not only be seen intuitively from the image, but also from the distribution of SMD in Figure 9 in which the mean SMD of the I-DBPI are higher than the DBPI-SLSR in those two datasets. Moreover, the dispersion degree of SMD decreases after the improvement, indicating that the improved method is more stable and precise. Why does the improved approach work better, one passible explain is that the self-supervised learning can learn much more feature information from a large amount of data than from only one image, which can help recover more image details and make distribution stable.

It is worth mentioning that the process time is greatly reduced by the I-DBPI, especially when dealing with large numbers of images. More intuitively, the DBPI-SLSR method needs around two minutes to super-resolve an image of size 500 × 500 on an NVIDIA TITAN RTX GPU, however, the time can be shortened to two seconds for per image by I-DBPI.

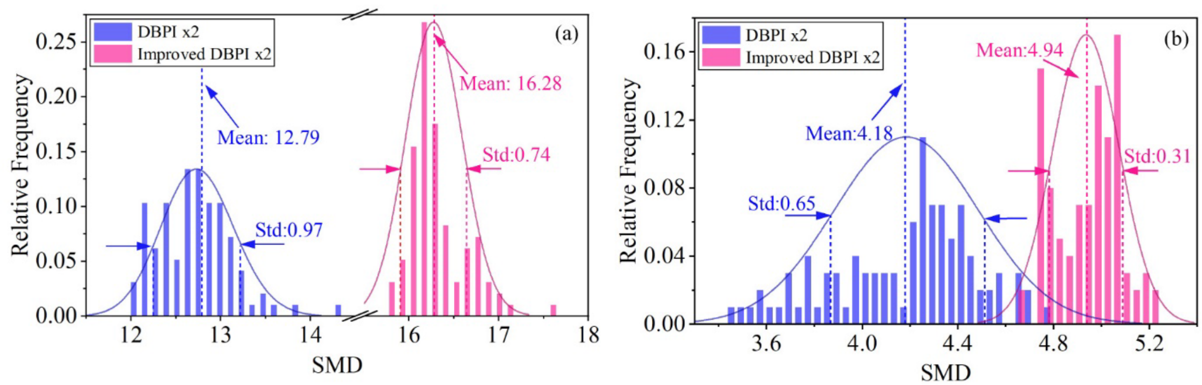

Figure 9. SMD comparison for 2× SR on the actually obtained HR images in *Berea Sandstone_Hyd(S2) (a), South China Sea sediment_Hyd(S3)* (b). The SMD scores are achieved by DBPI-SLSR and I-DBPI.

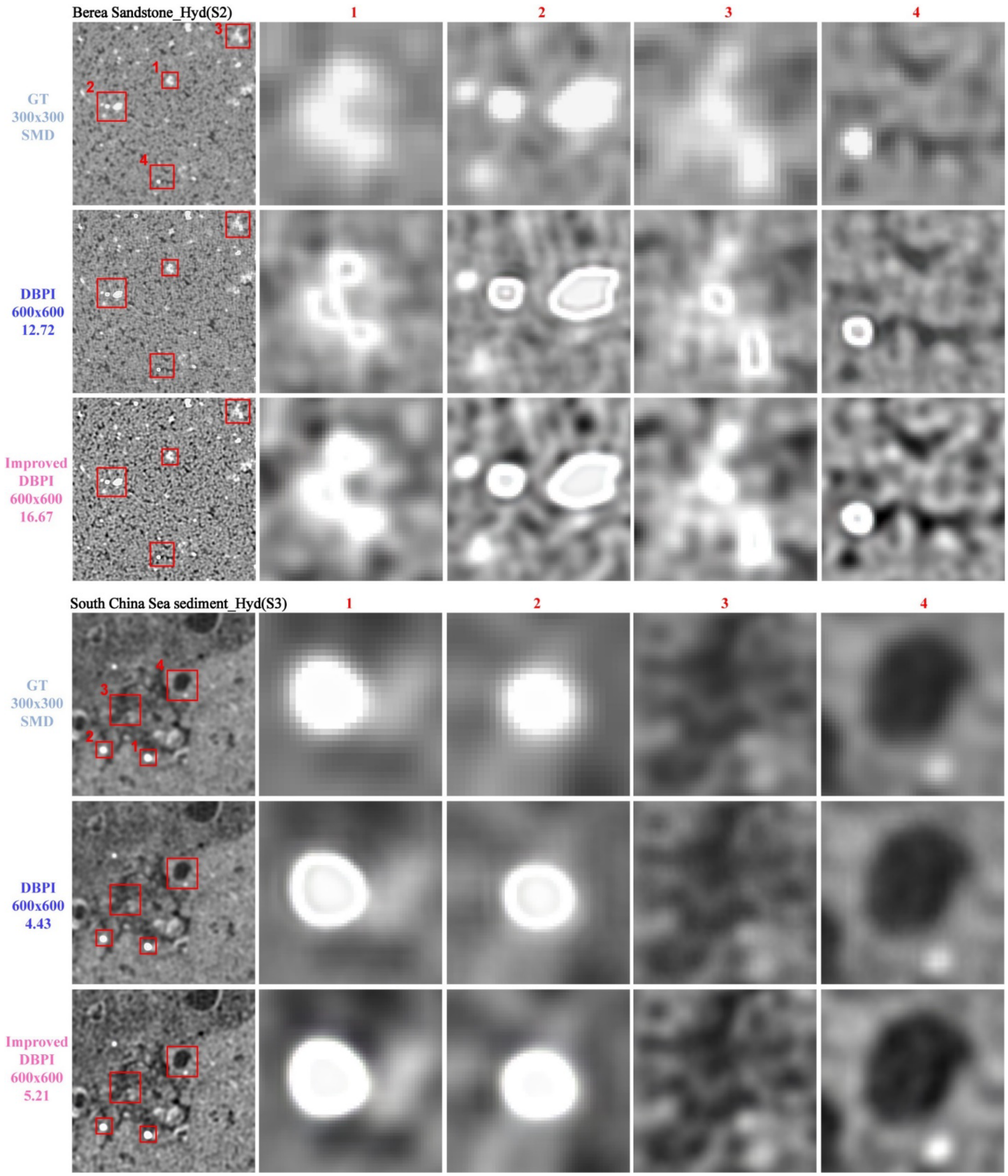

Figure 10. Zoom in visual comparison for 2× SR on the actually obtained HR micro-CT images. The images in every other cols are an enlargement of the red box of the first col. The numbers below the different SR methods are the resolution of images and SMD, respectively.

3.4 Porosity Comparison of *Berea Sandstone_Hyd(S3)* Images

The previous studies have demonstrated that high quality micro-CT images can be produced by the I-DBPI. Theoretically, high quality micro-CT images will helpful to the precise image segmentation, and the accurate calculation of physical property parameters. In order to verify this, the original HR images, the result images of bicubic interpolation, DBPI-SLSR, and theI-DBPI of *S3* were used to calculate the porosity for a comparison.

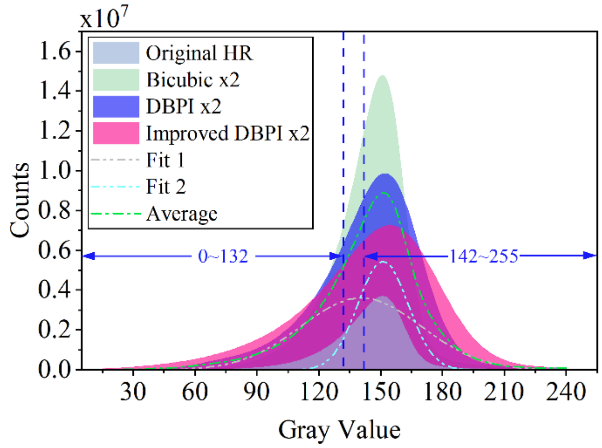

Figure 11. Image gray value distribution of *Berea Sandstone_Hyd(S3)* images by different SR methods. The curves filled with different colors are the original distribution. The dark green curve is the average of those original curves and the gray and light green curves are the fitting results of gaussian function. The gray range 0-132 and 142-255 are the initial segmentation threshold of pore and skeleton.

The gray value distribution of different SR methods is shown in Figure 11. However, there is no obvious trough in the gray value distribution, the watershed image segmentation algorithm is used as the segmentation threshold of pore and skeleton cannot be selected from the gray value distribution directly. In order to unify the initial threshold of watershed segmentation algorithm, all the gray value distribution were averaged firstly which is illustrated by the dark green curve in Figure 11. After that, the average distribution curve was bimodally fitted by gaussian function. The two curves obtained by the fitting are shown as gray and light green curves in Figure 11, whose intersection point is near the gray value of 140. Overall consideration, the initial segmentation threshold of pore and skeleton were set to 0-132 and 142-255 respectively, and the watershed algorithm is responsible for automatically inflating the remaining regions to different boundaries. The results of watershed image segmentation show that the boundary between the pore and skeleton of the image processed by DBPI and the improved DBPI is more obvious compared with the original image and bicubic upscaling image, which leads to more accurate segmentation results.

Based on the segmentation results, the porosity of 2D slices and the average porosity of 3D image were calculated by Avizo numerical simulation software. The porosity distribution of 2D slices is shown in Figure 12. From the figure, it can be seen that although the porosity values calculated by different methods varies greatly, the distribution trends of porosity are basically the same. Among them, the porosity distribution of the improved DBPI is more precise. Figure 13 shows the 3D pore skeleton image of the *Berea Sandstone Hyd* sample. The average porosity of the 3D sample are 0.2814, 0.2970, 0.3363 and 0.3658 respectively, which differ widely. For a better comparison, we measured the mean porosity by the saturated liquid weighing method which is a physical means. The porosity measured by the saturated liquid weighing method was 0.3648 that is basically consistent with the result of the improved DBPI, which indicates that the improved DBPI can help to distinguish the porosity and skeleton better, so as to calculate more accurate porosity parameters.

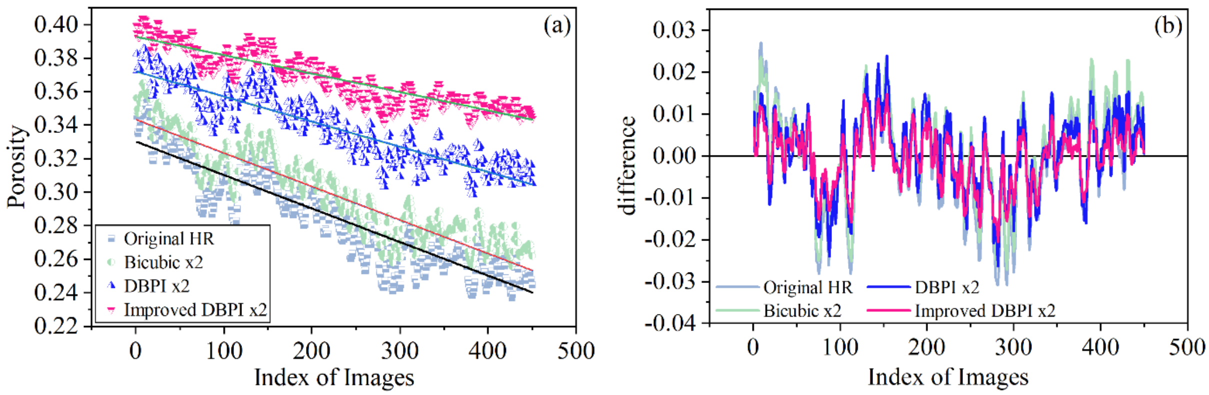

Figure 12. Porosity distribution of 2D slices of the *Berea Sandstone Hyd(S3)* sample by different SR methods, and the linear lines are the fitted trend of porosity distribution (a); Difference of the porosity and its corresponding fitted lines (b).

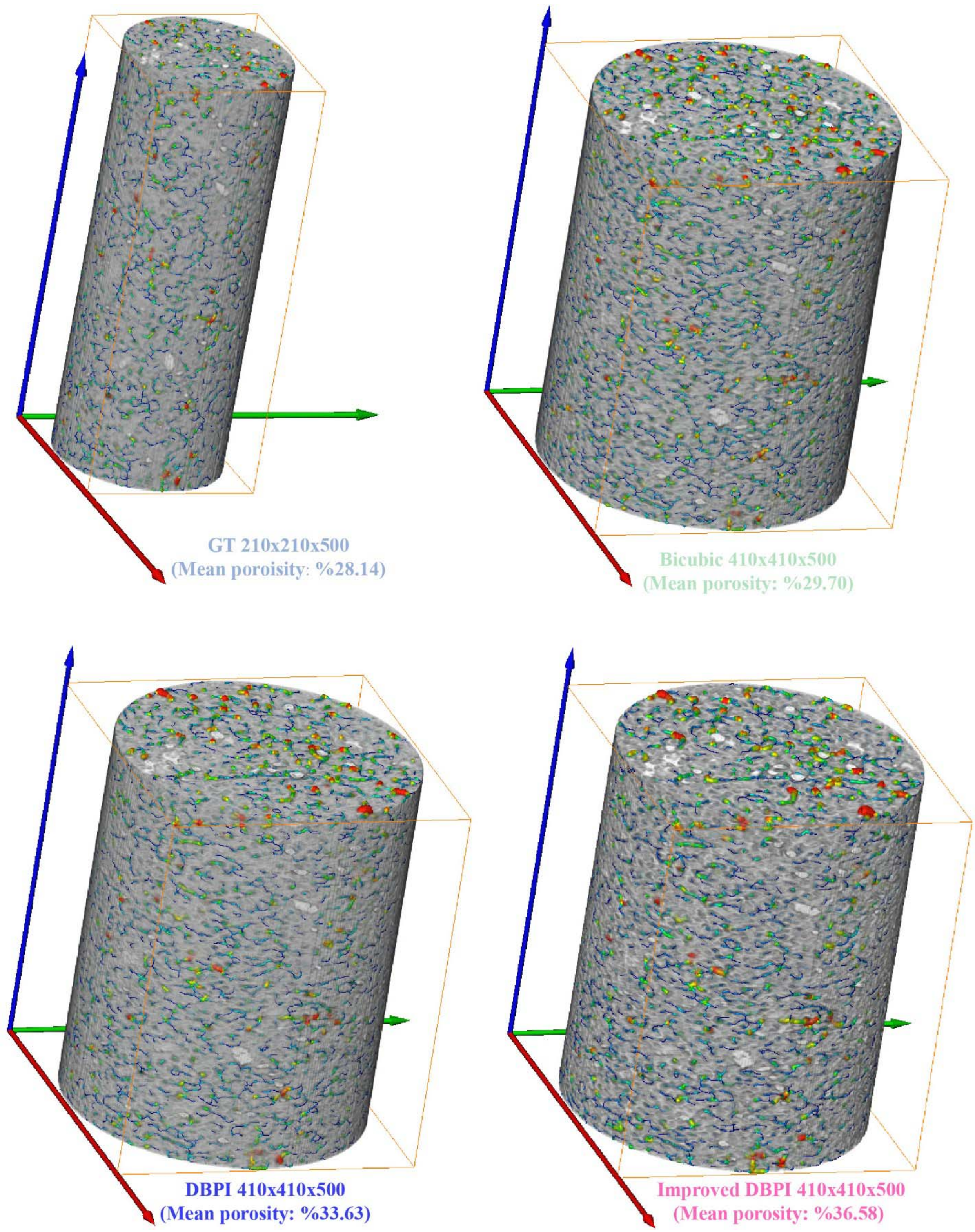

Figure 13. The 3D pore skeleton image of the *Berea Sandstone Hyd* sample by different SR methods.

4 Conclusions

We break through the hardware limits to enhance the resolution and enlarge the pore spatial structure of hydrate CT images by developing the Self-Supervised Learning-Based SR (SLSR) methods. Image Pairs-Based SR (PSR) methods are also used as the most useful SR means in the past. We do that by first training the PSR with synthesized LR-HR Micro-CT images and the SLSR with every single input micro-CT image based on decreased porosity hydrate CT image datasets S1, S2 and S3. The image quality of SLSR results exceed the PSR a lot on synthesized LR images (LR-HR) as well as the actually obtained HR images (HR-SR), as SLSR are not limited by the actually obtained HR images. SLSR are affected by the grain size and porosity of samples a little. Qualitative and quantitative evaluation tell us the DBPI-SLSR has the best performance on visual sense and clarity, however, it has certain limitations in terms of that it is greatly affected by image noise, leading to a large change degree of gray value near the center of the sand component in CT image before and after reconstruction. Our method trained on big datasets improved from DBPI-SLSR further improve the sharpness and contrast of the images while mitigating the above shortcomings. This research has made some efforts on the issue of exceeding the limitation of imaging systems on FOV and resolution.

After that, the watershed image segmentation algorithm was used to segment the pores and skeletons in the original CT image, the bicubic interpolation image, the results obtained by DBPI-SLSR and the results obtained by the improved DBPI-SLSR. The segmentation results showed that the watershed algorithm could expand from the set threshold to the boundary between the pores and the skeleton more accurately, since the improved DBPI-SLSR increased the gray difference between the pores and the skeleton in the micro-CT images. Finally, the porosity parameter was calculated based on the segmentation results, and the mean porosity of the result of the improved DBPI-SLSR was closest to that measured by the saturated liquid weighing method.

Taken together, the results of this study suggest that compared to PSR methods, the SLSR methods can promote the hydrate-bearing sediment micro-CT images qualities in resolution and clarity, which benefits accurate segmentation and calculation of physical property parameters, especially the improved DBPI-SLSR that proposed by us.

Acknowledgments

This work was supported by National Natural Science Foundation of China (Grant No. 41976205) and open research fund program of Zhoushan Field Scientific Observation and Research Station for Marine Geo-hazards, China Geological Survey (No. ZSORS-22-12).

References

Bell-Kligler, S., Shocher, A., & Irani, M. (2019). Blind super-resolution kernel estimation using an internal-gan. *Advances in Neural Information Processing Systems*, *32*.

Cai, J., Zeng, H., Yong, H., Cao, Z., & Zhang, L. (2019). Toward real-world single image super-resolution: A new benchmark and a new model. In *Proceedings of the IEEE/CVF International Conference on Computer Vision* (pp. 3086-3095). https://doi.org/10.1109/ICCV.2019.00318

Chen, C., Xiong, Z., Tian, X., Zha, Z. J., & Wu, F. (2019). Camera lens super-resolution. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 1652-1660). https://doi.org/10.1109/CVPR.2019.00175.

Chen, H., He, X., Teng, Q., Sheriff, R. E., Feng, J., & Xiong, S. (2020). Super-resolution of real-world rock microcomputed tomography images using cycle-consistent generative adversarial networks. *Physical Review E*, *101*(2), 023305. https://doi.org/10.1103/PhysRevE.101.023305.

Da Wang, Y., Armstrong, R. T., & Mostaghimi, P. (2019). Enhancing resolution of digital rock images with super resolution convolutional neural networks. *Journal of Petroleum Science and Engineering*, *182*, 106261. https://doi.org/10.1016/j.petrol.2019.106261.

Dong, C., Loy, C. C., & Tang, X. (2016, October). Accelerating the super-resolution convolutional neural network. In *European conference on computer vision* (pp. 391-407). Springer, Cham. https://doi.org/10.1007/978-3-319-46475-6_25.

Emad, M., Peemen, M., & Corporaal, H. (2021). DualSR: Zero-shot dual learning for real-world super-resolution. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision* (pp. 1630-1639). https://doi.org/10.1109/WACV48630.2021.00167.

by prior and perceptual information. *Computers & Geosciences*, *157*, 104939. https://doi.org/10.1016/j.cageo.2021.104939.

Janssens, N., Huysmans, M., & Swennen, R. (2020). Computed tomography 3D super-resolution with generative adversarial neural networks: Implications on unsaturated and two-phase fluid flow. *Materials*, *13*(6), 1397. https://doi.org/10.3390/ma13061397.

Kim, G., Park, J., Lee, K., Lee, J., Min, J., Lee, B., ... & Ko, H. (2020). Unsupervised real-world super resolution with cycle generative adversarial network and domain discriminator. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops* (pp. 456-457). https://doi.org/10.1109/CVPRW50498.2020.00236.

Kim, J., Jung, C., & Kim, C. (2020). Dual back-projection-based internal learning for blind super-resolution. *IEEE Signal Processing Letters*, *27*, 1190-1194. https://doi.org/10.1109/LSP.2020.3005043.

Kim, J., Lee, J. K., & Lee, K. M. (2016). Accurate image super-resolution using very deep convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 1646-1654). https://doi.org/10.1109/CVPR.2016.182.

Ledig, C., Theis, L., Huszár, F., Caballero, J., Cunningham, A., Acosta, A., & Shi, W. (2017). Photo-realistic single image super-resolution using a generative adversarial network. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 4681-4690). https://doi.org/10.1109/CVPR.2017.19.

Li, Z., Teng, Q., He, X., Yue, G., & Wang, Z. (2017). Sparse representation-based volumetric super-resolution algorithm for 3D CT images of reservoir rocks. *Journal of Applied Geophysics*, *144*, 69-77. https://doi.org/10.1016/j.jappgeo.2017.04.013.

Liu, C., Meng, Q., Hu, G., Li, C., Sun, J., He, X., ... & Liang, J. (2017). Characterization of hydrate-bearing sediments recovered from the Shenhu area of the South China Sea. *Interpretation*, *5*(3), SM13-SM23. https://doi.org/10.1190/INT-2016-0211.1.

Niu, Y., Jackson, S. J., Alqahtani, N., Mostaghimi, P., & Armstrong, R. T. (2021). A comparative study of paired versus unpaired deep learning methods for physically enhancing digital rock image resolution. *arXiv preprint arXiv:2112.08644*. https://doi.org/10.48550/arXiv.2112.08644

Niu, Y., Wang, Y. D., Mostaghimi, P., Swietojanski, P., & Armstrong, R. T. (2020). An innovative application of generative adversarial networks for physically accurate rock images with an unprecedented field of view. *Geophysical Research Letters*, *47*(23), e2020GL089029. https://doi.org/10.1029/2020GL089029.

Priest, J. A., Best, A. I., & Clayton, C. R. (2005). A laboratory investigation into the seismic velocities of methane gas hydrate-bearing sand. *Journal of geophysical research: solid earth*, *110*(B4). https://doi.org/10.1029/2004JB003259.

Ren, S. R., Liu, Y., Liu, Y., & Zhang, W. (2010). Acoustic velocity and electrical resistance of hydrate bearing sediments. *Journal of petroleum science and engineering*, *70*(1-2), 52-56. https://doi.org/10.1016/j.petrol.2009.09.001.

Shocher, A., Cohen, N., & Irani, M. (2018). “zero-shot” super-resolution using deep internal learning. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 3118-3126). https://doi.org/10.1109/CVPR.2018.00329.

Tai, Y., Yang, J., & Liu, X. (2017). Image super-resolution via deep recursive residual network. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 3147-3155). https://doi.org/10.1109/CVPR.2017.298.

Wang, D., Wang, C., Li, C., Liu, C., Lu, H., Wu, N., & Meng, Q. (2018). Effect of gas hydrate formation and decomposition on flow properties of fine-grained quartz sand sediments using X-ray CT based pore network model simulation. *Fuel*, *226*, 516-526. https://doi.org/10.1016/j.fuel.2018.04.042.

Wang, X., Yu, K., Wu, S., Gu, J., Liu, Y., Dong, C., & Change Loy, C. (2018). Esrgan: Enhanced super-resolution generative adversarial networks. In *Proceedings of the European conference on computer vision (ECCV) workshops* (pp. 0-0). https://doi.org/10.1007/978-3-030-11021-5_5.

Wildenschild, D., & Sheppard, A. P. (2013). X-ray imaging and analysis techniques for quantifying pore-scale structure and processes in subsurface porous medium systems. *Advances in Water resources*, *51*, 217-246. https://doi.org/10.1016/j.advwatres.2012.07.018.

You, C., Li, G., Zhang, Y., Zhang, X., Shan, H., Li, M., ... & Wang, G. (2019). CT super-resolution GAN constrained by the identical, residual, and cycle learning ensemble (GAN-CIRCLE). *IEEE transactions on medical imaging*, *39*(1), 188-203. https://doi.org/10.1109/TMI.2019.2922960.

Yuan, Y., Liu, S., Zhang, J., Zhang, Y., Dong, C., & Lin, L. (2018). Unsupervised image super-resolution using cycle-in-cycle generative adversarial networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops* (pp. 701-710). https://doi.org/10.1109/CVPRW.2018.00113.

Zhang, L., Ge, K., Wang, J., Zhao, J., & Song, Y. (2020). Pore-scale investigation of permeability evolution during hydrate formation using a pore network model based on X-ray CT. *Marine and Petroleum Geology*, *113*, 104157. https://doi.org/10.1016/j.marpetgeo.2019.104157.

Zhang, X., Chen, Q., Ng, R., & Koltun, V. (2019). Zoom to learn, learn to zoom. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 3762-3770). https://doi.org/10.1109/CVPR.2019.00388.

Zhu, J. Y., Park, T., Isola, P., & Efros, A. A. (2017). Unpaired image-to-image translation using cycle-consistent adversarial networks. In *Proceedings of the IEEE international conference on computer vision* (pp. 2223-2232). https://doi.org/10.1109/ICCV.2017.244.