Super-resolution of real-world rock microcomputed tomography images using cycle-consistent generative adversarial networks

Honggang Chen [1,*] Xiaohai He [1,2,†] Qizhi Teng [1,‡] Raymond E. Sheriff [3,§] Junxi Feng [1,‖] and Shuhua Xiong [1,¶]

[1]*College of Electronics and Information Engineering, Sichuan University, Chengdu 610065, China*
[2]*Key Laboratory of Wireless Power Transmission of Ministry of Education, Chengdu 610065, China*
[3]*School of Engineering, University of Bolton, Bolton BL35AB, United Kingdom*

Digital rock imaging plays an important role in studying the microstructure and macroscopic properties of rocks, where microcomputed tomography (MCT) is widely used. Due to the inherent limitations of MCT, a balance should be made between the field of view (FOV) and resolution of rock MCT images—a large FOV at low resolution (LR) or a small FOV at high resolution (HR). However, large FOV and HR are both expected for reliable analysis results in practice. Super-resolution (SR) is an effective solution to break through the mutual restriction between the FOV and resolution of rock MCT images, for it can reconstruct an HR image from a LR observation. Most of the existing SR methods cannot produce satisfactory HR results on real-world rock MCT images. One of the main reasons for this is that paired images are usually needed to learn the relationship between LR and HR rock images. However, it is challenging to collect such a dataset in a real scenario. Meanwhile, the simulated datasets may be unable to accurately reflect the model in actual applications. To address these problems, we propose a cycle-consistent generative adversarial network (CycleGAN)-based SR approach for real-world rock MCT images, namely, SRCycleGAN. In the off-line training phase, a set of unpaired rock MCT images is used to train the proposed SRCycleGAN, which can model the mapping between rock MCT images at different resolutions. In the on-line testing phase, the resolution of the LR input is enhanced via the learned mapping by SRCycleGAN. Experimental results show that the proposed SRCycleGAN can greatly improve the quality of simulated and real-world rock MCT images. The HR images reconstructed by SRCycleGAN show good agreement with the targets in terms of both the visual quality and the statistical parameters, including the porosity, the local porosity distribution, the two-point correlation function, the lineal-path function, the two-point cluster function, the chord-length distribution function, and the pore size distribution. Large FOV and HR rock MCT images can be obtained with the help of SRCycleGAN. Hence, this work makes it possible to generate HR rock MCT images that exceed the limitations of imaging systems on FOV and resolution.

I. INTRODUCTION

X-ray microcomputed tomography (MCT) is a widely used imaging technology to obtain three-dimensional (3D) images of porous media, such as rock, soil, wood, and ceramic. For digital rock imaging, MCT plays an important role. The 3D images produced by MCT show the microstructures of rock samples and therefore can be used to analyze the macroscopic properties of rocks, such as permeability and conductivity [1–4]. To obtain reliable analysis results, rock MCT images with high resolution (HR) and a large field of view (FOV) are desired. However, this is challenging in practice because of the inherent limitations of MCT. More specifically, a trade-off should be made between the FOV and resolution of MCT

images [5,6]. Actually, rock MCT images with a large FOV are typically of low resolution (LR), which causes increased difficulty of property analysis as many fine structures (e.g., pores) cannot be captured. On the contrary, HR images are usually with a small FOV, thereby resulting in decreased representativeness. Recent studies [7–11] show that this bottleneck can be addressed to some extent via super-resolution (SR), which maps a LR input to a space of higher resolution [12]. HR rock MCT images with a large FOV can be obtained by applying SR algorithms to the collected LR MCT images.

SR is an active research direction in the image and video processing area and extensive studies [13–45] have been done on this topic in recent years. Existing SR methods can be roughly classified into the following categories based on the object: Single image SR [13–40], multiframe SR [41,42], and video SR [43–45]. Single image SR refers to the reconstruction of an image with higher resolution from a single LR observation. For multiframe SR, a set of correlated LR images is used as the input to estimate an HR image. Correspondingly, video SR aims to produce an HR video from the acquired LR video. Overall, single image SR has received greater attention among the three groups for it is more practical in

----
*honggang_chen@yeah.net
†Corresponding author: hxh@scu.edu.cn
‡qzteng@scu.edu.cn
§R.Sheriff2@bolton.ac.uk
‖fengjx2011@gmail.com
¶xiongsh@scu.edu.cn

most cases. Traditional single image SR methods [13–19] usually combine the imaging model of LR observations and the prior knowledge of HR images to formulate an objective function, which is optimized to produce an HR estimate. Commonly used priors include sparsity [13], local smoothness [14], nonlocal similarity [15], low rank [17,18], gradient information [18,19], etc. Learning-based single image SR approaches [20–40] have become mainstream in recent years. This kind of approach learns the mapping between LR-HR image pairs in the model training stage, and LR test images are super-resolved using the learned mapping in the testing phase. Neighbor embedding [20], sparse representation [21], neighbor regression [22,23], random forest [24], and deep neural networks [25–40] are effective models for characterizing the mapping from LR to HR images. In particular, deep convolutional neural networks (CNNs) have shown excellent performance on a variety of computer vision tasks [46–59], and SR is no exception. Therefore, this work mainly focuses on deep CNN-based single image SR approaches.

The super-resolution convolutional neural network (SRCNN) proposed by Dong *et al.* [25] is one of the representative works for deep CNN-based SR algorithms. SRCNN consists of three convolutional layers, which perform block extraction, mapping, and reconstruction, respectively. SRCNN achieved state-of-the-art performance when first proposed, although its structure is simple. Subsequently, several of the networks were developed for single image SR. Typically, a deconvolutional layer used for upsampling was incorporated into the fast super-resolution convolutional neural network (FSRCNN) [26]. Shi *et al.* [27] developed a subpixel layer to produce the HR output from LR features. The deconvolutional layer and the subpixel layer were widely used in subsequent studies [32–34] because they are helpful for improving efficiency. Kim *et al.* [28] presented a very deep (20 layers) SR network (VDSR) using residual learning and gradient clipping. This work shows that within a certain

range, the quality of the super-resolved result improves as the network depth increases. Overall, follow-up networks for SR became deeper and deeper. In order to avoid excessive parameters of deep neural networks, recursive structures were used in the deeply-recursive convolutional network (DRCN) [29] and the deep recursive residual network (DRRN) [30]. Lim *et al.* [31] greatly improved the model capacity via enhancing network depth and width, and the designed enhanced deep residual network for SR (EDSR) won the New Trends in Image Restoration and Enhancement (NTIRE) challenge 2017 [60]. Zhang *et al.* [33] combined the dense block and the residual block, and presented a deep residual dense network (RDN) for single image SR. Subsequently, the authors developed a deep residual channel attention network (RCAN, over 400 layers) for SR via combining the residual in residual structure and the channel attention mechanism [34]. The methods mentioned above focus more on the objective parameters [e.g., peak signal-to-noise ratio (PSNR) and structural similarity index (SSIM)] of the recovered HR images. RDN [33] and RCAN [34] are state-of-the-art approaches in terms of PSNR and SSIM. However, these indexes may not be consistent with the human visual system in some cases. To address this problem, a series of perceptual-driven SR approaches [35–40] have been presented. The pioneer of this kind of algorithm is the generative adversarial network-based SR method (SRGAN) [35]. Aiming to improve the visual quality of the HR estimate, SRGAN was trained with perceptual and adversarial losses. The loss functions for network training affect SR performance greatly. Hence, a variety of losses were designed, including the local texture matching loss [36] and the contextual loss [37]. Park *et al.* [38] introduced an additional discriminator into a GAN-based SR network to produce more realistic results. Yuan *et al.* [39] and You *et al.* [40] developed unsupervised and semisupervised SR frameworks using the cycle-consistent generative adversarial network (CycleGAN) [61]. Overall, compared with objective
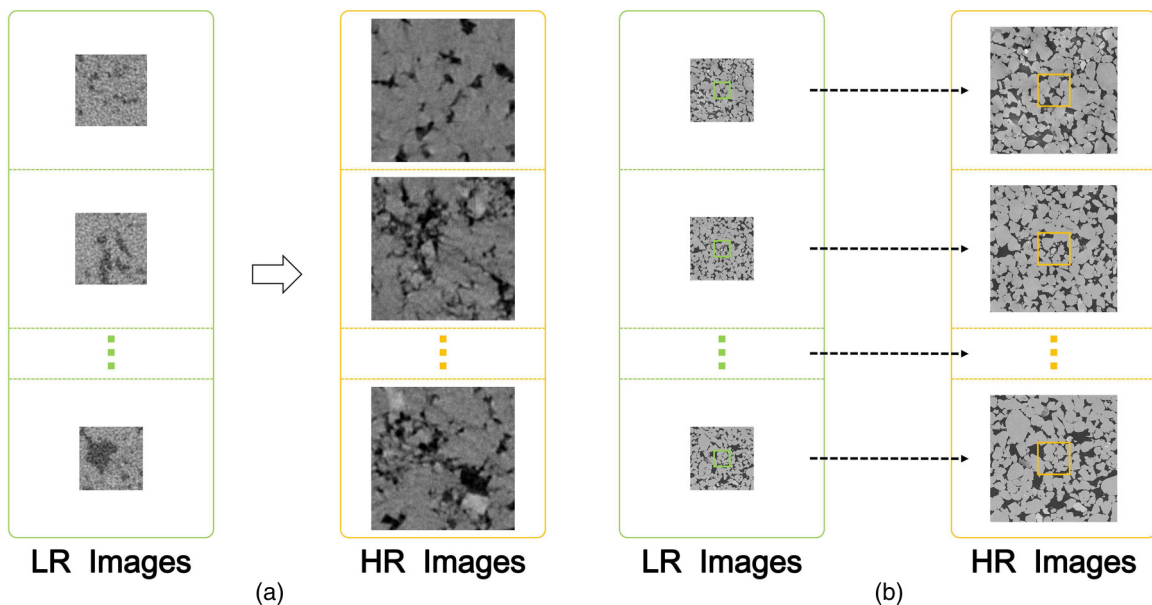

FIG. 1. Examples of unpaired and paired LR-HR images. (a) Unpaired LR-HR images in the real world. (b) Paired LR-HR images produced by simulation. Please zoom in to view details and make comparisons.
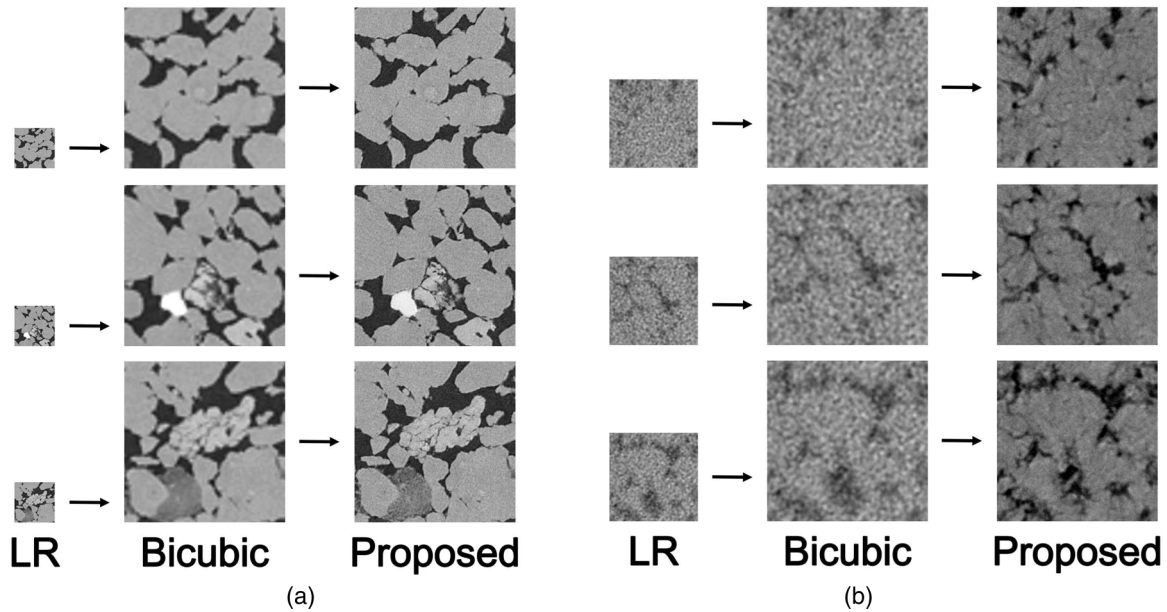
FIG. 2. Examples of rock MCT images SR. (a) Results on simulated LR images (SR factor: 4×; from left to right: LR, bicubic interpolation, proposed SRCycleGAN). (b) Results on real-world LR images (SR factor: 1.8×; from left to right: LR, bicubic interpolation, proposed SRCycleGAN). Please zoom in to view details and make comparisons.

parameters-driven SR methods, perceptual-driven approaches can produce visually more pleasant HR images.

With the rapid development of SR, the SR of rock MCT images is attracting more and more attention for its significant application potential in digital rock imaging. Wang *et al.* [7] improved the resolution of MCT images of rock samples using neighbor embedding. For the reconstructed HR rock image, the low-frequency part is provided by the LR MCT input, and the HR scanning electron microscopy (SEM) provides high-frequency information. Wang *et al.* [8,9] trained a set of networks including SRCNN, EDSR, and SRGAN for rock MCT images SR. There are also some studies for the SR of 3D rock images. For example, Li *et al.* [10] presented a sparse representation-based 3D volumetric SR framework. Wang *et al.* [11] extended the VDSR for 3D rock MCT images SR via introducing 3D convolution. These studies demonstrate that the resolution of rock MCT images can be enhanced by performing SR.

To the best of our knowledge, paired LR-HR training images are necessary for most of the existing learning-based SR approaches for rock MCT images [7–11]. However, it is challenging to collect paired LR-HR images of the same rock sample in a real scenario. Even though we can scan the same sample using different devices to generate images with different resolutions [as shown in Fig. 1(a)], accurate registration between images is very difficult. One solution for this problem is to produce LR-HR image pairs artificially [as shown in Fig. 1(b)]. Given an HR rock image, the corresponding LR image can be generated by downsampling. Producing accurately paired LR-HR images is convenient and efficient in this way. However, the relationship between simulated LR-HR image pairs may not reflect the actual mapping because the real-world LR images usually suffer from more complex degradations (nonideal blur, sensor noise,

compression artifacts, etc.), resulting in poor SR performance in practical applications.

This paper proposes an effective SR approach for real-world rock MCT images, namely, SRCycleGAN. Some SR results both on the simulated and real-world LR rock MCT images are illustrated in Fig. 2. It can be observed that the super-resolved results by the proposed SRCycleGAN are much better than LR images as well as the results of bicubic interpolation, with more details and clearer edges. Apparently, the SR process would be helpful for subsequent process and analysis. The main contributions of this work are as follows:

(1) To overcome the absence of paired training examples, we propose to consider the SR of real-world rock MCT images as the unpaired image-to-image translation. LR rock images and the associated HR rock images are assumed to belong to two related domains.

(2) We present a CycleGAN-based SR approach that is well suited to the SR of real-world rock MCT images, where paired LR-HR training images are difficult to obtain.

(3) Extensive experiments are performed to demonstrate the effectiveness and superiority of the proposed SRCycleGAN. In particular, a set of statistical parameters is compared between the SR results and targets to validate the reliability of SRCycleGAN.

(4) This work shows that, with an effective SR method, the inherent hardware limitations of digital rock imaging systems on the FOV and resolution can be compensated to some extent.

The remainder of this paper is organized as follows. Section II presents the proposed SRCycleGAN in detail. Experimental results and discussion are shown in Sec. III. Section IV concludes this study.
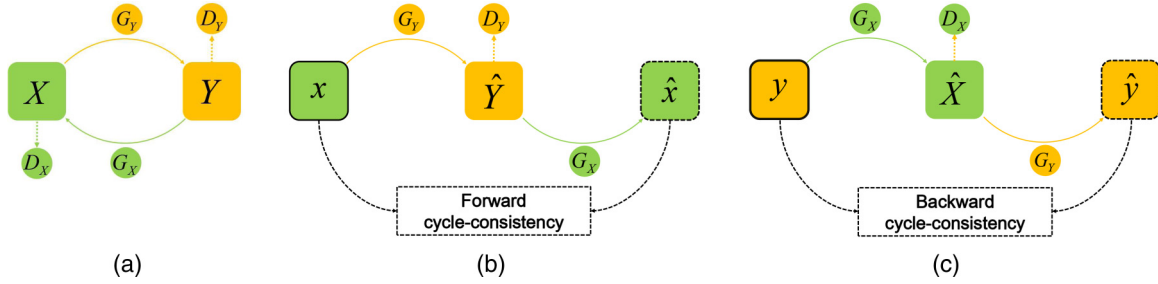
FIG. 3. The architecture of CycleGAN [61]. (a) CycleGAN comprises two generators ($G_X : Y \to X$ and $G_Y : X \to Y$) and two associated discriminators (D_X and D_Y). (b) Forward cycle consistency: $x \approx G_X(G_Y(x))$. (c) Backward cycle consistency: $y \approx G_Y(G_X(y))$.

II. SRCycleGAN: CycleGAN-BASED SR ALGORITHM FOR REAL-WORLD ROCK MCT IMAGES

A. Motivation

To address the problem of the mismatch between LR and HR training images (unpaired data), we propose to use CycleGAN [61] to learn the mapping between LR and HR images for the following reasons:

(1) Although it is difficult to obtain exactly matched LR-HR rock MCT images in a real scenario, collecting the images of the same sample at different resolutions is feasible. Therefore, unpaired LR-HR images of the same rock sample (or similar rock samples) can be obtained.

(2) CycleGAN is particularly suitable for translating the input from a source domain X to a target domain Y in the absence of paired input-output examples, making the distribution of the translated results from images in X indistinguishable from the distribution of images in Y.

We assume that LR rock images and the associated HR rock images belong to two related domains, i.e., different renderings of the same underlying scene. Correspondingly, our goal is to learn the underlying relationship between the LR domain and the HR domain, which can be achieved by using CycleGAN.

It should be noted that the goal of this work is to address the conflict between the FOV and the resolution of computed tomography (CT) images in real-world applications. To be more accurate, we aim to obtain an HR CT image with a large FOV from a LR observation using SR techniques. The absence of paired LR-HR examples for SR model training is one of the biggest challenges in a real scenario. To address this problem, this work proposes to consider the SR problem as the unpaired image-to-image translation. Because the target of this work is not to develop new networks, we adopt one of the most well-known algorithms for unpaired image-to-image translation, i.e., the CycleGAN [61], as the backbone of the proposed SR method.

B. An overview of CycleGAN

Figure 3 presents the schematic diagram of CycleGAN [61], which aims to translate an image in domain X to a target domain Y given one training set of images in X and another image set in Y. As shown in Fig. 3(a), CycleGAN consists of two generators (i.e., G_X and G_Y) and two adversarial discriminators (i.e., D_X and D_Y). G_X, G_Y, D_X, and D_Y are deep neural networks. Specifically, G_X models the mapping

$Y \to X$, such that the translated image $\hat{x} = G_X(y)$, $y \in Y$, is indistinguishable from images $x \in X$ by the associated discriminator D_X. D_X is trained to distinguish $\hat{x}$ (generated data in the X domain) from x (real data in the X domain). G_Y is the inverse of G_X; i.e., it learns the relationship $X \to Y$, making the output $\hat{y} = G_Y(x)$, $x \in X$, similar to images $y \in Y$ in the view of D_Y. D_Y aims to distinguish between $\hat{y}$ and y. The cycle-consistency property shown in Figs. 3(b) and 3(c) is one of the main characteristics of CycleGAN, which can guarantee the consistency between G_X and G_Y, i.e., $G_X(G_Y(x)) \approx x$ [forward cycle consistency, Fig. 3(b)] and $G_Y(G_X(y)) \approx y$ [backward cycle consistency, Fig. 3(c)]. Inspired by the impressive performance of CycleGAN on unpaired image-to-image translation, we take it as a backbone to develop the SR method for real-world rock MCT images in this work.

C. Proposed method

Figure 4 illustrates the framework of the proposed CycleGAN-based SR method for real-world rock MCT images, namely, SRCycleGAN, which consists of two stages: Off-line training (bottom) and on-line testing (top). G_Y aims to translate LR rock images from domain X into an image domain Y of higher resolution against D_Y that attempts to distinguish real and fake HR images. Meanwhile, G_X tries to produce LR rock images that belong to the LR domain X against D_X that differentiates fake LR images from real samples in X.

1. Off-line training

Without loss of generality, we denote LR and HR training samples as $\{x_i\}_{i=1}^{P}$ ($x_i \in X$ and $x_i \in \mathfrak{R}^{M \times N}$) and $\{y_j\}_{j=1}^{Q}$ ($y_j \in Y$ and $y_j \in \mathfrak{R}^{M \times N}$), respectively. It is important to note that, as shown in Fig. 4, the original LR training images are upsampled to the expected resolution before being fed into the network. More specifically, LR training samples $\{x_i\}_{i=1}^{P}$, which have the same resolution (i.e., $M \times N$) as HR training samples $\{y_j\}_{j=1}^{Q}$, are the upscaled results of the original LR images $\{x_i^{\mathrm{ori}}\}_{i=1}^{P}$ ($x_i^{\mathrm{ori}} \in \mathfrak{R}^{\frac{M}{s} \times \frac{N}{s}}$), where s denotes the upsampling factor. Although upsampling results $\{x_i\}_{i=1}^{P}$ ($x_i \in X$) have the same size as expected, we still call them LR training images for low quality.

Unpaired LR and HR rock images $\{x_i\}_{i=1}^{P}$, $\{y_j\}_{j=1}^{Q}$ are used to train the four networks G_X, G_Y, D_X, and D_Y in
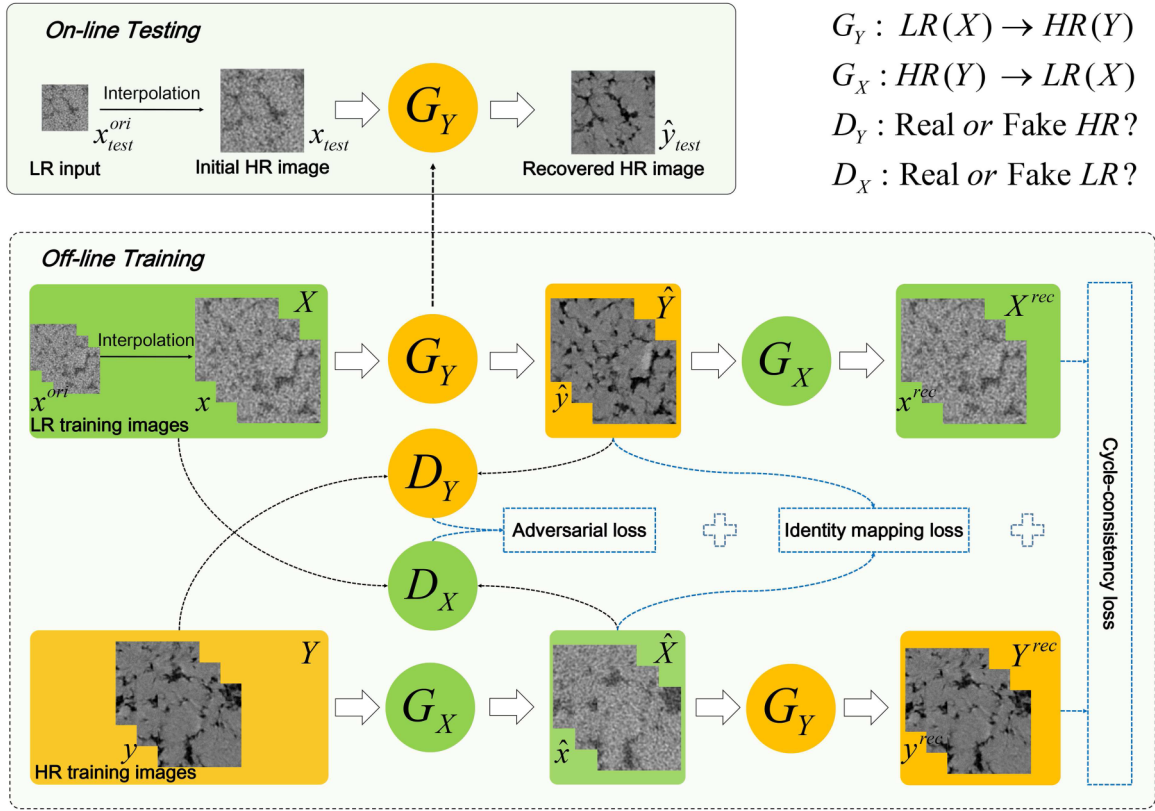
FIG. 4. The off-line training phase (bottom) and the on-line testing phase (top) of the proposed SRCycleGAN.

SRCycleGAN following the objective

$$\mathcal{L}(G_X, G_Y, D_X, D_Y, X, Y)$$
$$= \lambda_1 \mathcal{L}_{\text{gan}}(G_X, D_X, G_Y, D_Y, X, Y)$$
$$+ \lambda_2 \mathcal{L}_{\text{cyc}}(G_X, G_Y, X, Y)$$
$$+ \lambda_3 \mathcal{L}_{\text{ide}}(G_X, G_Y, X, Y), \tag{1}$$

where $\mathcal{L}_{\text{gan}}$, $\mathcal{L}_{\text{cyc}}$, and $\mathcal{L}_{\text{ide}}$ represent the adversarial loss, the cycle-consistency loss, and the identity mapping loss, respectively. λ_1, λ_2, and λ_3 are used to adjust the relative importance of the three terms on the right-hand side. Three loss functions are detailed in the following.

The adversarial loss $\mathcal{L}_{\text{gan}}$ makes the distribution of the translated data ($\hat{x}$ and $\hat{y}$) accord with the distribution of the data (x and y) in the target domain (X and Y). To be more specific, G_Y (G_X) tries to produce HR (LR) rock images that look like real HR (LR) images in domain Y (X). The generated HR (LR) results by generator G_Y (G_X) should be indistinguishable from real HR (LR) rock images for the corresponding discriminator D_Y (D_X). The objective is formulated as

$$\mathcal{L}_{\text{gan}}(G_X, D_X, G_Y, D_Y, X, Y)$$
$$= \mathcal{L}_{\text{gan}_X}(G_X, D_X, X, Y) + \mathcal{L}_{\text{gan}_Y}(G_Y, D_Y, Y, X), \tag{2}$$

where $\mathcal{L}_{\text{gan}_X}(G_X, D_X, X, Y) = \mathbb{E}_{x \sim p_{\text{data}}(x)}[\log D_X(x)] + \mathbb{E}_{y \sim p_{\text{data}}(y)}[\log(1 - D_X(G_X(y)))]$ and $\mathcal{L}_{\text{gan}_Y}(G_Y, D_Y, Y, X) = \mathbb{E}_{y \sim p_{\text{data}}(y)}[\log D_Y(y)] + \mathbb{E}_{x \sim p_{\text{data}}(x)}[\log(1 - D_Y(G_Y(x)))]$. $x \sim p_{\text{data}}(x)$ and $y \sim p_{\text{data}}(y)$ denote the distribution of LR and HR rock samples, respectively. Taking $\mathcal{L}_{\text{gan}_X}(G_X, D_X, X, Y)$ as an

example, G_X aims to produce a LR rock image $\hat{x} = G_X(y)$ that looks like images in the LR domain X, while D_X tries to make a distinction between the translated result $\hat{x}$ and real LR samples in X. Mathematically, G_X attempts to minimize $\mathcal{L}_{\text{gan}_X}(G_X, D_X, X, Y)$ against D_X that aims to maximize the same objective, i.e., $\min_{G_X} \max_{D_X} \mathcal{L}_{\text{gan}_X}(G_X, D_X, X, Y)$. The $\mathcal{L}_{\text{gan}_Y}(G_Y, D_Y, Y, X)$ for generator G_Y and its discriminator D_Y is similar to $\mathcal{L}_{\text{gan}_X}(G_X, D_X, X, Y)$.

The cycle-consistency loss $\mathcal{L}_{\text{cyc}}$ encourages the two generators (G_X and G_Y) to be inverse of each other. As presented in Fig. 4, for each LR rock image x in X, G_X should be able to bring the translated HR rock image $\hat{y} = G_Y(x)$ back to the original one, i.e., $x^{rec} = G_X(G_Y(x)) \approx x$. This constraint is called *forward cycle consistency*. Similarly, the *backward cycle consistency* is defined as $y^{rec} = G_Y(G_X(y)) \approx y$. The cycle consistency is achieved using the following loss:

$$\mathcal{L}_{\text{cyc}}(G_X, G_Y, X, Y) = \mathbb{E}_{x \sim p_{\text{data}}(x)}[\|G_X(G_Y(x)) - x\|_1]$$
$$+ \mathbb{E}_{y \sim p_{\text{data}}(y)}[\|G_Y(G_X(y)) - y\|_1]. \tag{3}$$

The identity mapping loss $\mathcal{L}_{\text{ide}}$ constrains the two generators (G_X and G_Y) to manifest as an identity mapping when the images from the target domain are used as the input. Specifically, G_X (G_Y) is expected to output the input when feeding LR (HR) rock images, i.e., $G_X(x) \approx x$ and $G_Y(y) \approx y$. This behavior is encouraged using the following
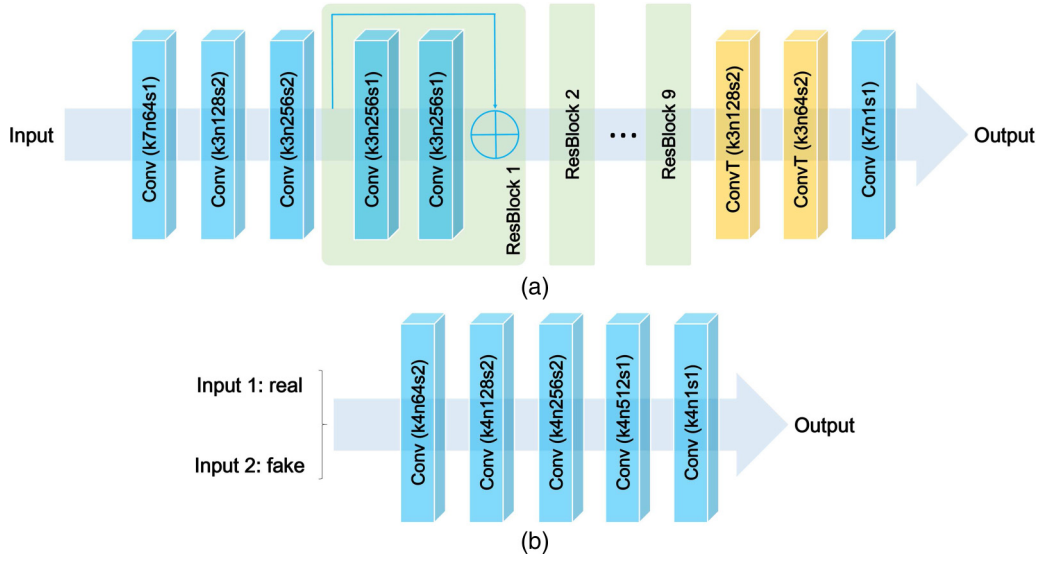
FIG. 5. Network architecture. (a) The architecture of generators (G_X and G_Y). (b) The architecture of discriminators (D_X and D_Y). Conv, convolutional layer; ConvT, deconvolutional layer. For each convolutional layer and deconvolutional layer, the numbers after symbols "k," "n," and "s" denote kernel size, number of kernels, and stride size, respectively. Note that nonlinear activation layers and normalization layers are omitted for simplicity.

loss:

$$\mathcal{L}_{\text{ide}}(G_X, G_Y, X, Y) = \mathbb{E}_{x \sim p_{\text{data}}(x)}[\|G_X(x) - x\|_1] \\ + \mathbb{E}_{y \sim p_{\text{data}}(y)}[\|G_Y(y) - y\|_1]. \quad (4)$$

Summarizing the above, the two generators (G_X and G_Y) and two discriminators (D_X and D_Y) can be optimized via solving

$$G_X{}^*, G_Y{}^*, D_X{}^*, D_Y{}^* \\ = \arg \min_{G_X, G_Y} \max_{D_X, D_Y} \mathcal{L}(G_X, G_Y, D_X, D_Y, X, Y). \quad (5)$$

Although four subnetworks in Eq. (5) are jointly optimized in the off-line training phase, only G_Y that can translate LR rock images into HR images is used at the on-line testing stage.

2. On-line testing

Given a LR rock image $x_{\text{test}}^{\text{ori}} \in \mathfrak{R}^{m \times n}$, as presented in Fig. 4, we first upsample it to the expected resolution ($ms \times ns$) using a simple interpolation method, like bicubic. The upsampling result is denoted as $x_{\text{test}} \in \mathfrak{R}^{ms \times ns}$. The learned generator G_Y is applied to x_{test} to produce an HR rock image $\hat{y}_{\text{test}} \in \mathfrak{R}^{ms \times ns}$ as

$$\hat{y}_{test} = G_Y(x_{test}). \quad (6)$$

D. Network architecture

Figure 5 shows the architectures of generators G_X, G_Y and discriminators D_X, D_Y. Note that only major components are presented, omitting nonlinear activation layers and normalization layers for simplicity. In Fig. 5, the convolutional layers and deconvolutional layers are expressed as "Conv" and "ConvT," respectively. For each convolutional layer and deconvolutional layer, the numbers after symbols "k," "n," and "s" represent kernel size, number of filters, and stride size, respectively. For example, "Conv (k3n128s2)" represents a

convolutional layer that contains 128 filters, and the filter size is 3×3 and the stride is set to 2.

The two generators G_X and G_Y in Fig. 4 share the same architecture, as shown in Fig. 5(a). The network illustrated in Fig. 5(a) consists of three parts. The first part is composed of three convolutional layers that gradually reduce feature size and increase the channel of features. And an activation layer is placed after each convolutional layer. The second part contains nine residual blocks, and each of then is composed of two convolutional layers and an activation layer in the middle. For the third part, two deconvolutional layers are adopted to increase feature size while decreasing the channel of features first, and then the last convolutional layer transforms features into the expected number of the output channel. The detailed settings of these convolutional and deconvolutional layers are indicated in the figure.

The two discriminators D_X and D_Y in Fig. 4 also share the same architecture, as presented in Fig. 5(b). The discriminator contains five convolutional layers, which are followed by an activation layer except for the last one. The first three layers increase the channel of features while reducing feature size, and the fourth layer further increases the number of feature channels to 512. The last convolutional layer compresses its input into one channel.

E. Datasets and training details

To more comprehensively test the ability of the proposed SR method for rock MCT images, SRCycleGAN is trained and tested on three datasets, i.e., *Sandstone_Real_1.8x*, *Sandstone_SiDe_4.0x*, and *Sandstone_SiUn_4.0x*.

a. Sandstone_Real_1.8x. This is a dataset we have built, which contains unpaired LR–HR rock MCT images of the same sandstone sample. HR rock images are imaged at a resolution of 5 μm and LR rock images are imaged at 9 μm. This dataset is composed of 900 LR images of size 1900 $\times$ 1900 and 512 HR images of size 520 $\times$ 520. Figure 6 presents
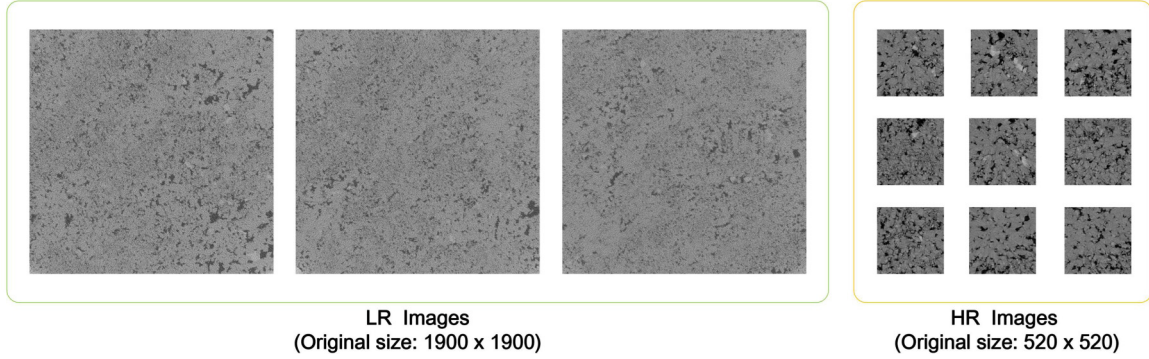
FIG. 6. Visualization of LR and HR images in the proposed dataset *Sandstone_Real_1.8x*. Please zoom in to view details and make comparisons.

some examples of this dataset. 2048 subimages (256×256) extracted from HR rock images are used as HR samples for network training. LR rock images are $1.8\times$ upsampled first to a resolution at 5 μm, and then 22 000 subimages (256×256) cropped from the interpolated LR images are used as LR samples for model training (20 000 images) and testing (2000 images). Note that there is no overlap between the LR samples for training and testing.

b. Sandstone_SiDe_4.0x. This dataset is provided by Wang *et al.* [62]. The size of 800 HR rock images in this dataset is 800×800. These HR images are $4\times$ downsampled to generate matched LR rock images using the imresize function in MATLAB, and the interpolation kernel is set to "bicubic." Hence, LR images are 200×200. In our experiments, 7200 HR samples (256×256) are extracted from HR rock images for network training. LR rock images are first $4\times$ upsampled to the size of 800×800, and 12 000 subimages (256×256) cropped from the interpolated LR images are used as LR samples for network training. It is important to note that although paired HR-LR rock images can be obtained for this dataset, we do not use paired data for network training. The extracted HR and LR training samples are not matched. Wang *et al.* [62] also provided a corresponding testing dataset, which contains 100 LR images (200×200) and their ground truth. We crop 9 subimages from each upscaled LR image (900 subimages in total) for testing, and 900 HR counterparts are cropped from the ground truth for performance evaluation. Similarly, there is no overlap between the LR samples for training and testing.

c. Sandstone_SiUn_4.0x. This dataset provided by Wang *et al.* [62] is almost the same as *Sandstone_SiDe_4.0x*, except a randomly selected interpolation kernel (including "box," "triangle," "cubic," "lanczos2," and "lanczos3") is applied to each HR image for downsampling. For this dataset, we use the same strategy as *Sandstone_SiUn_4.0x* to extract HR and LR training and testing samples.

We use PyTorch for training and testing on an NVIDIA GeForce GTX 1080Ti GPU. The Adam solver [63] is used to optimize our network SRCycleGAN with a batch size 1. All models are trained for 50 epochs from scratch with the initial learning rate 0.0002. The learning rate for the first 25 epochs is set to 0.0002 and it linearly decays to zero over the next 25 epochs. λ_1, λ_2, and λ_3 in Eq. (1) are set to 1, 10, and 5, respectively [64].

III. RESULTS AND DISCUSSION

A. Performance evaluation criteria

Objective and subjective comparisons are combined to evaluate the performance of the proposed SR approach SR-CycleGAN.

For the experiments on simulated LR rock images from *Sandstone_SiDe_4.0x* and *Sandstone_SiUn_4.0x*, in addition to visual effect comparisons, the PSNR is used to quantitatively assess the quality of the reconstructed HR rock images for the HR ground truth of each LR test image available. The PSNR is a full-reference image quality assessment index that measures the difference between two images by a comparison of pixel values [65]. PSNR is widely used for SR performance evaluation [9,11,33], and it is defined as

$$\text{PSNR} = 10\log_{10}\frac{255^2}{\text{MSE}},$$

$$\text{MSE} = \frac{1}{\text{HW}}\sum_{i=1}^{\text{H}}\sum_{j=1}^{\text{W}}[\text{I}_{\text{ref}}(i, j)-\text{I}(i, j)]^2, \quad (7)$$

where $\text{I} \in \mathfrak{R}^{\text{H}\times\text{W}}$ and $\text{I}_{\text{ref}} \in \mathfrak{R}^{\text{H}\times\text{W}}$ denote the recovered image and the corresponding ground truth, respectively. Generally, higher PSNR scores indicate better SR performance.

For the experiments on real-world rock MCT images, we first present the visual quality of the recovered HR images, and then quantitatively measure the accuracy and reliability of these results using a set of evaluation parameters.

Since there are no LR-HR image pairs in this test, the entropy is used to quantitatively evaluate the results produced by bicubic interpolation and our method SRCycleGAN. The entropy is a no-reference quality assessment index that measures the uncertainty, localization, and concentration [66]. The definition of the entropy is

$$\text{Entropy} = -\sum_{i=0}^{L-1} p_i\log_2 p_i, \quad (8)$$

where p_i represents the probability density function of the ith gray level, and L denotes the number of gray levels.

In addition, the porosity [67], the local porosity distribution [67,68], the two-point correlation function [2], the lineal-path function [69], the two-point cluster function [70], the chord-length distribution function [71], and the pore size distribution
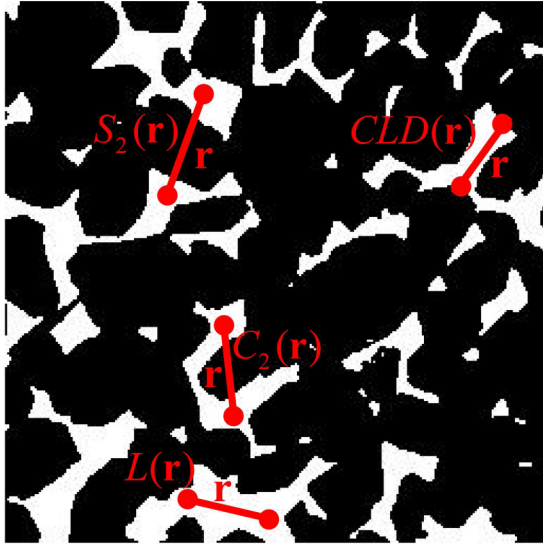
FIG. 7. Illustration of the two-point correlation function $S_2(r)$, the lineal-path function $L(r)$, the two-point cluster function $C_2(r)$, and the chord-length distribution function $CLD(r)$. The white part denotes the phase of interest and the black part is background.

[72] are employed to evaluate the consistency between the binarization results of the reconstructed HR rock images and the referred HR images (i.e., HR training images) from different perspectives. It should be noted that these parameters have been widely used to evaluate the performance of the algorithms for porous media reconstruction [3,4,47,59,73]. For two-dimensional images, the porosity refers to the ratio between the area of the pores and the sample area [67]. The local porosity distribution presents the probabilities of different local porosity values [67,68], and it can provide valuable information for measuring the physical properties (e.g., the permeability and diffusion quality) of porous media. Figure 7 illustrates the definitions of four morphological functions on a binary image. Given a line of length $r = \|\mathbf{r}\|$, the two-point correlation function $S_2(\mathbf{r})$ denotes the probability that the two endpoints of the line lie in the same phase [2]. Similarly, the two-point cluster function $C_2(\mathbf{r})$ represents the probability that the two endpoints are in the same cluster [70]. The lineal-path function $L(\mathbf{r})$ represents the probability that the line is located entirely within the same phase [69]. The chord-length distribution function $CLD(\mathbf{r})$ is an important index that measures the shape and size distribution of microscale constituents [71]. As shown in Fig. 7, a chord refers to a line segment whose midpoints are wholly in the phase of interest and endpoints are situated in the juncture of different phases. The pore size distribution PRD is also used as an index to measure the consistency between the recovered HR images and the ground truth [72]. In our experiments, the pores of different shapes are simplified in the shape of circles while calculating the pore radius. Mathematically, the pore radius r is defined as $\sqrt{A/\pi}$, where A denotes the area of a pore.

B. Experiments on simulated LR rock MCT images

We first conduct experiments on simulated LR rock MCT images for more intuitive and more accurate comparisons

between the reconstructed HR rock images and the corresponding ground truth. It should be noted that each LR testing image has an HR counterpart in this simulation experiment. Therefore, both qualitative and quantitative comparisons between the SR results and the real HR images can be conducted to evaluate the performance of the proposed SR method. As previously mentioned, 900 LR rock MCT images from *Sandstone_SiDe_4.0x* and 900 from *Sandstone_SiUn_4.0x* are used in this test, and the SR factor is set to 4. Figures 8 and 9 present some examples of the recovered HR rock images and the ground truth counterparts. Specifically, Fig. 8 shows the results on *Sandstone_SiDe_4.0x* and Fig. 9 shows the results on *Sandstone_SiUn_4.0x*. It can be observed from Figs. 8 and 9 that the resolution of the input (the first row) is so low that small structures are imperceptible. Although the results produced by bicubic interpolation (the second row) are larger than the corresponding inputs, they are blurred to some extent, especially in the areas of edges and textures. The HR rock images generated by the proposed SRCycleGAN (the third row), by contrast, not only have higher spatial resolution but also contain clearer edges and richer details. The results by our method are more visually appealing than the images produced by bicubic interpolation. And more importantly, it can be seen from Figs. 8 and 9 that the recovered HR images by the proposed SR method are very close to the corresponding targets. Even for the LR images with multiple components (e.g., different kinds of minerals), the SR results by SRCycleGAN are in good agreement with the original HR images. These results demonstrate the effectiveness and reliability of the proposed method.

Figure 10 presents the PSNR scores achieved by bicubic interpolation and the proposed SRCycleGAN, which are calculated between the recovered HR rock images and the corresponding ground truth. In general, higher PSNR values indicate better SR performance because the SR results are closer to the original HR images. The distributions of PSNR scores on *Sandstone_SiDe_4.0x* and *Sandstone_SiUn_4.0x* are illustrated in Figs. 10(a) and 10(b), respectively. Overall, the performance of the proposed SRCycleGAN is better than bicubic interpolation in terms of PSNR. The average PSNR scores achieved by the proposed method (26.34 and 26.14 dB) are higher than that of bicubic interpolation (25.95 and 25.74 dB) on both testing datasets. The PSNR gains achieved by SRCycleGAN quantitatively verify that the HR images produced by our approach are more reliable than the results by bicubic interpolation. Note that the average PSNR scores on *Sandstone_SiDe_4.0x* by bicubic interpolation and our method are about 0.2 dB higher than that on *Sandstone_SiUn_4.0x*. The main reason is that LR training and testing images are generated in different ways for the two datasets. For *Sandstone_SiDe_4.0x*, HR images are downsampled to generate LR counterparts using the "imresize" function, with the fixed interpolation kernel "bicubic." However, for *Sandstone_SiUn_4.0x*, a randomly selected interpolation kernel is applied to each HR image to generate a LR image. Clearly, the setting on *Sandstone_SiDe_4.0x* is relatively simpler, and thus better upsampling performance can be achieved. This phenomenon shows that the degradation model (i.e., the way to generate LR images) would affect SR performance.
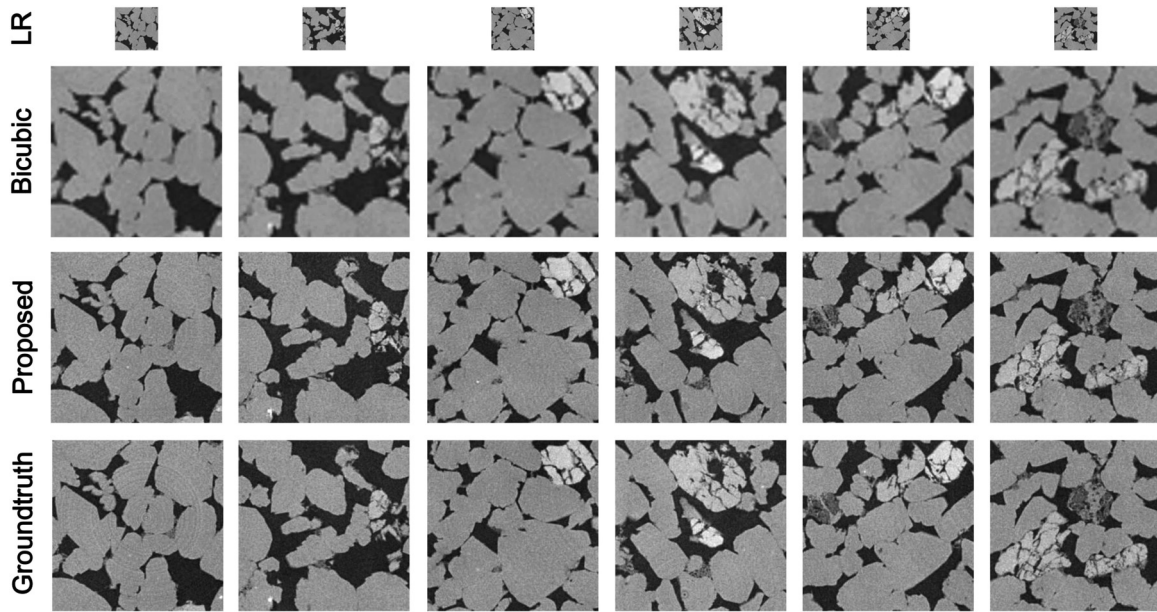
FIG. 8. Visual comparison for $4\times$ SR on the testing images in *Sandstone_SiDe_4.0x*. From top to bottom: LR rock images, the HR results produced by bicubic interpolation, the HR results produced by the proposed SRCycleGAN, and the ground truth. Please zoom in to view details and make comparisons.

The comparisons on visual quality and PSNR of the recovered HR images verify the effectiveness and superiority of the proposed SR framework for rock MCT images. It is important to highlight that unpaired samples are used for model training in this test, although we have LR-HR image pairs. That is, the one-to-one relationships between LR and HR training images are ignored to simulate a real-world scenario. The good agreement between the SR results by SRCycleGAN and the original HR images demonstrates that the proposed method can recover a reliable HR image from a LR observation, even without paired LR-HR images for model training.

C. Experiments on real-world LR rock MCT images

This section presents the experimental results on real-world rock MCT images. Figure 11 compares the perceptual quality of LR images (the first row), the recovered HR rock images by bicubic interpolation (the second row), and the proposed SRCycleGAN (the third row). In this test, there are no paired HR images corresponding to the LR testing images observed in a real-world scenario. Therefore, some HR training images (denoted as "Groundtruth" in Fig. 11) are shown in the fourth row for more intuitive comparisons. It is
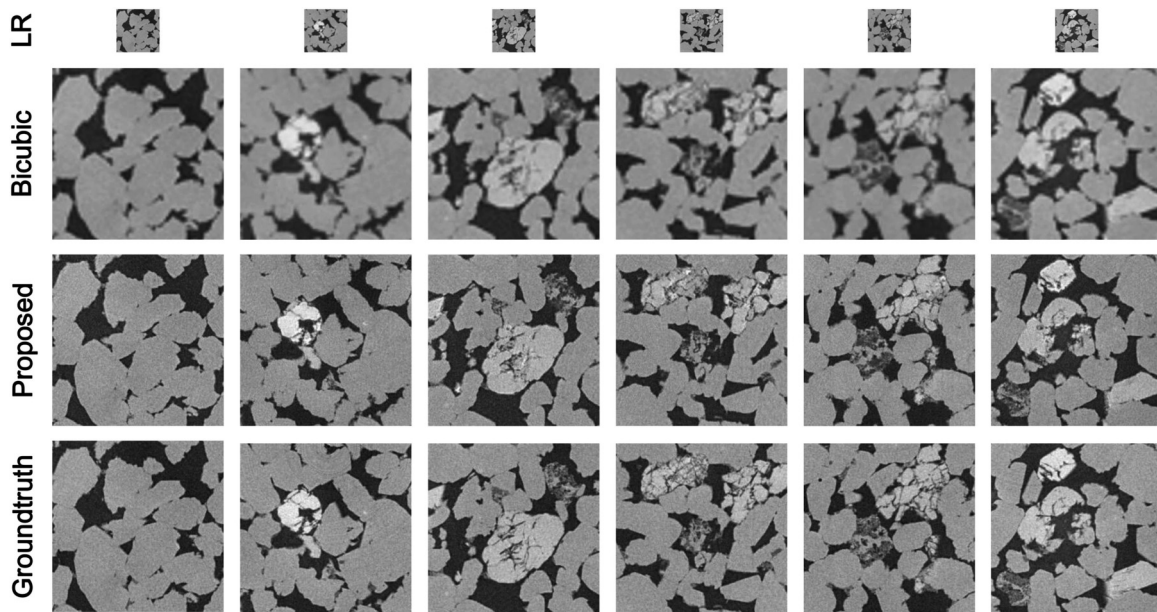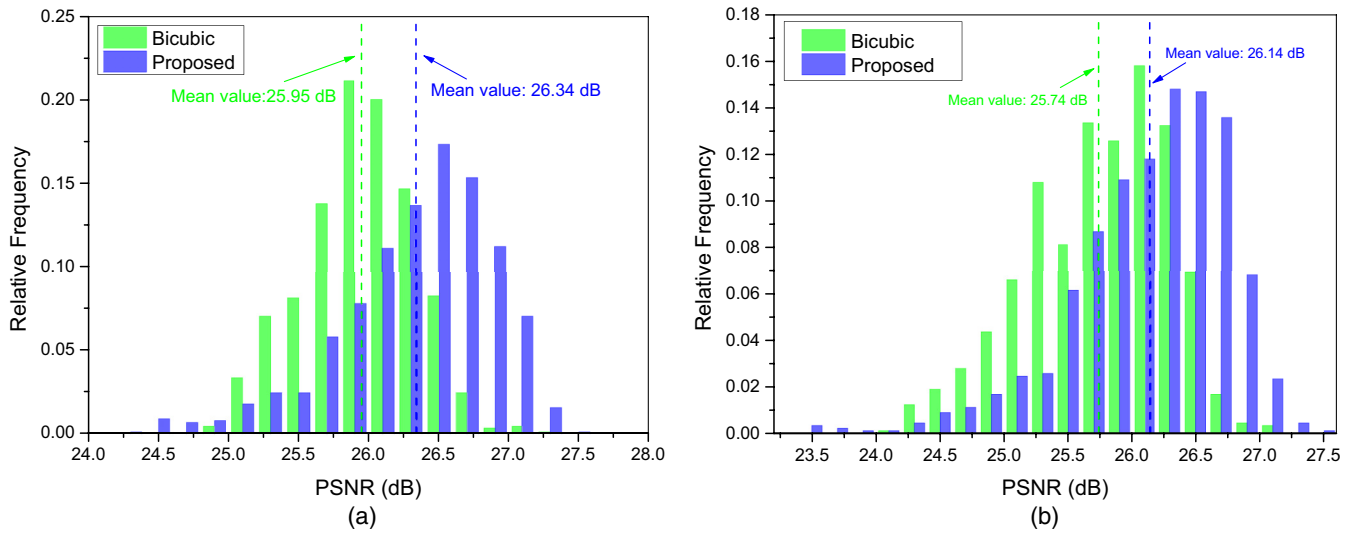

FIG. 9. Visual comparison for $4\times$ SR on the testing images in *Sandstone_SiUn_4.0x*. From top to bottom: LR rock images, the HR results produced by bicubic interpolation, the HR results produced by the proposed SRCycleGAN, and the ground truth. Please zoom in to view details and make comparisons.

FIG. 10. PSNR comparison for 4× SR on simulated LR images. (a) The PSNR scores achieved by bicubic interpolation and the proposed SRCycleGAN on 900 testing images in *Sandstone_SiDe_4.0x*. (b) The PSNR scores achieved by bicubic interpolation and the proposed SRCycleGAN on 900 testing images in *Sandstone_SiUn_4.0x*.

important to note that the HR images shown in the fourth row ("Groundtruth") do not match the images in other rows. In the following figures, we use similar expressions. It can be seen that LR images and the upsampling results by bicubic interpolation suffer from a variety of artifacts, resulting in the overlapping of the pore part and the stone part. For the HR images produced by the proposed framework SRCycleGAN, they are much clearer and lots of pores are well recovered. Meanwhile, the pore part and the stone part are distinguished better, which helps the subsequent processing and analysis.

In addition, although in this test we do not have the real HR counterparts of the input as the reference, one can observe that the recovered HR rock images by the proposed SRCycleGAN are very similar to HR training images, i.e., our targets. Specifically, it can be observed from the last column of Fig. 11 that there is a bright region in the HR image reconstructed by the proposed SRCycleGAN. As we understand it, there are two main reasons. First, the LR testing image in the last column contains a bright area, which can also be found in the SR result by bicubic interpolation. Second, bright areas
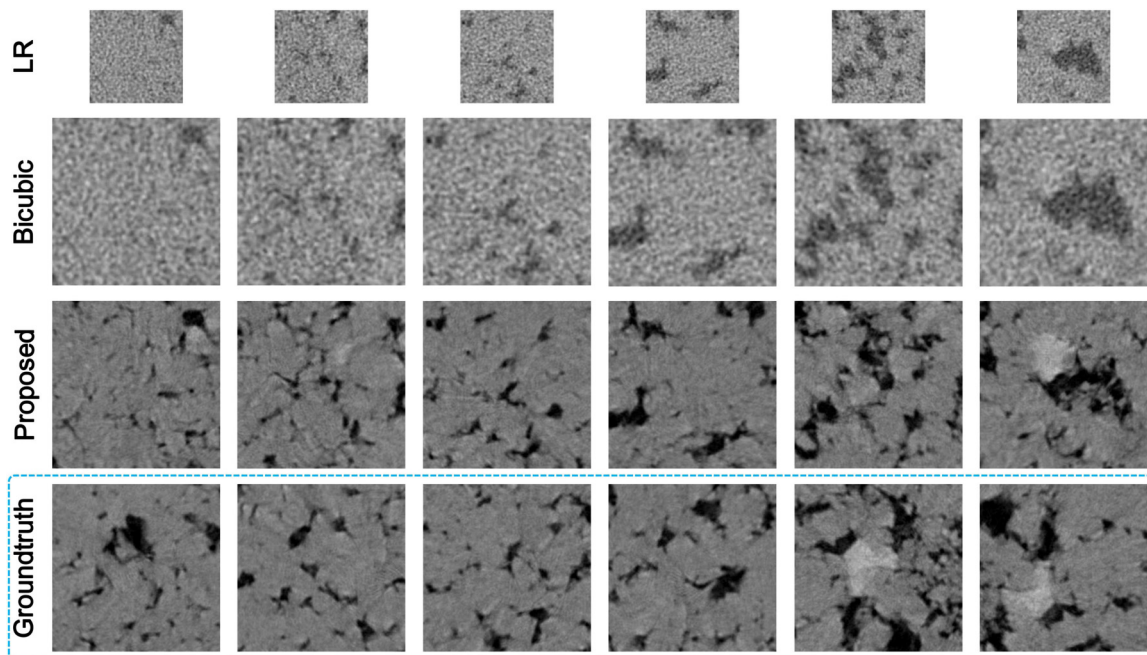

FIG. 11. Visual comparison for 1.8× SR on the testing images in *Sandstone_Real_1.8x*. From top to bottom: LR rock images, the HR results produced by bicubic interpolation, the HR results produced by the proposed SRCycleGAN, and the ground truth. Note that the ground truth in this figure refers to HR training images, which do not match LR images. In the following figures, we use similar expressions. Please zoom in to view details and make comparisons.
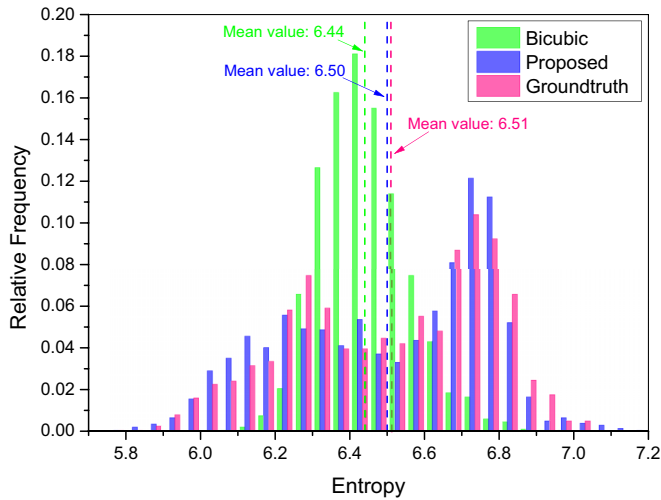
FIG. 12. Entropy comparison for $1.8\times$ SR on 2000 real-world LR images. The entropy scores are achieved by bicubic interpolation, the proposed SRCycleGAN, and the ground truth on the testing images in *Sandstone_Real_1.8x*.

appear in some LR and HR training images, as shown in the last line of Fig. 11. What is more, the bright regions in HR training images generally are brighter than those in LR testing and training images. The bright region in the result by SRCycleGAN is more visible because the HR image by SRCycleGAN is closer to the ground truth than the result by bicubic interpolation and the corresponding LR testing image.

The distribution of entropy values shown in Fig. 12 is used to perform quantitative evaluation due to the absence of paired LR-HR images in practical applications. It can be seen from Fig. 12 that the entropy distribution of the HR images recovered by bicubic interpolation is quite different from that of our targets (denoted as "Groundtruth"). By contrast, the entropy distribution achieved by the proposed method shows a good agreement with that of the targets.

To further evaluate the performance of the proposed SR algorithm for rock MCT images, a set of commonly used
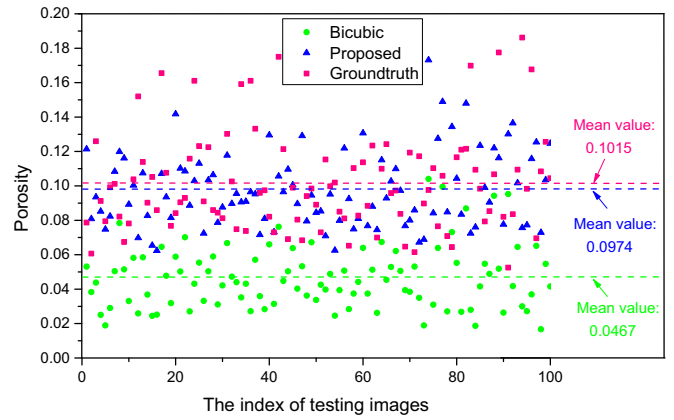

FIG. 14. Comparison of the porosity values achieved by bicubic interpolation, the proposed SRCycleGAN, and the ground truth.

statistical parameters (i.e., the porosity [67], the local porosity distribution *LPD* [67,68], the two-point correlation function S_2 [2], the lineal-path function *L* [69], the two-point cluster function C_2 [70], the chord-length distribution function *CLD* [71], and the pore size distribution *PRD* [72]) are compared among the bicubic interpolation, the proposed approach, and the targets.

To calculate these parameters, the targets (denoted as "Groundtruth") and the recovered HR images are binarized first. Some binary images are shown in Fig. 13. The binarized images of the ground truth in Fig. 13 are used as the reference. Obviously, compared with bicubic interpolation, our method distinguishes the pore part and the stone part better. Correspondingly, we can conclude that the results produced by the proposed SRCycleGAN are more conducive to the analysis of the microstructure and macroscopic properties of rock samples.

Figure 14 presents the porosity values achieved by different methods. It can be observed from Fig. 14 that the porosity values of the real HR rock CT images and that of the HR images reconstructed by SRCycleGAN are almost in the interval from 0.06 to 0.19, and these two sets of data have
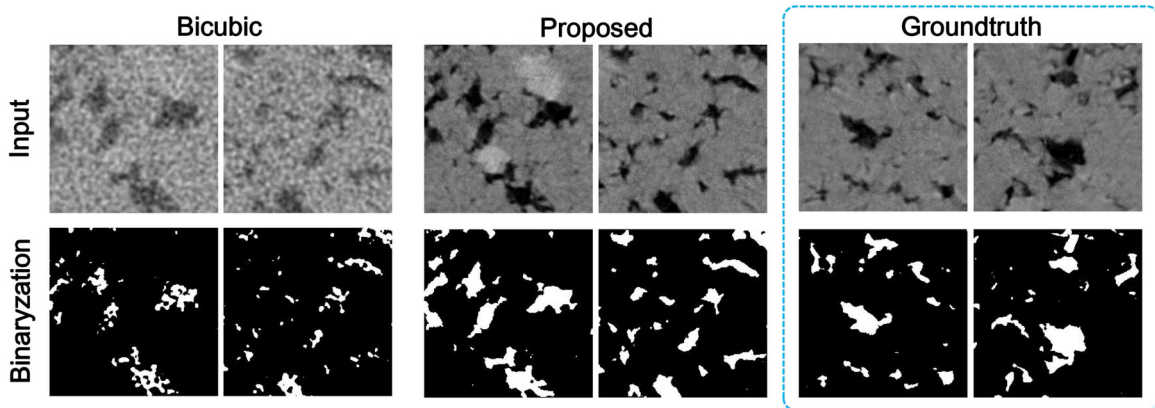

FIG. 13. Visual comparison for binarization on the original HR images and the recovered HR rock images by bicubic interpolation and the proposed SRCycleGAN. From left to right: The results produced by bicubic interpolation, the results produced by the proposed SRCycleGAN, and the ground truth. From top to bottom: The gray images and the corresponding binary images. The white part denotes the phase of interest and the black part is background. Please zoom in to view details and make comparisons.
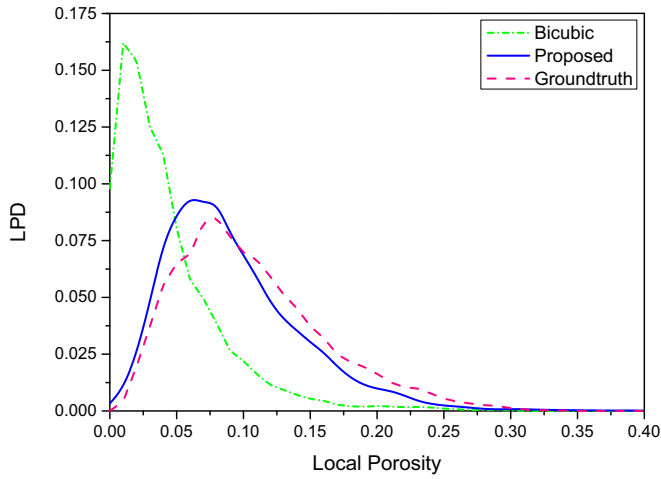
FIG. 15. Comparison of the averaged local porosity distributions *LPD* achieved by bicubic interpolation, the proposed SRCycleGAN, and the ground truth.

FIG. 17. Comparison of the averaged lineal-path functions *L* achieved by bicubic interpolation, the proposed SRCycleGAN, and the ground truth.

similar distributions. By contrast, the porosity values achieved by bicubic interpolation are markedly different. Specifically, the average porosity values of bicubic interpolation, the proposed SRCycleGAN, and the ground truth are 0.0467, 0.0974, and 0.1015, respectively. These results demonstrate that the HR images by SRCycleGAN are quite close to the targets in terms of the porosity. Actually, the images shown in Figs. 11 and 13 can also support this conclusion. The proposed SRCycleGAN not only improves the resolution, it also makes the pores more distinguishable from other phases. The local porosity distribution is considered a valuable index for analyzing the geometric properties and physical properties (e.g., permeability) of porous media. Therefore, Fig. 15 further compares the local porosity distribution. One can see that the local porosity distributions of SRCycleGAN and the ground truth display a good agreement, which manifests the superiority and reliability of the HR images reconstructed by the proposed method again.
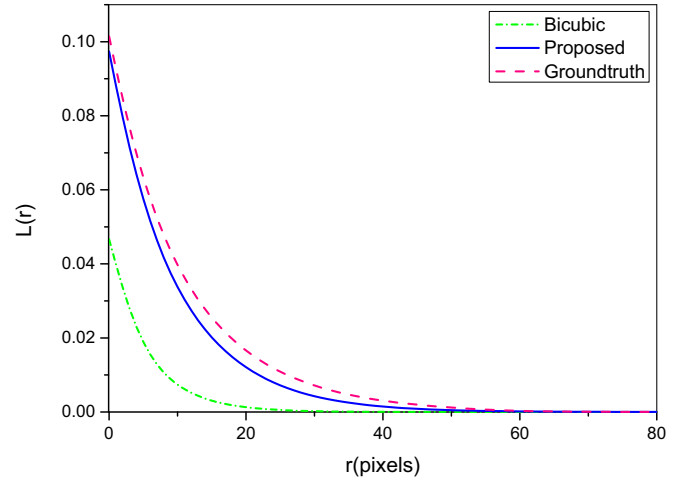
In order to validate the reliability of the proposed SR method from more comprehensive perspectives, Figs. 16–20 show the two-point correlation function S_2, the lineal-path function L, the two-point cluster function C_2, the chord-length distribution function *CLD*, and the pore size distribution *PRD*, respectively. It can be observed from Figs. 16–20 that there is a big gap between the statistical parameters of the images recovered by bicubic interpolation and the expected results (denoted as "Groundtruth"). The statistical parameters of the HR rock images generated by the proposed SRCycleGAN are in good agreement with that of the real HR observations, which shows the consistency between the recovered HR images and the targets on the morphology. These results also verify the accuracy and reliability of the HR images reconstructed by SRCycleGAN.

The above comparisons manifest that the HR images recovered by the proposed SRCycleGAN are close to the targets in terms of both morphological structures and statistical
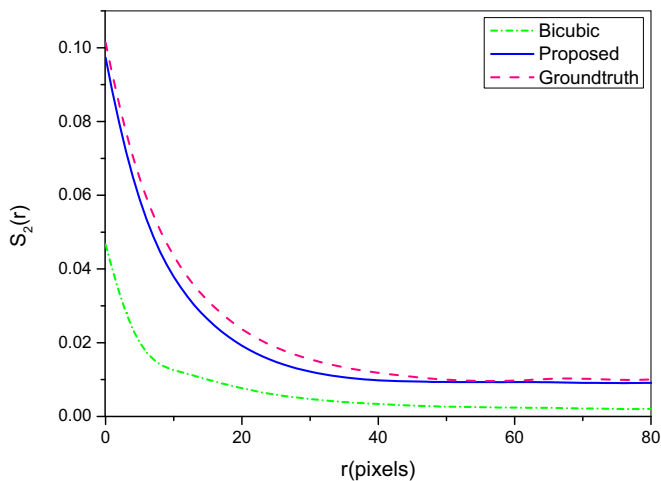

FIG. 16. Comparison of the averaged two-point correlation functions S_2 achieved by bicubic interpolation, the proposed SRCycle-GAN, and the ground truth.
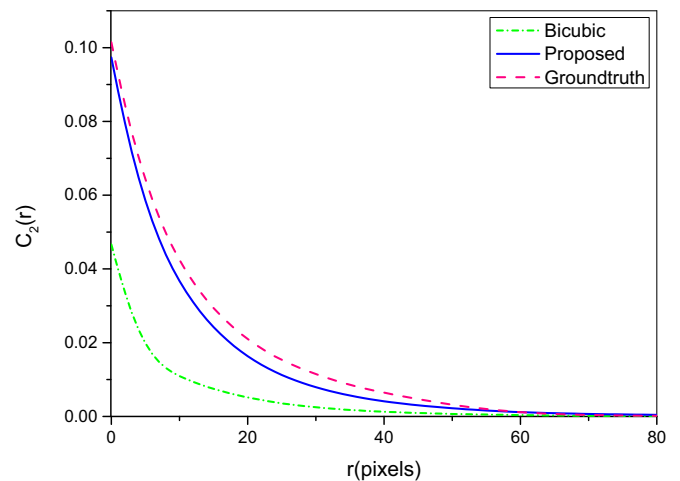

FIG. 18. Comparison of the averaged two-point cluster functions C_2 achieved by bicubic interpolation, the proposed SRCycleGAN, and the ground truth.
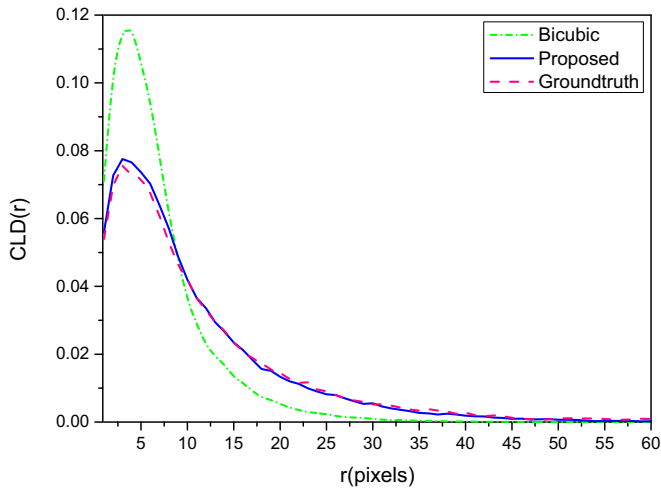
FIG. 19. Comparison of the averaged chord-length distributions *CLD* achieved by bicubic interpolation, the proposed SRCycleGAN, and the ground truth.

properties, which demonstrates that the proposed SRCycle-GAN performs well on the SR of real-world rock MCT images and the reconstructed HR images are reliable. Therefore, given a LR rock MCT image that has a large FOV, a reliable HR image with the same FOV can be produced by the proposed SRCycleGAN. Figure 21 presents an example. The LR rock MCT image shown in Fig. 21(a), which is captured in a real scenario, has a large FOV. However, its resolution is too low for subsequent processing and analysis. As presented in Fig. 21(b), the proposed SRCycleGAN greatly enhances the quality of Fig. 21(a) while keeping the same FOV. Compared with the LR rock MCT image shown in Fig. 21(a), the HR image Fig. 21(b) produced by SRCycleGAN shows the internal structure of the rock sample more clearly, which is helpful to analyze macroscopic properties.
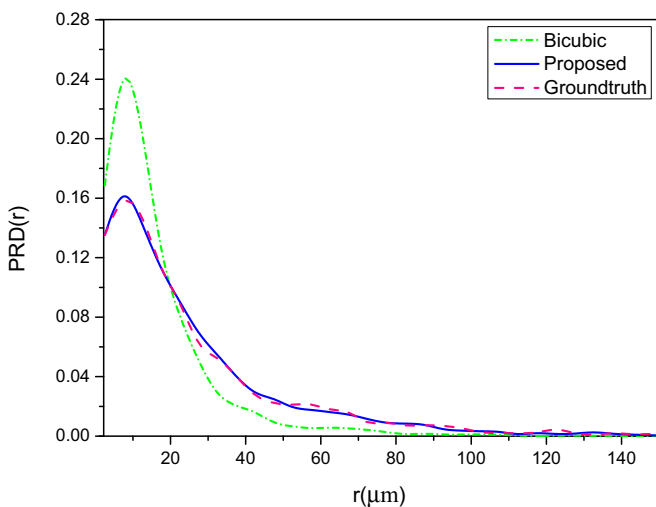

FIG. 20. Comparison of the averaged pore radius distributions *PRD* achieved by bicubic interpolation, the proposed SRCycleGAN, and the ground truth.
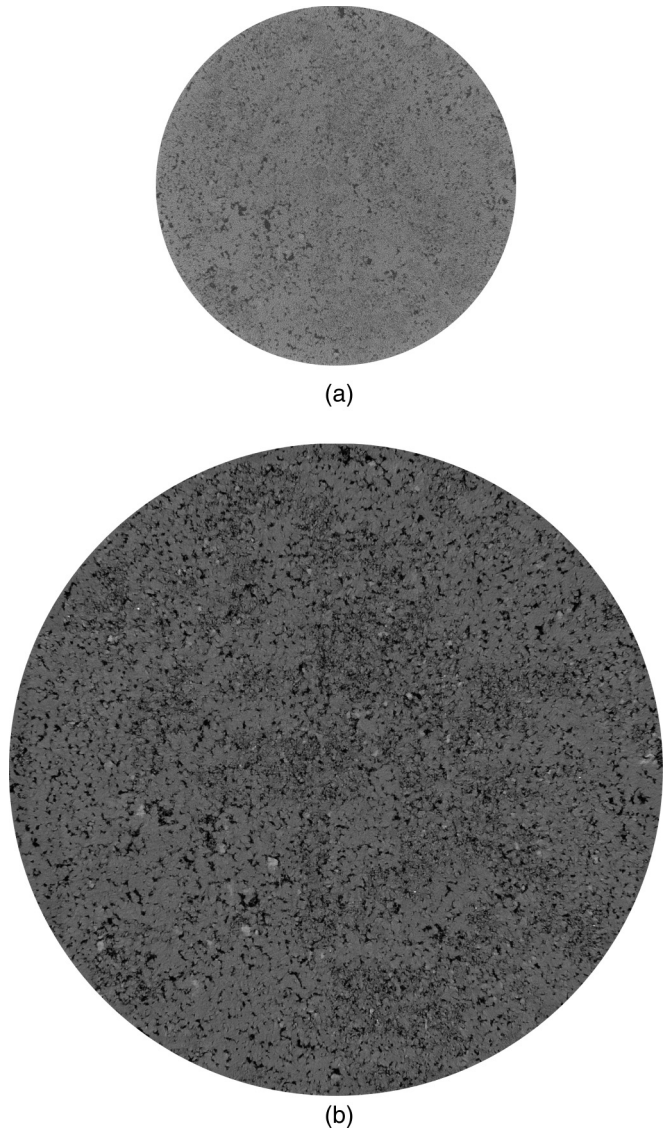

FIG. 21. Visual comparison for $1.8\times$ SR on a rock MCT image captured in a real scenario. (a) LR rock MCT image. (b) The HR rock MCT image produced by the proposed SRCycleGAN. Please zoom in to view details and make comparisons.

IV. CONCLUSION

In this work, we propose an effective SR method named SRCycleGAN for real-world rock MCT images, using cycle-consistent adversarial networks. To address the problem of the absence of paired LR-HR training samples, LR and HR rock samples are innovatively assumed to be located in two related domains, i.e., different renderings of the same underlying scene. Then, instead of learning the mapping between paired rock images, we learn the relationship between the LR domain and the HR domain without paired data. Experiments on simulated as well as real-world LR rock images demonstrate the effectiveness and superiority of the proposed SRCycleGAN, in terms of both qualitative and quantitative evaluation. The proposed SRCycleGAN can reconstruct a reliable HR rock image from the observed LR MCT image while keeping the same FOV. This work shows that, with the help of the

proposed SR algorithm, it is possible to obtain HR rock MCT images that exceed the limitations of imaging systems on FOV and resolution.

While the proposed SRCycleGAN achieves excellent SR performance on real-world MCT images, there still remain some problems for further study. First, this work focuses on the SR of two-dimensional rock MCT images, ignoring the information in the third dimension. It is interesting to investigate the extension of the proposed SRCycleGAN to 3D data, including running computational fluid dynamics simulations on the SR results and the original HR data. Second, the architectures of the generators and discriminators and the objective functions for model training will be studied further, specifically for the SR of real-world CT images. Finally, the applications of the proposed framework in other fields will

also be our future works, such as medical imaging and remote sensing.

[1] M. Sahimi, *Flow and Transport in Porous Media and Fractured Rock: From Classical Methods to Modern Approaches* (John Wiley & Sons, New York, 2011).

[2] S. Torquato, *Random Heterogeneous Materials: Microstructure and Macroscopic Properties* (Springer, Berlin, 2013), Vol. 16.

[3] K. Ding, Q. Teng, Z. Wang, X. He, and J. Feng, Phys. Rev. E **97**, 063304 (2018).

[4] Y. Li, Q. Teng, X. He, C. Ren, H. Chen, and J. Feng, Phys. Rev. E **99**, 062134 (2019).

[5] M. V. Karsanina, K. M. Gerke, E. B. Skvortsova, A. L. Ivanov, and D. Mallants, Geoderma **314**, 138 (2018).

[6] K. M. Gerke, M. V. Karsanina, T. O. Sizonenko, X. Miao, D. R. Gafurova, and D. V. Korost, in *SPE Russian Petroleum Technology Conference* (Society of Petroleum Engineers, Moscow, 2017).

[7] Y. Wang, S. S. Rahman, and C. H. Arns, Physica A **493**, 177 (2018).

[8] Y. Da Wang, R. T. Armstrong, and P. Mostaghimi, arXiv:1907.07131.

[9] Y. Da Wang, R. T. Armstrong, and P. Mostaghimi, J. Pet. Sci. Eng. **182**, 106261 (2019).

[10] Z. Li, Q. Teng, X. He, G. Yue, and Z. Wang, J. Appl. Geophys. **144**, 69 (2017).

[11] Y. Wang, Q. Teng, X. He, J. Feng, and T. Zhang, Comput. Geotech. **133**, 104314 (2019).

[12] K. Nasrollahi and T. B. Moeslund, Mach. Vis. Appl. **25**, 1423 (2014).

[13] W. Dong, L. Zhang, G. Shi, and X. Li, IEEE Trans. Image Process. **22**, 1620 (2013).

[14] H. Chen, X. He, Q. Teng, and C. Ren, Signal Process. Image Commun. **43**, 68 (2016).

[15] H. Chen, X. He, L. Qing, and Q. Teng, IEEE Trans. Multimedia **19**, 1702 (2017).

[16] K. Chang, P. L. K. Ding, and B. Li, IEEE Signal Process. Lett. **25**, 596 (2018).

[17] T. Li, X. He, L. Qing, Q. Teng, and H. Chen, IEEE Trans. Multimedia **20**, 1305 (2018).

[18] K. Chang, X. Zhang, P. L. K. Ding, and B. Li, Signal Process. **161**, 36 (2019).

[19] T. Li, X. Dong, and H. Chen, Neurocomputing **355**, 105 (2019).

[20] H. Chang, D.-Y. Yeung, and Y. Xiong, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (IEEE, New York, 2004), pp. 275–282.

[21] J. Yang, J. Wright, T. S. Huang, and Y. Ma, IEEE Trans. Image Process. **19**, 2861 (2010).

[22] R. Timofte, V. De Smet, and L. Van Gool, in *Proceedings of the Asian Conference on Computer Vision (ACCV)* (Springer, New York, 2014), pp. 111–126.

[23] H. Chen, X. He, L. Qing, Q. Teng, and C. Ren, Signal Process. Image Commun. **66**, 1 (2018).

[24] J.-J. Huang, T. Liu, P. Luigi Dragotti, and T. Stathaki, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)* (IEEE, New York, 2017), pp. 71–79.

[25] C. Dong, C. C. Loy, K. He, and X. Tang, IEEE Trans. Pattern Anal. Mach. Intell. **38**, 295 (2016).

[26] C. Dong, C. C. Loy, and X. Tang, in *Proceedings of the European Conference on Computer Vision (ECCV)* (Springer, Berlin, 2016), pp. 391–407.

[27] W. Shi, J. Caballero, F. Huszár, J. Totz, A. P. Aitken, R. Bishop, D. Rueckert, and Z. Wang, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (IEEE, New York, 2016), pp. 1874–1883.

[28] J. Kim, J. Kwon Lee, and K. Mu Lee, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (IEEE, New York, 2016), pp. 1646–1654.

[29] J. Kim, J. Kwon Lee, and K. Mu Lee, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (IEEE, New York, 2016), pp. 1637–1645.

[30] Y. Tai, J. Yang, and X. Liu, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (IEEE, New York, 2017), pp. 2790–2798.

[31] B. Lim, S. Son, H. Kim, S. Nah, and K. M. Lee, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)* (IEEE, New York, 2017), pp. 1132–1140.

[32] H. Chen, X. He, C. Ren, L. Qing, and Q. Teng, Neurocomputing **285**, 204 (2018).

[33] Y. Zhang, Y. Tian, Y. Kong, B. Zhong, and Y. Fu, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (IEEE, New York, 2018), pp. 2472–2481.

[34] Y. Zhang, K. Li, K. Li, L. Wang, B. Zhong, and Y. Fu, in *Proceedings of the European Conference on Computer Vision (ECCV)* (Springer, Munich, 2018), pp. 286–301.

[35] C. Ledig, L. Theis, F. Huszár, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang *et al.*, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (IEEE, New York, 2017), pp. 4681–4690.

[36] M. S. Sajjadi, B. Scholkopf, and M. Hirsch, in *Proceedings of the IEEE International Conference on Computer Vision (ICCV)* (IEEE, New York, 2017), pp. 4491–4500.

[37] R. Mechrez, I. Talmi, F. Shama, and L. Zelnik-Manor, in *Asian Conference on Computer Vision (ACCV)* (Springer, Berlin, 2018), pp. 427–443.

[38] S.-J. Park, H. Son, S. Cho, K.-S. Hong, and S. Lee, in *Proceedings of the European Conference on Computer Vision (ECCV)* (Springer, Munich, 2018), pp. 439–455.

[39] Y. Yuan, S. Liu, J. Zhang, Y. Zhang, C. Dong, and L. Lin, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops* (IEEE, New York, 2018), pp. 701–710.

[40] C. You, G. Li, Y. Zhang, X. Zhang, H. Shan, M. Li, S. Ju, Z. Zhao, Z. Zhang, W. Cong *et al.*, IEEE Trans. Med. Imaging **39**, 188 (2020).

[41] X. Liu, L. Chen, W. Wang, and J. Zhao, IEEE Trans. Image Process. **27**, 4971 (2018).

[42] F. Li, P. Ruiz, O. Cossairt, and A. K. Katsaggelos, in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (IEEE, New York, 2019), pp. 2327–2331.

[43] W. Wang, C. Ren, X. He, H. Chen, and L. Qing, IEEE Access **6**, 23767 (2018).

[44] Y. Jo, S. Wug Oh, J. Kang, and S. Joo Kim, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (IEEE, New York, 2018), pp. 3224–3232.

[45] M. S. Sajjadi, R. Vemulapalli, and M. Brown, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (IEEE, New York, 2018), pp. 6626–6634.

[46] N. Lubbers, T. Lookman, and K. Barros, Phys. Rev. E **96**, 052111 (2017).

[47] L. Mosser, O. Dubrule, and M. J. Blunt, Phys. Rev. E **96**, 043309 (2017).

[48] L. Mosser, O. Dubrule, and M. J. Blunt, arXiv:1802.05622.

[49] J. Feng, Q. Teng, X. He, and X. Wu, Acta Mater. **159**, 296 (2018).

[50] H. Chen, X. He, L. Qing, S. Xiong, and T. Q. Nguyen, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)* (IEEE, New York, 2018), pp. 711–720.

[51] G. Cheng, J. Han, P. Zhou, and D. Xu, IEEE Trans. Image Process. **28**, 265 (2018).

[52] X. Cao, F. Zhou, L. Xu, D. Meng, Z. Xu, and J. Paisley, IEEE Trans. Image Process. **27**, 2354 (2018).

[53] S. Karimpouli and P. Tahmasebi, Comput. Geosci. **126**, 142 (2019).

[54] S. Karimpouli, P. Tahmasebi, and E. H. Saenger, Nat. Resour. Res. 1 (2019), doi: 10.1007/s11053-019-09536-y.

[55] H. Chen, X. He, C. An, and T. Q. Nguyen, IEEE Signal Process. Lett. **26**, 79 (2019).

[56] S. Guo, Z. Yan, K. Zhang, W. Zuo, and L. Zhang, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (IEEE, New York, 2019), pp. 1712–1722.

[57] S. Karimpouli and P. Tahmasebi, Neural Networks **111**, 89 (2019).

[58] J. Zimmermann, B. Langbehn, R. Cucini, M. Di Fraia, P. Finetti, A. C. LaForge, T. Nishiyama, Y. Ovcharenko, P. Piseri, O. Plekan *et al.*, Phys. Rev. E **99**, 063309 (2019).

[59] J. Feng, X. He, Q. Teng, C. Ren, H. Chen, and Y. Li, Phys. Rev. E **100**, 033308 (2019).

[60] R. Timofte, E. Agustsson, L. Van Gool, M.-H. Yang, and L. Zhang, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)* (IEEE, New York, 2017), pp. 114–125.

[61] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, in *Proceedings of the IEEE International Conference on Computer Vision (ICCV)* (IEEE, New York, 2017), pp. 2223–2232.

[62] Y. Da Wang, P. Mostaghimi, and R. T. Armstrong, https://www.digitalrocksportal.org/projects/211.

[63] D. P. Kingma and J. Ba, in 3rd *International Conference for Learning Representations (ICLR 2015) San Diego, CA*.

[64] The source code of our SRCycleGAN is available at https://github.com/III-SCU/SRCycleGAN.

[65] A. Hore and D. Ziou, in *20th International Conference on Pattern Recognition (ICPR)* (IEEE, Washington, 2010), pp. 2366–2369.

[66] Z. H. Sun, in *IEEE International Conference on Image Processing (ICIP)* (IEEE, New York, 2017), pp. 3565–3569.

[67] R. Hilfer, Phys. Rev. B **45**, 7115 (1992).

[68] J. Hu and P. Stroeven, Cem. Concr. Res. **35**, 233 (2005).

[69] B. Lu and S. Torquato, Phys. Rev. A **45**, 922 (1992).

[70] S. Torquato, J. Beasley, and Y. Chiew, J. Chem. Phys. **88**, 6540 (1988).

[71] S. Torquato and B. Lu, Phys. Rev. E **47**, 2950 (1993).

[72] Y. Su, M. Zha, X. Ding, J. Qu, X. Wang, C. Yang, and S. Iglauer, Mar. Pet. Geol. **89**, 761 (2018).

[73] Y. Wu, P. Tahmasebi, C. Lin, L. Ren, and C. Dong, Mar. Pet. Geol. **109**, 9 (2019).

Correction: The ORCID identifier for the second author was presented incorrectly and has been fixed.