

Geometrical-Based Generative Adversarial Network to Enhance Digital Rock Image Quality

Yufu Niu,¹ Ying Da Wang^{✉,1} Peyman Mostaghimi,¹ James E. McClure,² Junqi Yin,³ and Ryan T. Armstrong^{1,*}

¹*School of Minerals and Energy Resources Engineering, The University of New South Wales, Sydney, NSW 2052, Australia*

²*Advanced Research Computing, Virginia Tech, Blacksburg, Virginia 24060, USA*

³*National Center for Computational Sciences, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831, USA*



(Received 14 February 2021; accepted 10 May 2021; published 14 June 2021)

X-ray microcomputed tomography (micro-CT) is a common tool for the study of porous media structures and properties. High-quality micro-CT data are required to accurately capture pore structures. Acquiring high-quality micro-CT data, however, is not always possible, owing to application limitations and experimental constraints. Therefore, we propose a geometrical-based generative adversarial network (GAN) to rapidly restore noisy micro-CT images to their clean counterparts. The training data and related ground-truth (GT) data are scanned for 7 min and 9.5 h, respectively. To evaluate the performance of the geometrical-based GAN, a 600³ voxel image that has never been used for training is reconstructed and compared with the corresponding GT image. Histogram matching and linear normalization are implemented to adjust the histogram of the reconstructed image to that of the GT image. A watershed-based segmentation method is then applied to delineate pore and solid phases. Lastly, we measure the Minkowski functionals and petrophysical properties, including absolute permeability, pore size distribution, drainage capillary pressure-saturation curve, and imbibition relative permeability, to estimate the physical accuracy of the denoised image. The results show that the proposed geometrical-based GAN can accurately restore noisy micro-CT data. By reducing the scanning time from 9.5 h to 7 min, the expenditure of collecting micro-CT can be decreased significantly. This is particularly important for applications where time-lapse images of a dynamic process are required, high-throughput imaging is necessary for real-time data analysis, or where the quantification of large sample volumes is required.

DOI: 10.1103/PhysRevApplied.15.064033

I. INTRODUCTION

Digital rock analysis has become an important technology in various research fields related to subsurface engineering applications [1,2]. Generally, three-dimensional (3D) digital images of porous rocks are obtained using x-ray microcomputed tomography (micro-CT), which are then filtered and segmented into pore space and solid volumes for direct numerical flow simulations [3–5], such as the lattice Boltzmann method [6,7] or pore network modeling [8–10]. The petrophysical properties predicted by numerical simulation depend on the micro-CT data quality; thus, the initial processing steps of filtering and segmentation are critical. Various filters have emerged as common standards; however, these filters are known to influence the physical properties of segmented images. In most cases, higher quality data collection is advised over image filtering for noise removal. However, it might not always

be practical to collect high-quality micro-CT data, such as when fast scans are required to collect time-lapsed data or when high throughput is required for real-time analysis. To address this issue, we introduce a generative adversarial network (GAN) with a geometrical-based loss function to restore digital rock images with significant noise.

Image noise relates to undesired signals, including arbitrary variation in color information or brightness caused by image acquisition and transmission [11,12]. In the past few decades, researchers have developed various image-denoising methods, such as spatial-domain filtering, transform-domain filtering, and statistical models. Spatial-domain-filtering methods include local filters, such as Gaussian [13] and Weiner filters [14], and nonlocal filters like the nonlocal means filter [15]. Unlike the spatial-domain approach, transform-domain-filtering methods contain the curvelet [16], discrete cosine [17], and wavelet domain transforms [18]. Other algorithms commonly used to reduce image noises include statistical models [19,20] or hybrid methods [21,22]. These methods,

*ryan.armstrong@unsw.edu.au

however, are usually limited to specific applications and have inherent side effects. For example, a spatial-domain filter may result in blurred data [23], and the transform-domain method may introduce ringing artifacts around the edges [24]. Therefore, it is crucial to develop advanced techniques for image denoising.

For micro-CT imaging, the quality of micro-CT data mainly depends on image contrast, spatial resolution, image noise, and artifacts [25]. In particular, the amount of noise in micro-CT data is directly dependent on the scanning time. This is highly relevant when 3D scans are required within a limited time constraint. With advances in micro-CT hardware and reconstruction algorithms, it is currently possible to collect 3D images at scan times of less than 1 min [26,27]. In addition, fast synchrotron-based micro-CT has been widely applied to achieve four-dimensional dynamic imaging of fluid flow in porous media on a timescale of tens of seconds [28,29]. Dynamic micro-CT has played an important role in understanding multiphase flow in porous geological materials for enhanced oil recovery and CO₂ sequestration [30]. Other studies with subminute scan times include those by Myers *et al.* [31], Youssef *et al.* [32], and Bultreys *et al.* [30]. Even though the current methods are promising for characterizing dynamics in 3D, there remain limitations in the degree to which the data can be quantified, the accuracy of quantification, and the timescales that can be achieved, which are directly associated with the amount of noise generated. Therefore, developing a physically accurate image-denoising technique is a critical step towards obtaining insights into fluid transport and other related transport processes in porous media.

Recently, deep learning has been widely applied to achieve state-of-the-art performance in image denoising, such as convolutional neural networks [33,34], autoencoders [35,36], and GANs [37–39]. However, there is only a limited amount of work on denoising micro-CT data related to porous media research. The denoising of micro-CT data with deep learning was briefly explored in previous superresolution studies [40,41] as an added feature achieved by synthetic augmentation, but this did not incorporate real noise from low-quality images directly obtained from the detector. In other studies with superresolution images [42], noise reduction was not analyzed within the context of reduced scan times and low x-ray dosage. Liu *et al.* [39] proposed a network called TomoGAN to reduce noise from low-dose micro-CT data. They demonstrated that TomoGAN achieved good performance in image denoising for both simulated data generated by the ASTRA toolbox and tomopy, and real experimental micro-CT data. However, the actual physical accuracy of the generated images were not evaluated, which is an important outcome when data are intended for geometrical measures and related bulk physical properties.

In this study, we propose a geometrical-based GAN (TomoGAN GVGG) to denoise very low quality micro-CT data, which traditional filtering methods, such as nonlocal means or anisotropic diffusion, cannot deal with (which is demonstrated in this study). A fine tuning process is implemented by adapting the VGG19 deep convolutional neural network into a geometrical-based network (GVGG) that aids in the training of TomoGAN. Overall, the trained network (TomoGAN GVGG) restores micro-CT data to a denoised state and maintains physical accuracy, in terms of common geometrical measures and bulk properties.

II. METHODOLOGY

GAN is a robust algorithm that was proposed by Goodfellow *et al.* [44] for unsupervised learning tasks. The GAN consists of two networks: a generator (G) and a discriminator (D). The generator creates a solution that is similar to the real data or ground truth (GT), while the discriminator aims to distinguish between fake and real data. This particular structure can form a two-player min-max competition from which the training process finds an optimum condition when both the generator and discriminator cannot be improved. At this point, the discriminator cannot discriminate between real and fake data [44]. The two-player min-max function between G and D is shown as follows:

$$\min_{\theta_G} \max_{\theta_D} V(D, G) = \mathbb{E}_{x \sim p_{\text{data}}(x)} [\log D(x)] + \mathbb{E}_{z \sim p_z(z)} (\log \{1 - D[G(z)]\}), \quad (1)$$

where θ_G and θ_D are the training weights and bias of G and D , respectively, and $p_{\text{data}}(x)$ and $p_z(z)$ are the corresponding probability distributions of real data and fake data from the generator, respectively. \mathbb{E} is the related expectation. We expect to obtain the optimized object $V(D, G)$, where the generator probability distribution [$p_z(z)$] matches the real probability distribution [$p_{\text{data}}(x)$].

A. Network formulation

We propose a Wasserstein GAN with a geometrical-based loss function (TomoGAN GVGG) to denoise micro-CT data. The GAN component follows previous works [39,45] and the geometrical-based loss function improves image quality based on the geometrical measures of volume, surface area, curvature, and Euler characteristic, which are collectively known as the Minkowski functionals (MFs). The networks are shown in Fig. 1 and are explained in the following paragraphs.

Figure 1 provides the structure of the generator and discriminator used for the TomoGAN GVGG. The generator is a modification of the U-Net architecture introduced by Ronneberger *et al.* [46]. The discriminator consists of

a standard number of convolutional layers that are condensed into a single probability output. The total loss for the generator is

$$L_G = \lambda_1 L_{\text{adv}} + \lambda_2 L_{\text{MSE}} + \lambda_3 L_{\text{GVGG}}, \quad (2)$$

where L_{adv} is the adversarial loss; L_{MSE} is the pixel-wise mean square error (MSE); L_{GVGG} is a geometrical-based VGG (GVGG) loss; and λ_1 , λ_2 , and λ_3 are set to 10, 0.5, and 2, respectively, as the weights of each loss function [39].

The adversarial loss is denoted as follows:

$$L_{\text{adv}} = - \sum_{i=1}^m D[G(I_i^{\text{input}})], \quad (3)$$

where m is the batch size, and I_i^{input} is the i th input noisy image from the batch size. Perceptual loss is the pixel-wise MSE loss [47]. However, solutions using the MSE loss often do not represent the spatial frequencies in images that bring about higher-level differences, such as structure and texture [48].

An approach to capture the higher-level features of an image is possible by using a VGG network as a perceptual

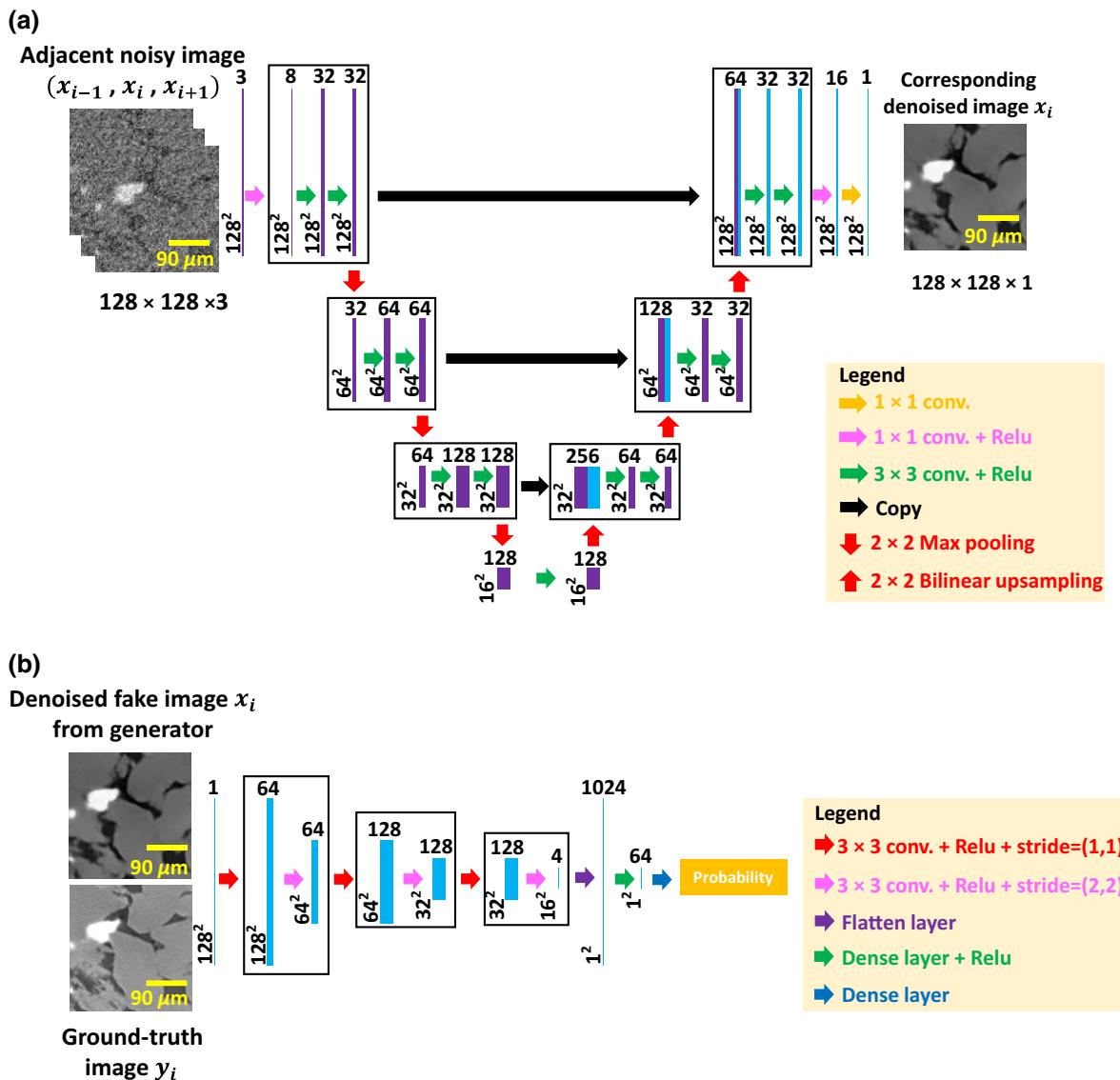


FIG. 1. (a) Generator structure of TomoGAN GVGG. Input is a three-channel adjacent gray-scale image (x_{i-1}, x_i, x_{i+1}) . Output is a one-channel gray-scale image corresponding to the middle input image (x_i) . (b) Discriminator structure of TomoGAN GVGG, inputs are the denoised fake image (x_i) and the related ground-truth image (y_i) . Output is a probability representing the discriminator performance.

loss [39,48,49]. A shortcoming is that current VGG networks are not trained on micro-CT images of rocks nor are they trained to capture the geometrical features of rocks that are important for digital rock applications. In general, the standard VGG19 loss is designed to compare features extracted from a pretrained VGG19 network [50]. Instead of the standard VGG19 loss, we introduce a geometrical-based VGG loss for digital rock images as the perceptual loss, which is defined by calculating the Euclidean distance between the geometrical-based feature representations of a denoised image, $V_{\theta_{\text{VGG}}}[G_{\theta_{\text{VGG}}}(I^{\text{input}})]$, from the generator and the related ground-truth image, $V_{\theta_{\text{VGG}}}(I^{\text{GT}})$, as follows:

$$L_{\text{GVGG}} = \sum_{i=1}^{W_f} \sum_{j=1}^{H_f} \{V_{\theta_{\text{VGG}}}(I^{\text{GT}})_{i,j} - V_{\theta_{\text{VGG}}}[G_{\theta_G}(I^{\text{input}})]_{i,j}\}^2, \quad (4)$$

where W_f and H_f represent the dimensions of the feature maps from VGG19, and $V_{\theta_{\text{VGG}}}$ is the extracted feature representation by fine-tuning the geometrical-based VGG19 network to identify geometrical information. Further details of this loss function are presented by Jiang and Li [51].

In addition, to optimize the discriminator loss, we add an additional gradient penalty proposed by Gulrajani *et al.* [52] based on the regular discriminator for stable training.

$$\begin{aligned} L_D = & \frac{1}{m} \sum_{i=1}^m \{D[G(I_i^{\text{input}})] - D(I_i^{\text{GT}})\} \\ & + \lambda_D \frac{1}{m} \{[||\nabla_{\alpha} D(\alpha)||_2 - 1]^2\}, \end{aligned} \quad (5)$$

where λ_D is set to ten for the gradient penalty, $\alpha = \epsilon G(I^{\text{input}}) + (1 - \epsilon) I^{\text{GT}}$, and ϵ is a random number between zero and one.

B. Data processing and training

The training and testing images are a series of three adjacent micro-CT image slices, as shown in Fig. 1(a), which provides 3D data to the network. The GT images are obtained by micro-CT imaging of Bentheimer sandstone with a relatively high number of projections, with many image accumulations per projected angle. The noisy images are collected by reducing the number of projections and image accumulations (number of radiographs for each angle). The specific imaging setting details are provided in

Table I; high-quality data take about 9.5 h to collect, while noisy data take 7 min. The resulting images are shown in Fig. 2.

We initially convert the 16-bit images into 8-bit images and normalize them by

$$\bar{p} = \frac{p - p_{\min}}{p_{\max} - p_{\min}}, \quad (6)$$

where p is the pixel value, and min and max denote the minimum and maximum pixel values after removing the extrema, respectively. In addition, the sandstone images are of a cylindrical core, and we add padding of the rock pattern to the edge regions to fill the background. The padding details can be found in the Supplemental Material [53]. Without the padding of the edge region, the network produces blurred textures in the edge region since these edge regions are not included in the training.

To create training and testing data sets, we divide each sample into two subvolumes and extract the image patches for training and testing from the corresponding subvolumes. The training data are shuffled for TomoGAN training. In this way, the testing data have the same probability distribution as the training data with different ROIs compared with the training data. In total, 40 000 image patches are cropped from the images and 5000 image patches for testing. The i th image patch is $128 \times 128 \times 3$ voxels by stacking the images of $i - 1$, i , and $i + 1$. The learning rate is set to 10^{-4} and then halved every 10 epochs to improve the capacity for learning complex patterns [54]. The Adam optimizer [55] is applied for generator and discriminator optimization. For each epoch, the generator is updated once while the discriminator is updated twice. The batch size is set to 32, and the number of training epochs is 200. All training is conducted on a single graphics processing unit (GPU), NVIDIA GeForce RTX 2080Ti, and the code is developed using TensorFlow.

After training, the TomoGAN GVGG network, including the model parameters and structures, is saved in a hierarchical data format. The image-reconstruction process can be implemented by reloading the trained model. Notably, the reconstructed images tested in this study are not from the training region, but have the same probability distribution as that of the training data. Using an NVIDIA GeForce RTX 2080Ti, we find that TomoGAN GVGG can reconstruct a 600×600 pixel image in 0.3 s.

TABLE I. Scan details of the training and testing micro-CT data.

Sample	Scan duration	Exposure time (s)	No. of projections	Accumulations	Resolution (μm)
GT	9 h 27 m	0.55	2880	20	2.15
Noising input	7 m	0.55	360	1	2.15

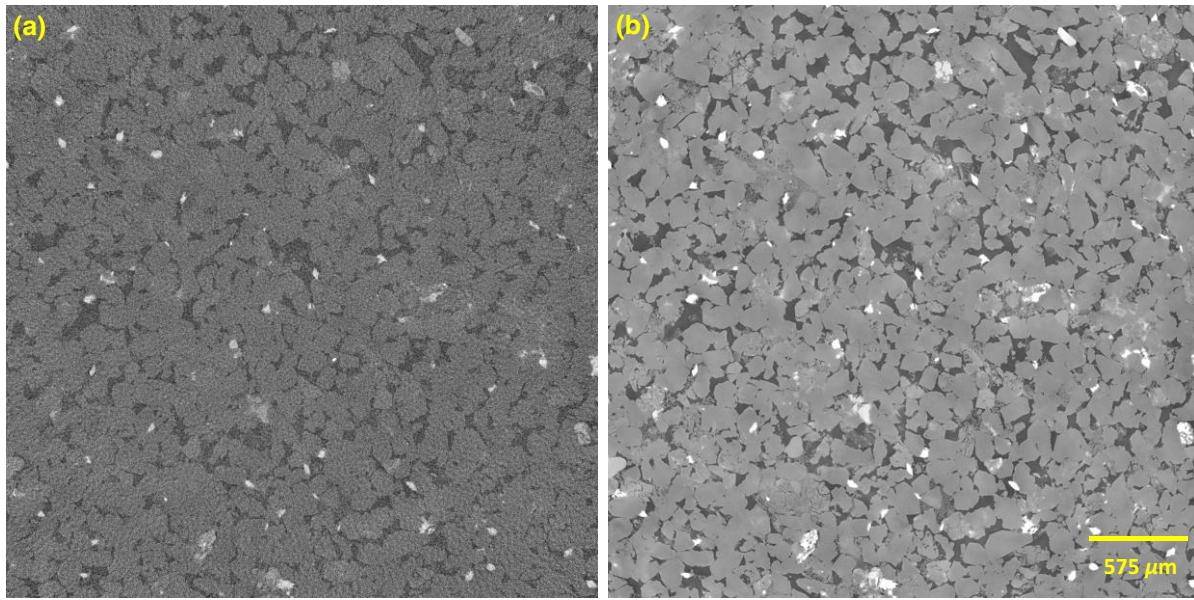


FIG. 2. Example of micro-CT data. (a) Noisy input data, (b) corresponding GT data.

C. Geometrical-based VGG

Deep convolutional neural networks (DCNNs), such as ResNet, GoogleNet, and VGG, are trained based on large public image repositories, such as ImageNet [56]. Pretrained DCNN models can usually be customized to effectively solve other given tasks. However, for digital rock data, the focus is on the pore structure and other geometrical features. These exclusive textures and patterns are different from those available in ImageNet. Therefore, directly applying pretrained DCNN models may cause issues in digital rock data. To learn the features of digital rock images, we implement a geometrical-based VGG that is specific for rock images, instead of using the standard VGG trained by ImageNet as the perceptual loss. This is achieved by fine-tuning the standard VGG19, which is a process in which the parameters of a model are trained for a given task and then adjusted to fit another task. The fine-tuning process is usually implemented on the top few layers of a network, so that the model parameters on the other layers are unchanged. An advantage of fine-tuning is that it significantly decreases the time required for training because it avoids training a new model from scratch. In addition, as a pretrained model is trained from a larger data set, such as ImageNet, the lower-layer features of the pretrained model usually perform well at capturing generic image features, such as edges [57]. Therefore, generic image features are captured more accurately, while fine-tuning adjusts the network to a specific task [58].

For fine-tuning, Bentheimer (resolution of 4.83 μm), Berea (resolution of 4.67 μm), and Leopard (resolution of 4.67 μm) sandstones are imaged at similar resolutions. The MFs [59,60] are used as the target objective for training.

The VGG is a pretrained two-dimensional (2D) network; therefore, the fine-tuning process is implemented in 2D. This means that the MFs are measured in the 2D space of the sandstone data sets. Indeed, a 3D VGG network could be designed and trained, but the benefit of having a pretrained network would be lost.

The MFs are a complete set of geometric measures that are commonly applied to quantify the structure of porous media [59–62]. Considering phase X embedded in Euclidean space Ω with surface δX , the MFs are $d + 1$ functionals, where d is the Euclidean space dimension. In 2D space, the geometric quantities of the MFs can be denoted as area, perimeter, and 2D Euler characteristic. The 2D MFs for the segmented images are calculated based on approximations of the Minkowski measures using the filtering and look-up table transformations proposed by Legland *et al.* [63].

The VGG19 network is assigned a regression task to learn the 2D MFs from gray-scale micro-CT images of the sandstones. The overall procedure can be explained in the following steps.

(a) Step 1, a watershed-based algorithm [64] is used to segment the micro-CT images into the pore space and solid phases.

(b) Step 2, both gray-scale and corresponding segmented images are cropped to image patches of 224×224 pixels (Fig. 3).

(c) Step 3, the 2D MFs are measured for each 2D image patch of the segmented image.

(d) Step 4, the gray-scale image patches and related MFs are applied to fine-tune the pretrained VGG19 network.

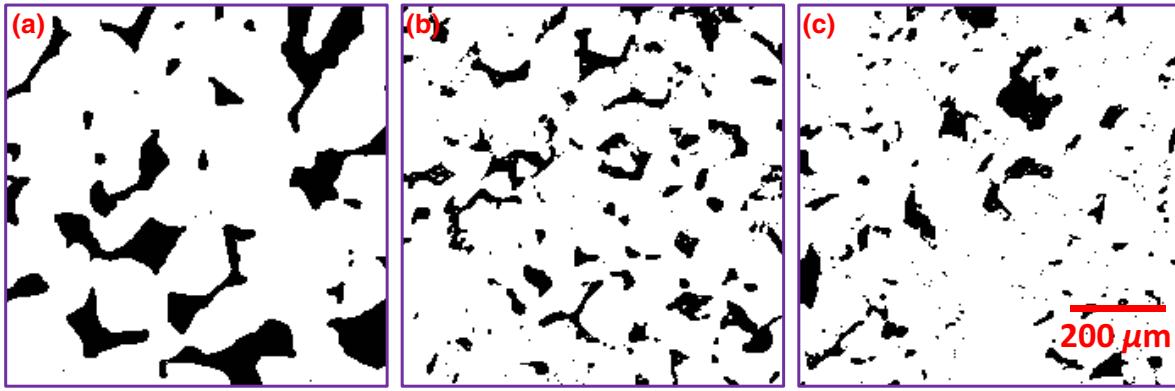


FIG. 3. 2D 224×224 -pixel segmented images of three sandstones for fine-tuning. (a) Bentheimer sandstone, (b) Berea sandstone, (c) Leopard sandstone. White is grain and black is void.

The training result is saved as $V_{\theta_{\text{VGG}}}$, which is applied to update the perceptual loss (L_{VGG}). The overall procedure is shown in Fig. 4(a).

In total, 27 650 image patches are created for training and 6910 image patches for testing. To avoid overfitting, the training epoch is set to 30, the batch size is 64, and the learning rate is 2×10^{-4} with Adam optimization. Figure 4(b) shows the fine-tuning of the VGG19 architecture. The red dotted box shows the trainable weights for fine-tuning. We unfreeze only the weights of the top few layers, which represent more specific features in VGG19.

III. DATA POSTPROCESSING

After training the geometrical-based VGG perceptual loss, we select a cubic region of 600^3 voxels that is not used for training to evaluate the network performance. Even though we add a perceptual loss to minimize the total loss between the output and GT, the perceptual loss results in an inconsistent reconstruction of image brightness [52,65]. Consequently, we implement a postprocessing procedure to deal with the reconstructed images. Figures 5 and 6 provide the workflow details and example images for the postprocessing steps. The related procedures are explained in the following steps.

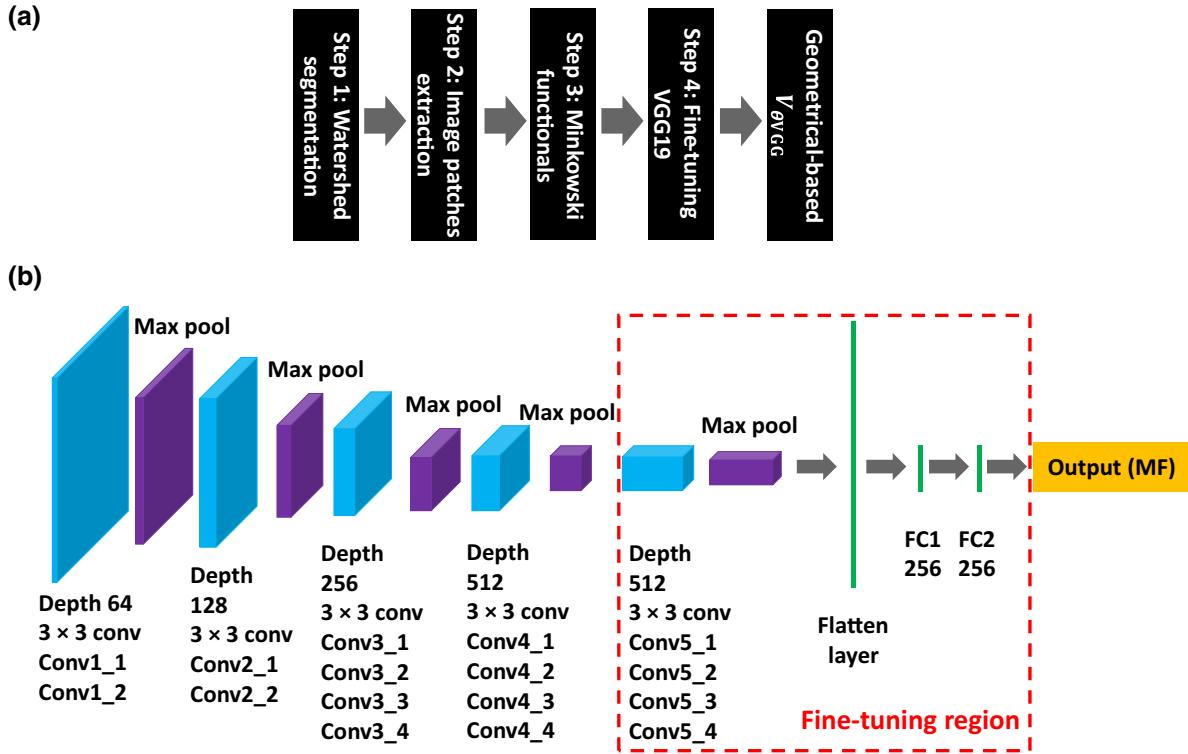


FIG. 4. (a) Geometrical-based VGG19 fine-tuning process for digital rock data, (b) fine-tuning VGG19 architecture.

(a) Step 1, use the trained TomoGAN GVGG to reconstruct a 600^3 voxel region of noisy data [Fig. 6(a)] that is not used for training. The reconstructed data are shown in Fig. 6(b), which we refer to as the initial TomoGAN GVGG.

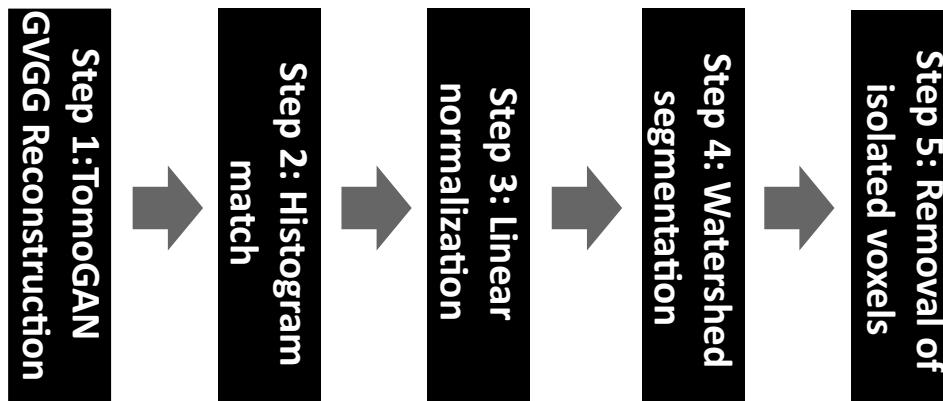
(b) Step 2, use histogram matching to provide consistent brightness for the TomoGAN GVGG data. We regard the first slice of TomoGAN GVGG as the reference slice and apply a histogram match to the remaining 599 slices.

(c) Step 3, use linear normalization on the TomoGAN GVGG result and GT, we calculate the average gray-scale values of each phase: pore space, clay, low-density minerals, and high-density minerals. Then, we build a linear relationship between the TomoGAN GVGG and GT data to obtain the final TomoGAN GVGG data. The linear relationship can be found in the Supplemental Material [53]. The gray-scale values of the final TomoGAN GVGG [Fig. 6(d)] are equivalent to those of the GT data [Fig. 6(c)].

(d) Step 4, use watershed-based segmentation to process both the final TomoGAN GVGG and GT data using the exact same parameter setting. The effective pore structures are then extracted by considering a connectivity algorithm (26 connectivity, faces, edges, or corners touching).

(e) Step 5, remove unrealistic isolated clusters of 30 voxels or less caused by image-segmentation errors. The resulting images are shown in Figs. 6(e)–6(f).

Here, we expand on Step 5. The segmentation process often causes isolated objects that can affect the pore connectivity. The Euler characteristic can be applied to evaluate the connectivity of a 3D binary image. We therefore measure the pore Euler characteristic versus the removal of objects below a given size; this is provided in the Supplemental Material [53]. When the removed object size is 30 voxels or less, the pore Euler characteristic becomes stable. Finally, we remove all objects in the TomoGAN GVGG and GT data sets that have fewer than 30 voxels.



Once the final data are prepared, as displayed in Figs. 6(e) and 6(f), the resulting geometrical measures and petrophysical properties are assessed. The geometrical measures are the MFs measured in 3D space. In addition to the previously explained MFs in 2D space, the MFs in 3D space are denoted as the volume M_0 , integral of the surface M_1 , integral of mean curvature M_2 , and Euler characteristic instead of M_3 .

The first functional, M_0 , is the volume of X . The second functional, M_1 , is the integral measure of the surface area of X , which is defined as

$$M_1(X) = \int_{\delta X} ds, \quad (7)$$

where ds is the surface element of X . The third functional is the integral of the mean curvature of the surface (δX), which is defined as

$$M_2(X) = \int_{\delta X} \left[\frac{1}{r_1} + \frac{1}{r_2} \right] ds, \quad (8)$$

where r_1 and r_2 are the principal radii of curvature of the surface element ds . The last functional is the integral of the Gaussian curvature of the surface (δX), which is defined as

$$M_3(X) = \int_{\delta X} \left[\frac{1}{r_1 r_2} \right] ds = 2\pi\chi(\delta X), \quad (9)$$

where $\chi(\delta X)$ is the Euler characteristic of the bounding surface [66]. The 3D Euler characteristic can also be defined as

$$\chi(X) = I - L + O, \quad (10)$$

where I is the number of isolated objects, L is the number of redundant loops, and O is the number of cavities [60], which provides a more intuitive understanding of the measure. A more negative Euler characteristic indicates more connectivity or loops. The 3D MFs are measured based on the algorithms proposed by Legland *et al.* [63].

FIG. 5. Data postprocessing workflow.

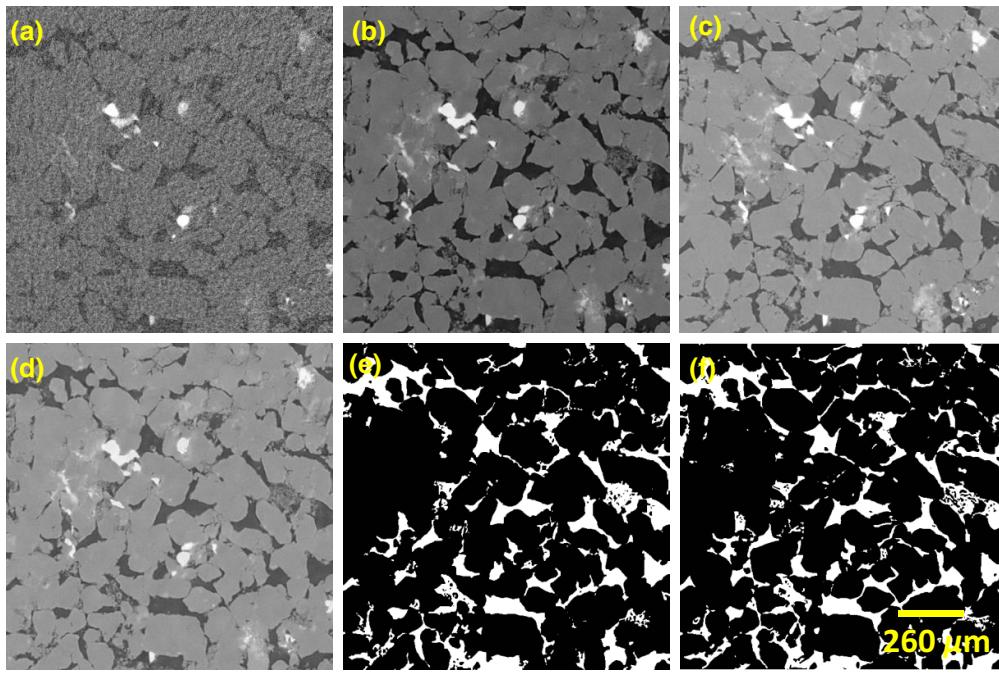


FIG. 6. Related 2D example images with post-processing. (a) input noisy image, (b) initial TomoGAN GVGG result, (c) ground truth, (d) final TomoGAN GVGG result, (e) final TomoGAN GVGG segmentation after removing objects with fewer than 30 voxels, (f) GT segmentation after removing objects with fewer than 30 voxels.

We also estimate the petrophysical properties of the images, such as absolute permeability, pore size distribution, drainage capillary pressure-saturation curve, and imbibition relative permeability. The absolute permeability is calculated using the single-phase lattice Boltzmann method [67,68]. The pore size distribution is estimated based on the local distance maximum method [68–70], which is similar to the watershed-based transform [71] or the maximum inscribed sphere method [72]. Capillary pressure-saturation curves are calculated on the segmented images using a quasistatic morphological approach based on the local maximum distance transform, as explained by Wang *et al.* [68]. By taking the inverse of the largest pore-filled radius (proportional to the Young-Laplace equation) at given invasion profiles, the normalized drainage-capillary-pressure curve can be calculated over the invasion-saturation profile. The imbibition relative permeability is calculated on the segmented images using direct two-phase flow simulation by the morphological lattice Boltzmann method (MorphLBM) [67,68]. Each image is morphologically initialized, and the domain is mirrored in the principal flow direction. Periodic boundary conditions and a body-force term are applied to mimic coinjection, and the fluids have equal density and viscosity. The system interfacial tension and body force are set to simulate flow at a capillary number of 10^{-5} , and saturation steps are made by shell aggregation, as specified in the MorphLBM method. These simulations are performed on the Gadi supercomputer (National Computational Infrastructure, supported by the Australian Government) using four Nvidia V100 GPUs. The simulations took five million time steps within 48 h of wall time.

IV. RESULTS AND DISCUSSION

A. Geometrical-based VGG by fine-tuning

First, we investigate the fine-tuning strategy for obtaining a geometrical-based VGG for digital rock data. Figure 7(a) shows the MSE for geometrical-based VGG training and testing. The training and testing MSE converge from the eighth epoch and become stable thereafter. To examine the model performance, we compare the predicted 2D MFs results [Figs. 7(b)–7(d)] with the GT data for the testing data sets. The 2D MFs results shown in Figs. 7(b)–7(d) illustrate that the geometrical-based VGG can achieve accurate area and perimeter predictions with R^2 values of 0.98 and 0.94, respectively. Although the Euler characteristic is more intractable to estimate immediately, its R^2 still reaches a value of 0.86. Overall, the proposed geometrical-based VGG can accurately predict the geometrical properties of digital rock images, owing to the adaptability of the fine-tuning strategy.

B. Noise removal

By introducing the geometrical-based VGG19 loss as a perceptual loss, we find that the reconstructed image quality can be improved. Figures 8(a)–8(c) shows the image results for the standard VGG19 (TomoGAN SVGG) and geometrical-based VGG19 (final TomoGAN GVGG) compared with the GT image. ROI images are also provided where finer clay structure details can be compared. It can be seen that TomoGAN SVGG produces blurred results [Fig. 8(d)], which do not recover the finer-detailed clay minerals that are evident in the GT image [Fig. 8(f)]. In contrast, TomoGAN GVGG captures finer

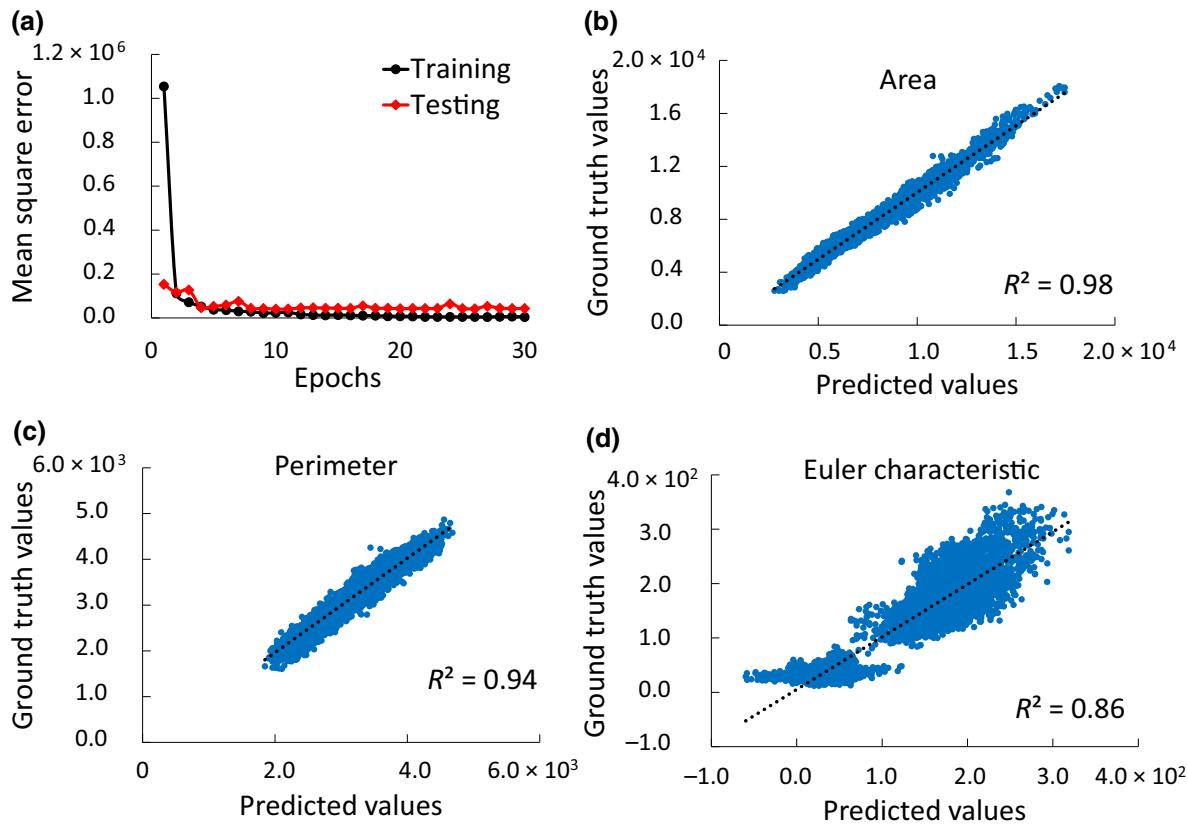


FIG. 7. Fine-tuning results for geometrical-based VGG. (a) Training and testing loss curve, (b) cross plot of predicted area and GT, (c) cross plot of predicted perimeter and GT, (d) cross plot of predicted Euler characteristic and GT.

details [Fig. 8(e)], which appear similar to the GT image [Fig. 8(f)]. A baseline comparison for noise removal is also provided in the Supplemental Material [53], whereby

a nonlocal means filter or anisotropic diffusion filter are applied to the noisy data. These standard filters are not able to restore the noisy data.

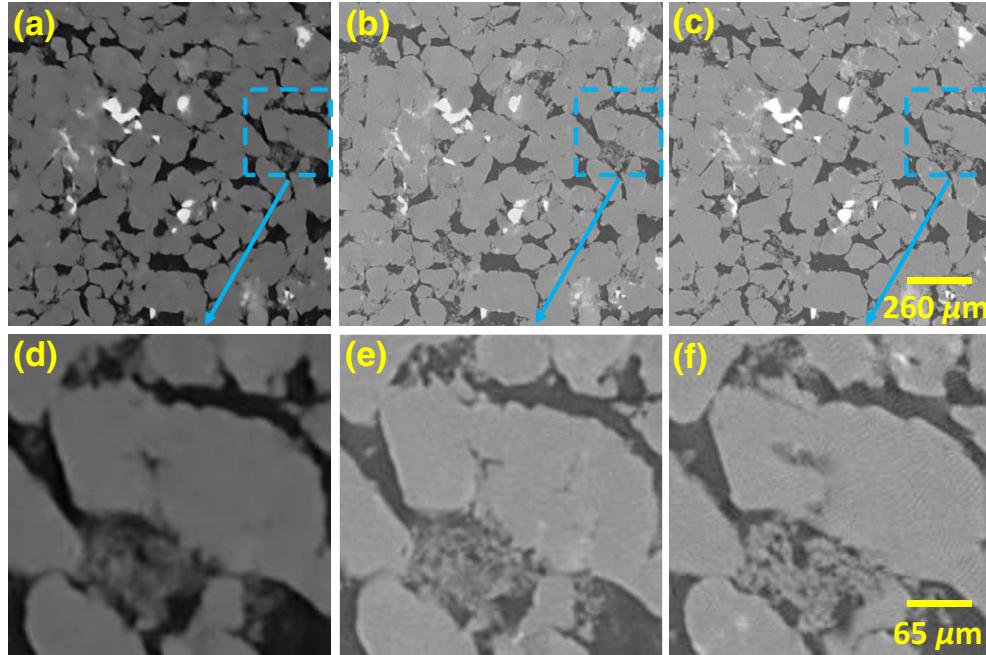


FIG. 8. Example 2D result comparison. (a) TomoGAN SVGG data, (b) final TomoGAN GVGG data, (c) GT data, (d) region of interest (ROI) for TomoGAN SVGG data, (e) ROI for final TomoGAN GVGG data, (f) ROI for GT data.

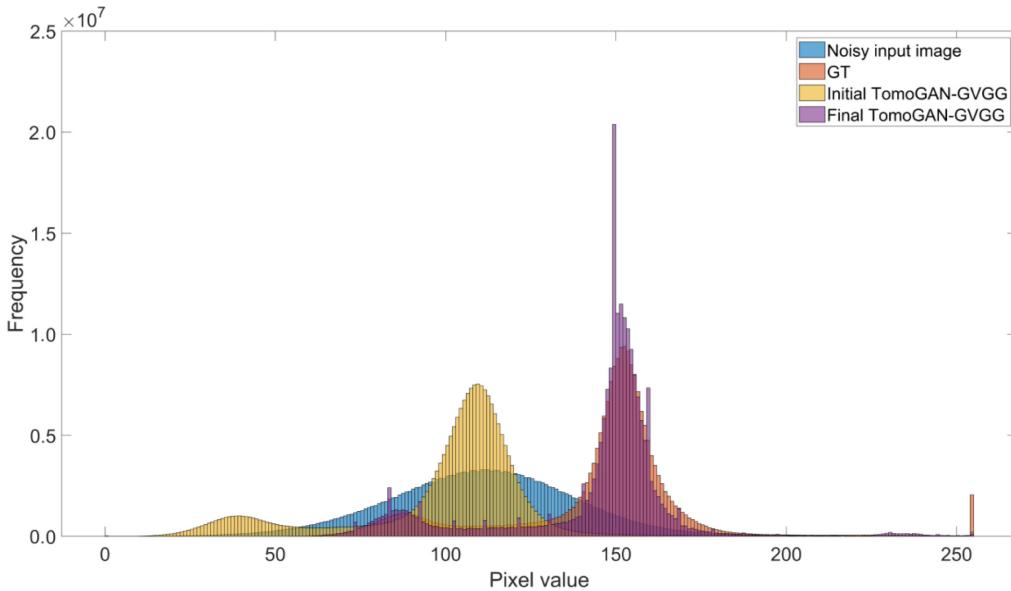


Figure 9 provides a histogram comparison of the 3D data sets. The noisy input image has only one peak on the histogram, which cannot be segmented into pores and solids, while the histogram of GT data provides double peaks for solids and pores. Due to the inconsistent brightness of the reconstructed data, the initial TomoGAN GVGG produces an offset in the pixel distribution. After histogram matching and linear normalization, however, the pixel distribution for the final TomoGAN GVGG result corresponds well with GT data. Furthermore, Table II shows the structural similarity index measure (SSIM) for the noisy input image, initial TomoGAN GVGG, and final TomoGAN GVGG versus GT. Results show that the final TomoGAN GVGG is closer to the GT than the noisy input image or initial TomoGAN GVGG. Overall, both data sets can be segmented with the same settings, as evident in the histogram data (Figure 9). In addition, image histogram results after applying a nonlocal means filter or anisotropic diffusion filter to the noisy data are provided in the Supplemental Material [53] for a baseline comparison. Neither of these traditional filtering methods can restore the double peaks associated with the pore and solid phases.

C. Geometrical analysis

Although perceptual loss is beneficial for obtaining more realistic images than pixel-wise loss, GAN-based

FIG. 9. Histogram comparison for the noisy input image, ground truth, initial TomoGAN GVGG, and final TomoGAN GVGG.

approaches with perceptual loss may generate unrealistic features [48,73]. Even though this effect is balanced by the other loss functions, it is important to evaluate whether the reconstructed images are physically equivalent to the GT images. To achieve this, we compare the physical accuracy of the final TomoGAN GVGG and GT data sets by assessing the segmented-data geometrical properties.

The 3D MF results for the TomoGAN GVGG and GT data sets are shown in Table III. The presented values are for the resulting pore space after image segmentation and extraction of the connected pore space. The largest percentage difference between the GT and TomoGAN GVGG data sets is for the integral of mean curvature. This results from higher-curvature surfaces in the TomoGAN GVGG data, which could result from the sharper high-contrast edges, as observed in Fig. 8(d), than those observed for the GT data [Fig. 8(e)]. All other geometrical measures provide a percentage difference of less than 8.36%.

D. Petrophysical analysis

Figures 10(a) and 10(b) provide 3D pore-structure renderings for the GT and final TomoGAN GVGG data sets. Based on the single-phase flow simulation, we obtain absolute permeability results of 57 and 76 mD for the TomoGAN GVGG and GT data sets, respectively. This results in a difference of 19 mD (25%) in permeability between the two data sets. Figure 10(c) shows that TomoGAN GVGG slightly overestimates the pore size in the medium-radius range (5–11 voxels) and underestimates it for other radii. We believe that the discrepancy in the petrophysical properties is mainly caused by the linear normalization step, which may slightly change the image histogram. In addition, the use of the perceptual loss of TomoGAN GVGG may also bias the reconstructed

TABLE II. SSIM for the noisy input image, initial TomoGAN GVGG, and final TomoGAN GVGG vs GT.

	Noisy input image vs GT	Initial TomoGAN GVGG vs GT	Final TomoGAN GVGG vs GT
SSIM	0.2238	0.5079	0.5957

TABLE III. 3D MFs for the 3D final TomoGAN GVGG and GT data sets.

Pore phase	Final TomoGAN GVGG	Ground truth	Percent difference (%)
Volume fraction (pore space)	0.18	0.19	5.56
Surface area, M_1 (μm^2)	7.28×10^7	7.80×10^7	6.70
Mean curvature, M_2 (μm)	8.00×10^5	6.51×10^5	22.91
Euler characteristic, $\chi(X)$	-4.46×10^4	-4.11×10^4	8.36

image towards unrealistic features. However, the permeability and pore size distributions for the two data sets are still identified to be equivalent. Lastly, the capillary pressure curve is a critical factor when considering multiphase flow in porous rocks [74]. The simulated drainage-capillary-pressure curve [Fig. 10(d)] shows that the TomoGAN GVGG data can produce a curve that is nearly equivalent to that provided by the GT data. The petrophysical accuracy can be further verified by the imbibition relative permeability results shown in Fig. 11.

The TomoGAN GVGG results underestimate the relative permeability in the oil phase compared with the GT result, while the relative permeability in the water phase for both results is equivalent. Nevertheless, TomoGAN GVGG data still provide physically accurate relative permeability data with a 5.49% difference in the crossover point and 0.64% difference in the end-point oil-phase relative permeability. The simulation results can also be observed visually from Fig. 12, which shows the steady-state velocity fields and phase fields at $S_w = 0.45$, near the relative

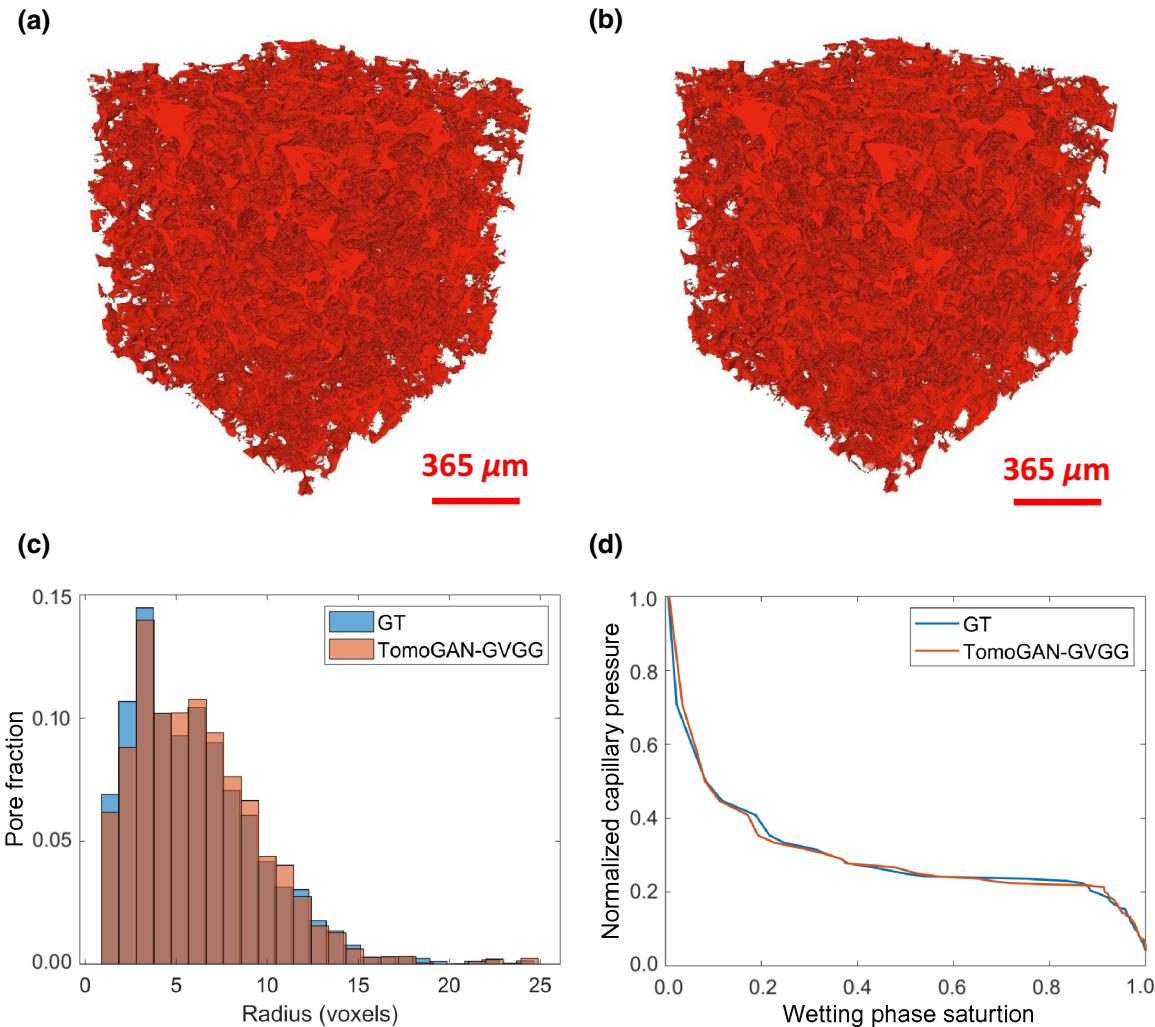


FIG. 10. Petrophysical analysis for TomoGAN GVGG and GT data. (a) 3D rendering of pore structure of TomoGAN GVGG data, (b) 3D rendering of pore structure of GT data, (c) pore size distribution, (d) drainage capillary pressure versus saturation curve.

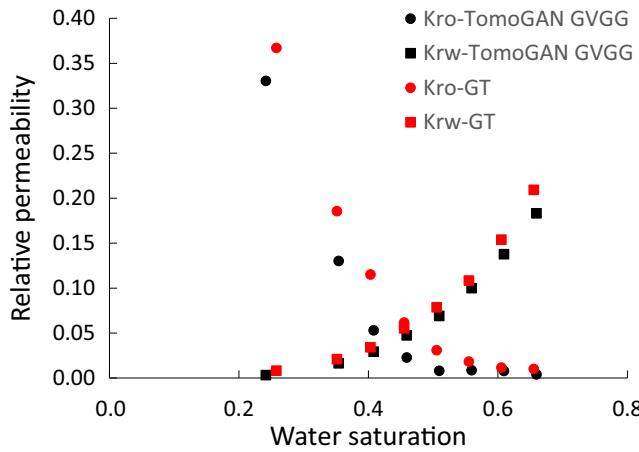


FIG. 11. Imbibition relative permeability versus water saturation for GT and TomoGAN GVGG data.

permeability crossover point. Overall, the proposed GAN-based approach provides a physically accurate solution that restores noisy digital rock data to a level that is comparable to the GT data.

V. CONCLUSION

A GAN with a geometrical-based loss function is presented for noise removal from digital rock images. A fine-tuning method is applied to obtain a geometrical-based VGG19 for digital rock data. The fine-tuning process transforms a classification problem into a regression problem to “highlight” important geometrical features of porous rock. We reconstruct a cubic region of 600^3 voxels that is not used for training to evaluate the performance of TomoGAN SVGG and TomoGAN GVGG. By comparing the TomoGAN GVGG results with those of TomoGAN SVGG, the results show that TomoGAN SVGG causes blurring, which indicates that the pretrained VGG19 from ImageNet cannot capture features related to digital rock images. In contrast, TomoGAN-GVGG can reconstruct more realistic images with finer details of mineralogical features within the pore space.

Although the GAN-based approach can generate photo-realistic images based on perceptual loss, the reconstructed data may contain unrealistic features. By analyzing the MFs and petrophysical results, we find that the proposed TomoGAN GVGG can provide a reliable and accurate way

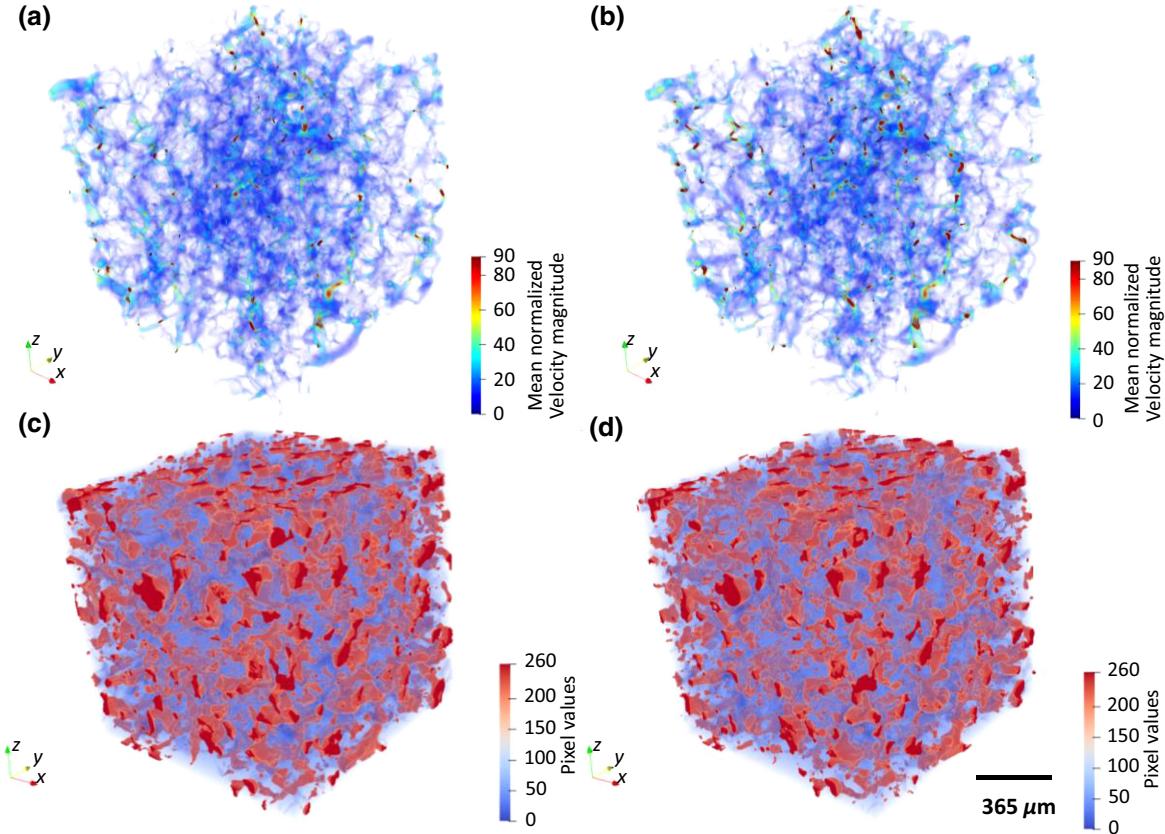


FIG. 12. (a) Visualization of steady-state velocity field of 3D GT data, (b) visualization of steady-state velocity field of 3D TomoGAN GVGG data, (c) visualization of phase fields at $S_w = 0.45$ (near the relative-permeability crossover point) for 3D GT data from imbibition relative permeability results, (d) visualization of phase fields at $S_w = 0.45$ (near the relative-permeability crossover point) for 3D TomoGAN GVGG data from imbibition relative permeability results.

to restore noisy digital rock images. By reducing the scanning time from 9.5 h to 7 min, the expenditure of collecting micro-CT can be decreased significantly. This is particularly important for applications where time-lapse images of a dynamic process are required, high-throughput imaging is necessary for real-time data analysis, or where the quantification of large sample volumes is required.

ACKNOWLEDGMENT

This research is undertaken with the assistance of resources and services from the National Computational Infrastructure (NCI), which is supported by the Australian Government.

- [1] M. J. Blunt, B. Bijeljic, H. Dong, O. Gharbi, S. Iglaue, P. Mostaghimi, A. Paluszny, and C. Pentland, Pore-scale imaging and modelling, *Adv. Water Res.* **51**, 197 (2013).
- [2] A. Heiko, N. Combaret, J. Dvorkin, E. Glatt, J. Han, M. Kabel, Y. Keehm, *et al.*, Digital rock physics benchmarks—part I: Imaging and segmentation, *Comput. Geosci.* **50**, 25 (2013).
- [3] J. H. Dunsuir, S. R. Ferguson, K. L. D'Amico, and J. P. Stokes, in *SPE Annual Technical Conference and Exhibition, Society of Petroleum Engineers* (1991).
- [4] P. Mostaghimi, M. J. Blunt, and B. Bijeljic, Computations of absolute permeability on micro-CT images, *Math. Geosci.* **45**, 103 (2013).
- [5] P. Mostaghimi, R. T. Armstrong, A. Gerami, Y. Hu, Y. Jing, F. Kamali, M. Liu, *et al.*, Cleat-scale characterisation of coal: An overview, *J. Nat. Gas Sci. Eng.* **39**, 143 (2017).
- [6] C. Manwart, U. Aaltosalmi, A. Koponen, R. Hilfer, and J. Timonen, Lattice-Boltzmann and finite-difference simulations for the permeability for three-dimensional porous media, *Phys. Rev. E* **66**, 016702 (2002).
- [7] S. Succi, M. Sbragaglia, and S. Ubertini, Lattice boltzmann method, *Scholarpedia* **5**, 9507 (2010).
- [8] I. Fatt, The network model of porous media, *Trans. AIME* **207**, 144 (1956).
- [9] L. Xavier, P. H. Valvatne, and M. J. Blunt, Predictive network modeling of single-phase non-newtonian flow in porous media, *J. Colloid Interface Sci.* **264**, 256 (2003).
- [10] A. G. Yiotis, I. N. Tsimpanogiannis, A. K. Stubos, and Y. C. Yortsos, Pore-network study of the characteristic periods in the drying of porous materials, *J. Colloid Interface Sci.* **297**, 738 (2006).
- [11] R. C. Gonzalez and R. E. Woods, *Digital Image Processing* (Prentice Hall, Upper Saddle River, NJ, USA, 2006), 3rd ed.
- [12] M. Raghav and S. Raheja, Image denoising techniques: Literature review, *Int. J. Eng. Comput. Sci.* **3**, 5637 (2014).
- [13] A. Aboshosha, M. Hassan, M. Ashour, and M. El Mashade, Image denoising based on spatial filters, an analytical study, *IEEE Int. Conf. Comput. Eng. Syst.*, 245 (2009).
- [14] J. Benesty, J. Chen, and Y. Huang, in *2010 IEEE International Conference on Acoustics, Speech and Signal Processing IEEE* (2010), pp. 205–208.
- [15] A. Buades, B. Coll, and J. M. Morel, A review of image denoising algorithms, with a new one, *Multiscale Model. Simul.* **4**, 490 (2005).
- [16] J. L. Starck, E. J. Candès, and D. L. Donoho, The curvelet transform for image denoising, *IEEE Trans. Image Process.* **11**, 670 (2002).
- [17] G. Yu and G. Sapiro, DCT image denoising: A simple and effective image denoising algorithm, *Image Process. Line* **1**, 292 (2011).
- [18] Z. Hou, Adaptive singular value decomposition in wavelet domain for image denoising, *Pattern Recognit.* **36**, 1747 (2003).
- [19] M. K. Mihcak, I. Kozintsev, and K. Ramchandran, in *IEEE International Conference on Acoustics, Speech, and Signal Processing Proceedings ICASSP99 (Cat. No. 99CH36258)* (1999), pp. 3253–3256.
- [20] S. Yin, L. Cao, Y. Ling, and G. Jin, Image denoising with anisotropic bivariate shrinkage, *Signal Process.* **91**, 2078 (2011).
- [21] C. Knous and M. Zwicker, Dual-domain filtering, *SIAM J. Imaging Sci.* **8**, 1396 (2015).
- [22] N. Wang, Y. Shang, Y. Chen, M. Yang, Q. Zhang, Y. Liu, and Z. Gui, A hybrid model for image denoising combining modified isotropic diffusion model and modified perona-malik model, *IEEE Access* **6**, 33568 (2018).
- [23] K. Gupta and S. K. Gupta, Image denoising techniques—a review paper, *IJITEE* **2**, 6 (2013).
- [24] B. Goyal, A. Dogra, S. Agrawal, B. S. Sohi, and A. Sharma, Image denoising review: From classical to state-of-the-art approaches, *Inf. Fusion* **55**, 220 (2020).
- [25] L. W. Goldman, Principles of CT: Radiation dose and image quality, *J. Nucl. Med. Technol.* **35**, 213 (2007).
- [26] S. Bonnet, A. Koenig, S. Roux, P. Hugonnard, R. Guillemaud, and P. Grangeat, Dynamic X-ray computed tomography, *Proc. IEEE* **91**, 1574 (2003).
- [27] G. R. Myers, A. M. Kingston, T. K. Varslot, M. L. Turner, and A. P. Sheppard, Dynamic tomography with a priori information, *Appl. Opt.* **50**, 3685 (2011).
- [28] S. Berg, H. Ott, S. A. Klapp, A. Schwing, R. Neiteler, N. Brussee, A. Makurat, L. Leu, F. Enzmann, J. O. Schwarz, and M. Kersten, Real-time 3D imaging of haines jumps in porous media flow, *Proc. Natl. Acad. Sci.* **110**, 3755 (2013).
- [29] R. T. Armstrong, A. Georgiadis, H. Ott, D. Klemin, and S. Berg, Critical capillary number: Desaturation studied with fast X-ray computed microtomography, *Geophys. Res. Lett.* **41**, 55 (2014).
- [30] T. Bultreys, M. A. Boone, M. N. Boone, T. De Schryver, B. Masschaele, L. Van Hoorebeke, and V. Cnudde, Fast laboratory-based micro-computed tomography for pore-scale research: Illustrative experiments and perspectives on the future, *Adv. Water Res.* **95**, 341 (2016).
- [31] G. R. Myers, A. M. Kingston, T. K. Varslot, M. L. Turner, and A. P. Sheppard, in *International Symposium of the Society of Core Analysts*, Austin, TX, (2011), pp. 1–12.
- [32] S. Youssef, H. Deschamps, J. Dautriat, E. Rosenberg, R. Oughanem, E. Maire, and R. Mokso, in *International Congress on* (2013).
- [33] L. Xu, J. S. Ren, C. Liu, and J. Jia, Deep convolutional neural network for image deconvolution, *Adv. Neural Inf. Process. Syst.* **27**, 1790 (2014).

- [34] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, Beyond a Gaussian denoiser: Residual learning of deep cnn for image denoising, *IEEE Trans. Image Process.* **26**, 3142 (2017).
- [35] J. Xie, L. Xu, and E. Chen, Image denoising and inpainting with deep neural networks, *Adv. Neural Inf. Process. Syst.* **25**, 341 (2012).
- [36] H. M. Li, Deep learning for image denoising, *international journal of signal processing, Image Process. Pattern Recognit.* **7**, 171 (2014).
- [37] J. M. Wolterink, T. Leiner, M. A. Viergever, and I. Isgum, Generative adversarial networks for noise reduction in low-dose CT, *IEEE Trans. Med. Imaging* **36**, 2536 (2017).
- [38] Q. Yang, P. Yan, Y. Zhang, H. Yu, Y. Shi, X. Mou, M. K. Kalra, Y. Zhang, L. Sun, and G. Wang, Low-dose CT image denoising using a generative adversarial network with wasserstein distance and perceptual loss, *IEEE Trans. Med. Imaging* **37**, 1348 (2018).
- [39] Z. Liu, T. Bicer, R. Kettimuthu, D. Gursoy, F. De Carlo, and I. Foster, TomoGAN: Low-dose synchrotron x-ray tomography with generative adversarial networks: Discussion, *J. Opt. Soc. Am. A* **37**, 422 (2020).
- [40] Y. D. Wang, R. T. Armstrong, and P. Mostaghimi, Boosting resolution and recovering texture of 2D and 3D micro-CT images with deep learning, *Water Resour. Res.* **56**, e2019WR026052 (2020).
- [41] Y. D. Wang, R. T. Armstrong, and P. Mostaghimi, Enhancing resolution of digital rock images with super resolution convolutional neural networks, *J. Pet. Sci. Eng.* **182**, 106261 (2019).
- [42] Y. Niu, Y. D. Wang, P. Mostaghimi, P. Swietojanski, and R. T. Armstrong, An innovative application of generative adversarial networks for physically accurate rock images with an unprecedented field of view, *Geophys. Res. Lett.* **47**, e2020GL089029 (2020).
- [43] Z. Fang, H. Gong, G. Yang, and T. Zhang, in *MIPPR: Medical Imaging, Parallel Processing of Images, and Optimization Techniques* (International Society for Optics and Photonics, Yichang, China, 2009), Vol. 7497, pp.74972B.
- [44] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, Generative adversarial networks, arXiv preprint [arXiv:1406.2661](https://arxiv.org/abs/1406.2661) (2014).
- [45] M. Arjovsky, S. Chintala, and L. Bottou, in *International Conference on Machine Learning*, PMLR, (2017), pp. 214–223.
- [46] O. Ronneberger, P. Fischer, and T. Brox, in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer, Cham, (2015), pp. 234–241.
- [47] C. Dong, C. C. Loy, K. He, and X. Tang, Image super-resolution using deep convolutional networks, *IEEE Trans. Pattern Anal. Mach. Intell.* **38**, 295 (2015).
- [48] C. Ledig, L. Theis, F. Huszár, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang, and W. Shi, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2017), pp. 4681–4690.
- [49] F. Milletari, N. Navab, and S. A. Ahmadi, in *2016 Fourth International Conference on 3D Vision (3DV)*, IEEE, (2016), pp. 565–571.
- [50] K. Simonyan and A. Zisserman, Very deep convolutional networks for large-scale image recognition, arXiv preprint [arXiv:1409.1556](https://arxiv.org/abs/1409.1556) (2014).
- [51] Y. Jiang and J. Li, Generative adversarial network for image super-resolution combining texture loss, *Appl. Sci.* **10**, 1729 (2020).
- [52] I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin, and A. Courville, Improved training of wasserstein gans, arXiv preprint [arXiv:1704.00028](https://arxiv.org/abs/1704.00028) (2017).
- [53] See the Supplemental Material at <http://link.aps.org/supplemental/10.1103/PhysRevApplied.15.064033> for padding details, the relation between the pore Euler characteristic and removal of the maximum object voxels, the noise removal by non-local mean filter and anisotropic diffusion filter, and the image histogram comparison after applying a non-local means filter or anisotropic diffusion filter to the noisy data.
- [54] K. You, M. Long, J. Wang, and M. I. Jordan, How does learning rate decay help modern neural networks?, arXiv preprint [arXiv:1908.01878](https://arxiv.org/abs/1908.01878) (2019).
- [55] D. P. Kingma and J. Ba, Adam: A method for stochastic optimization, arXiv preprint [arXiv:1412.6980](https://arxiv.org/abs/1412.6980) (2014).
- [56] J. Deng, W. Dong, R. Socher, L. J. Li, K. Li, and L. Fei-Fei, in *2009 IEEE Conference on Computer Vision and Pattern Recognition*, IEEE, (2009), pp. 248–255.
- [57] M. E. Peters, S. Ruder, and N. A. Smith, in *Proceedings of the 4th Workshp on Representation Learning for NLP, Florence, Italy*, [arXiv:1903.05987](https://arxiv.org/abs/1903.05987) (2019).
- [58] E. C. Too, L. Yujian, S. Njuki, and L. Yingchun, A comparative study of fine-tuning deep learning models for plant disease identification, *Comput. Electron. Agric.* **161**, 272 (2019).
- [59] J. Serra, *Image Analysis and Mathematical Morphology* (Academic Press, Inc, 1983).
- [60] R. T. Armstrong, J. E. McClure, V. Robins, Z. Liu, C. H. Arns, S. Schlüter, and S. Berg, Porous media characterization using minkowski functionals: Theories, applications and future directions, *Transp. Porous Media* **130**, 305 (2019).
- [61] K. R. Mecke, in *Statistical Physics and Spatial Statistics* (Springer, Berlin, Heidelberg, 2000), pp. 111–184.
- [62] C. H. Arns, M. A. Knackstedt, and K. R. Mecke, Characterisation of irregular spatial structures by parallel sets and integral geometric measures, *Colloids Surf., A* **241**, 351 (2004).
- [63] D. Legland, K. Kiêu, and M. F. Devaux, Computation of minkowski measures on 2D and 3D binary images, *Image Anal. Stereol.* **26**, 83 (2007).
- [64] S. Beucher, in *Proceedings of the International Workshop on Image Processing*, CCETT, (1979).
- [65] X. Wang, K. Yu, S. Wu, J. Gu, Y. Liu, C. Dong, Y. Qiao, and C. Change Loy, in *Proceedings of the European Conference on Computer Vision (ECCV) Workshops* (2018).
- [66] K. Michielsen and H. De Raedt, Integral-geometry morphological image analysis, *Phys. Rep.* **347**, 461 (2001).
- [67] J. E. McClure, J. F. Prins, and C. T. Miller, A novel heterogeneous algorithm to simulate multiphase flow in porous media on multicore CPU-GPU systems, *Comput. Phys. Commun.* **185**, 1865 (2014).

- [68] Y. D. Wang, T. Chung, R. T. Armstrong, J. McClure, T. Ramstad, and P. Mostaghimi, Accelerated computation of relative permeability by coupled morphological and direct multiphase flow simulation, *J. Comput. Phys.* **401**, 108966 (2020).
- [69] V. Shabro, C. Torres-Verdín, F. Javadpour, and K. Sepehrnoori, Finite-difference approximation for fluid-flow simulation and calculation of permeability in porous media, *Transp. Porous Media* **94**, 775 (2012).
- [70] T. Chung, Y. D. Wang, R. T. Armstrong, and P. Mostaghimi, Voxel agglomeration for accelerated estimation of permeability from micro-CT images, *J. Pet. Sci. Eng.* **184**, 106577 (2020).
- [71] A. N. Strahler, Quantitative analysis of watershed geomorphology, *Eos, Transactions American Geophysical Union* **38**, 913 (1957).
- [72] D. Silin and T. Patzek, Pore space morphology analysis using maximal inscribed spheres, *Phys. A* **371**, 336 (2006).
- [73] C. Wang, C. Xu, C. Wang, and D. Tao, Perceptual adversarial networks for image-to-image transformation, *IEEE Trans. Image Process.* **27**, 4066 (2018).
- [74] V. P. Dimri, R. P. Srivastava, and N. Vedanti, *Frac-tal Models in Exploration Geophysics: Applications to Hydrocarbon Reservoirs* (Elsevier, Oxford, UK, 2012), Vol. 41.