Generalizable Framework of Unpaired Domain Transfer and Deep Learning for the Processing of Real-Time Synchrotron-Based X-Ray Microcomputed Tomography Images of Complex Structures

Kunning Tang ⊕,[1,*] Ying Da Wang,[1] James McClure,[2] Cheng Chen,[3] Peyman Mostaghimi,[1] 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, USA*
[3]*Department of Mining and Minerals Engineering, Virginia Tech, Blacksburg, Virginia, USA*

Mitigating greenhouse gas emissions by underground carbon dioxide storage or by coupling intermittent renewable energy with underground hydrogen storage are solutions essential to the future of energy. Of particular importance to the success of underground storage is the fundamental understanding of geochemical reactions with mineralogical phases and flow behavior at the length scale at which interfaces are well resolved. Fast synchrotron-based three-dimensional x-ray microcomputed tomography (μCT) of rocks is a widely used technique that provides real-time visualization of fluid flow and transport mechanisms. However, fast imaging results in significant noise and artifacts that complicate the extraction of quantitative data beyond the basic identification of solid and void regions. To address this issue, an image-processing workflow is introduced that begins with unpaired domain transfer by cycle-consistent adversarial network, which is used to transfer synchrotron-based μCT images containing fast-imaging-associated noise to long-scan high-quality μCT images that have paired ground truth labels for all phases. The second part of the workflow is multimineral segmentation of images using convolutional neural networks (CNNs). Four CNNs are trained using the transferred dynamic-style μCT images. A quantitative assessment of physically meaningful parameters and material properties is carried out. In terms of physical accuracy, the results show a high variance for each network output, which indicates that the segmentation performance cannot be fully revealed by pixel-wise accuracy alone. Overall, the integration of unpaired domain transfer with CNN-based multimineral segmentation provides a generalizable digital material framework to study the physics of porous materials for energy-related applications, such as underground CO_2 and H_2 storage.

I. INTRODUCTION

The adoption of the Paris Agreement in 2015 renewed enthusiasm toward greenhouse gas emission mitigation and the transition from fossil fuels to renewable-energy-based systems [1–3]. The capture and geological storage of emitted CO_2 (CCS) is a promising method for reducing atmospheric greenhouse gas, and underground hydrogen storage (UHS), which stores excess renewable energy in the form of hydrogen in geological structures, is an emerging means of resolving renewable energy intermittency [4–8]. When it comes to the research regarding containment, capillary trapping and hydrogen loss are two key factors that determine the feasibility of CCS and UHS. The performance of CO_2 capillary trapping significantly depends on the wettability of the surrounding minerals [8,9], and hydrogen loss is mainly controlled by the mineral types and their dissolution and precipitation processes [10–12]. Therefore, understanding reservoir rock mineralogy and the corresponding flow behavior is essential for deciphering the CO_2 and H_2 storage efficiencies.

With the aid of digital rock physics (DRP), rock mineralogy and flow dynamics can be characterized comprehensively on the length scale at which interfaces are well resolved [10,13–15]. The workflow for DRP provides a generalizable technique for mineralization and flow analyses in complex micrometer-sized structures; it starts with image acquisition, image processing, and subsequent physical measures and/or numerical simulation of flow and transport mechanisms [13,16–19]. To date, the focus has extended from static to dynamic analysis, such as dynamic image acquisition by synchrotron-based x-ray

_____
*z5189000@ad.unsw.edu.au

microcomputed tomography (μCT), which provides three-dimensional (3D) real-time images of fluid flow and transport phenomena in porous media [20,21]. This dynamic imaging technique can be potentially used to capture the dynamics of CO_2 and H_2 distribution, flow behavior, and reaction with minerals in porous rocks. However, the fast acquisition time results in a relatively low-intensity signal in comparison with standard imaging techniques, where acquisition times are an order of magnitude greater. The low signal-to-noise ratio results in images that are challenging to process, especially for image segmentation. Therefore, these images are segmented into void and solid phases; however, for many physical processes, the actual distribution of the constitutive minerals and their interactions with the flowing fluids are important. With multiphase flow, the varying mineralogy and associated wetting properties are important. Thus, without the development of an image-processing workflow for dynamic data, it is challenging to capture accurate pore structures for the simulation of physical processes within a digital framework.

Three-dimensional μCT is a noninvasive and nondestructive imaging tool that has been used to capture rock features and fluid and/or gas distributions to support compositional characterization, pore network modeling, and the assessment of mineralogy and other petrophysical parameters [22–25]. Numerical simulations of fluid flow at the pore scale, such as multimineral reactive transport and single-phase and two-phase flow, which are highly related to underground CO_2 and H_2 storage, can be directly performed on μCT images [26,27]. Several studies have substantiated that numerical simulations of 3D μCT images agree reasonably well with experimental results [28–31]. However, several hours are required to obtain high-quality 3D μCT images, which limits their utilization for the imaging of dynamic processes [32,33]. Important fundamental scientific questions remain regarding the role of transient processes during multiphase flow [20,34]. To address these issues, synchrotron-based μCT imaging has attracted attention owing to its extremely high photon flux, which makes it possible to capture full 3D images within seconds [35]. Therefore, real-time 3D dynamic imaging can be achieved with synchrotron-based μCT [20] but at the expense of image quality [36].

Accurate feature characterization and flow simulation of μCT images rely on image segmentation. Image segmentation in DRP refers to the process of partitioning μCT images into multiple mineral phases, pores, and fluid and/or gas phases [18]. To date, several studies have performed μCT image segmentation by proposing machine learning methods, especially the use of advanced convolutional neural networks (CNNs) [37–41]. A CNN is built within a deep learning framework that performs multiphase segmentation with the benefit of eliminating user judgment of the parameters associated with segmentation.

CNNs can capture informative features and semantics from images with receptive fields by stacking several convolutional layers with nonlinearities and down-sampling layers. CNNs have been widely used in the field of computer vision, including object detection [42], image classification [43,44], superresolution [45], image-to-image translation [46], and image segmentation [47,48]. In regards to CNN-based semantic segmentation, compared to traditional segmentation techniques, including edge-detection methods, thresholding methods, or region-based methods, CNN relies less on the voxel intensity frequency distribution and reduces the requirement of expert intervention [49,50]. A study of segmentation with both a watershed segmentation technique and CNN methods demonstrated that CNN gives a better segmentation result in terms of phase boundaries and connectivity [51]. A study of dual-energy x-ray absorptiometry images to distinguish between bone and soft tissue compared the segmentation accuracy between Otsu's thresholding and a deep learning method [52]. The pixel-based accuracy using deep learning was significantly higher than Otsu's thresholding method. These results suggest that CNNs have matured to the point that they outperform other methods for semantic segmentation [52–54].

Specifically to DRP, multimineral segmentation of μCT images can be a complex and time-consuming task using traditional segmentation algorithms because the voxel values for different minerals are incompletely differentiated owing to similar x-ray attenuation coefficients and Poisson-Gaussian noise [55]. Therefore, most studies treat all minerals as a single solid phase and pores as the other phase [56,57]. However, this is not appropriate in all cases, especially for CCS and UHS. Certain types of minerals will cause hydrogen loss, which is observed in the reaction between H_2-saturated brine and calcite [12]. In addition, CO_2 trapping is characterized by the wettability of different types of minerals [25,58,59]. The utilization of CNN-based multimineral segmentation could be one approach to resolve this issue.

For CCS and UHS studies, high pixel-wise accuracy does not necessarily guarantee that petrophysical parameters such as permeability or relative permeability are accurately captured. As shown in previous studies, petrophysical parameters are highly sensitive to small segmentation errors [39]. Therefore, the sensitivity of quantitative measurements to these errors must be considered. A segmentation of six mineral phases on a sandstone sample was performed using several CNNs, including SegNet and U-Net [37]; segmentation errors were evaluated in terms of both pixel-wise accuracy and physical accuracy. Overall, 95% pixel-wise accuracy was achieved, whereas the segmentation results displayed high variance in terms of physical measurements. The commonly occurring phases such as quartz are the main contributors to pixel-wise accuracy, while the accuracy of the less commonly occurring

mineral phases is lower than the overall accuracy. For example, a study reported that in zircon the phase accuracy was only 60% with an overall pixel-wise accuracy of 94% [38]. The issue is that networks tend to overestimate commonly occurring phases and underestimate the less commonly occurring phases when the training data of each phase are imbalanced. This issue was exemplified in the data set used by Refs. [37,38], which contained 61% quartz phase but only 0.21% mica phase as a volume fraction. Less commonly occurring phases are likely to be recognized as noise by the network and networks; therefore, they train for the commonly occurring phases that dominate the accuracy. It is difficult to balance sparse phases because these minerals are rare in rocks, and obtaining training data in the first place is difficult because of the high expense. Data augmentation of sparsely occurring phases and judicious selection of the loss function during training can be carried out to reduce the imbalance to an extent; for example, focus loss reduces the weight of the easy-to-segment phases and forces the network to focus more on the loss of less common phases [60]. However, because the CNN-based segmentation method is a data-driven task, such imbalances are unavoidable when it comes to less commonly occurring phases.

In addition, μCT scanning noise, which is commonly regarded as a random signal and is characterized by a probability density function, is known to significantly influence segmentation results [61,62]. The noise mainly includes Gaussian and Poisson processes [55]. Gaussian noise results from the random distribution of independent signals, whereas Poisson noise is commonly found in situations where photons are accumulated over a detector, such as charge-coupled device cameras [63]. The boundary of each mineral phase is difficult to define with Poisson and Gaussian noise because edge detection is highly sensitive to noise, and the image quality is reduced significantly. Therefore, it is common to leverage several noise removal filters before segmentation. These filters mainly include the nonlocal mean filter and Gaussian filter, which were introduced as edge-preserving denoising and blurring filters to remove additive Gaussian noise [64]. Laplacian, Canny, and Sobel filters are sharpening and edge-detection filters commonly used for boundary detection and feature extraction for supervised machine learning segmentation [65]. However, denoising filters can cause a certain degree of degradation of the details in images and remove the "real information" and fine structures, especially for less-common mineral phases [66]. Although these degraded effects might be acceptable for binary segmentation, they should be strictly avoided in multimineral or multiphase segmentation because these fine structures need to be preserved for segmenting the less prevalent minerals. For real-time synchrotron-based μCT imaging, noise associated with dynamic imaging occurs across all images because the exposure time for each collected radiograph

is significantly reduced. Therefore, it is difficult to segment the images. Moreover, CNN-based segmentation is becoming increasingly difficult to perform because of the limited availability of real-time ground-truth data sets due to the time and cost expense. Overall, performing accurate multimineral segmentation on real-time data with common Gaussian and Poisson noise is essential for fine structure characterization and dynamic image processing.

The aim of this study is to develop an advanced digital material workflow that is generalizable to various physical problems, particularly imaging systems that have complex micrometer-sized structures composed of multiple phases. Our target system is sandstone rock for CCS and UHS applications, while other applicable systems include but are not limited to fuel cells [67], negative compressibility materials [68], flexible metal-organic frameworks [69], and high-thermal-conductivity porous media [70]. We investigate the potential of unpaired domain transfer between real-time synchrotron-based μCT images with associated noise and long-scan traditional μCT images to provide a framework for the imaging of dynamic processes. Domain transfer is performed to provide a robust framework because ground-truth segmented data for real-time synchrotron-based μCT images are not always readily available. Cycle-consistent adversarial network (Cycle-GAN) is widely used for unpaired image-to-image translation [71] and is thus used for transferring synchrotron-based noise to a long-scan μCT data set with a ground-truth counterpart. A total of four CNN architectures are trained to segment the real-time data into six mineral phases. In terms of error assessment, in addition to the commonly used pixel-wise accuracy, region-based and physical accuracy are essential metrics. This is because the purpose of segmentation is to facilitate the quantitative assessment of physically meaningful quantities, such as interface determination, topological connectivity, and permeability, which are key design parameters for CCS and UHS applications. Overall, this study evaluates a digital material platform for the quantitative assessment of complex porous materials using an unpaired domain transfer method for dynamic synchrotron-based μCT, providing real-time analysis of physical processes where multiple phases and complex structures are present.

II. MATERIALS AND METHODS

A. Data sets

Unpaired domain transfer is performed between raw μCT data and synchrotron-based μCT data. A miniplug from the Mt. Simon sandstone reservoir is cored to 5 mm in length and 3 mm in diameter and scanned using μCT at the University of New South Wales. The surface of the sample is polished for subsequent mineral classification by quantitative evaluation of minerals using scanning electron microscopy (QEMSCAN). The two-dimensional (2D)

QEMSCAN mineral maps are then registered to the corresponding cross section of the three-dimensional (3D) μCT voxel [26]. Following this process, 3D mineral segmentation in the μCT image is performed based on the x-ray intensity differences of minerals as guided by the registered 2D QEMSCAN images, which serve as ground truth (GT) for the 3D Mt. Simon sandstone μCT data. Further details on the data preparation can be found in Ref. [37]. In total, the full-size raw μCT data and corresponding GT data are $1100 \times 1100 \times 2200$ voxels, as shown in Figs. 1(a) and 1(b). The GT images are comprised of six phases, labeled from 0 to 5: pore, clay, quartz, feldspar, micas, and a mixed group of less common high-density minerals.

Synchrotron-based μCT images are obtained from GeoSoilEnviroCARS Sector 13 at the Argonne National Laboratory Advanced Photon Source (APS). Miniplug Bentheimer sandstone is obtained by cutting the rock to 5 mm in diameter and 10 mm in length. The real-time 3D-synchrotron-based μCT scan is then conducted with x-ray photon fluxes of approximately 10^{12}–10^{14} photons s^{-1} during a waterflooding experiment. The 3D images are collected in approximately 20 s at a resolution of 3.5 μm; further details are provided in Ref. [32]. It is noted that there is always a trade-off between sample size and image resolution. While a higher resolution image may capture fine features of geometry, it limits the field of view and spatial information at a larger scale, thus affecting determination of properties of porous media. For typical flow characterization of sandstone, the resolution

is often around 2 to 5 μm [72–74]. The synchrotron-based μCT image of the dry Bentheimer sandstone is $1000 \times 1000 \times 1200$ voxels in size, as shown in Fig. 1(c). This data set contains fast-imaging-associated noise that exists across all phases, which makes any attempt at multiphase segmentation challenging. Table I gives a summary of the samples and data sets used in this study.

B. Domain transfer by CycleGAN and image degradation

Noise from the real-time synchrotron-based μCT data is transferred to the raw μCT data using CycleGAN. CycleGAN comprises two generators and two discriminators that perform unpaired image-style transfer. The generators are based on an encoder-decoder structure that applies three convolutional layers in the down-sampling steps, followed by nine residual blocks and sequential decoding steps using up-sampling layers instead of transposed convolutional layers to avoid prediction artifacts [75]. Instance normalization is used because of its advantage in image-style transfer, which normalizes each image individually without considering the image content of the entire batch [76]. Using this process, the features of the two μCT sandstone data sets are captured and transferred between each other by the generators. Two styles of fake μCT sandstone images and real μCT sandstone images are then processed through two discriminators; PatchGAN [46] is used as a discriminator to determine whether an $N \times N$ output is
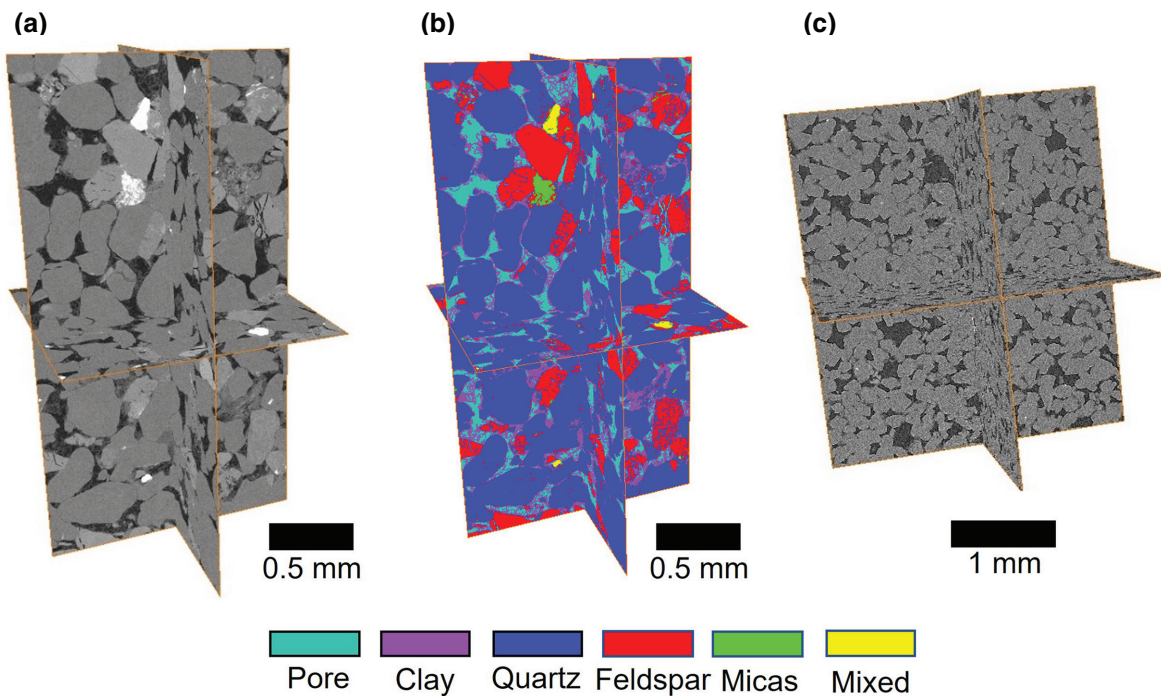

FIG. 1. (a) Full-size raw μCT data with a voxel size of $1100 \times 1100 \times 2200$. (b) Segmented GT data set with six phases presented. (c) Full-size synchrotron-based μCT data containing fast-imaging-associated noise, measuring $1000 \times 1000 \times 1200$.

TABLE I. Basic information for the samples and data sets used in this study. The nomenclature of each data set is used throughout this paper.

Samples	Data set	Nomenclature	Voxel size
Mt. Simon Sandstone	Long-scan Mt. Simon sandstone μCT data	Raw μCT data	$1100 \times 1100 \times 2200$
Mt. Simon Sandstone	Domain-transferred Mt. Simon μCT data	Dynamic-based μCT data	$1100 \times 1100 \times 2200$
Bentheimer sandstone	Synchrotron-based Bentheimer sandstone μCT data	Synchrotron-based μCT data	$1000 \times 1000 \times 1200$
Bentheimer sandstone	Domain-transferred Bentheimer synchrotron-based μCT data	Static-styled synchrotron data	$1000 \times 1000 \times 1200$

fake or real. Overall, two discriminators are trained to distinguish the fake and real images, while the two generators are trained to produce fake images that appear similar to real images. The detailed workflow of CycleGAN is shown in Fig. 2.

The objective loss function for training CycleGAN contains two types of loss: (1) adversarial loss and (2) cycle consistency loss. Adversarial loss acts as the loss for the discriminator, whereas the cycle consistency loss acts as the loss for the generator. The total objective loss function is

$$L(G, F, D_X, D_Y) = L_{\text{GAN}}(G, D_Y, X, Y) \\ + L_{\text{GAN}}(F, D_X, X, Y) + L_{\text{cyc}}(G, F),$$

(1)

where X, Y are the two different domains. $L_{\text{GAN}}(G, D_Y, X, Y)$ includes the loss between the fake image generated by the mapping function G and the real image that needs to be minimized, and the loss for the discriminator D_Y to distinguish the fake and real image that needs to be maximized. A similar objective loss, $L_{\text{GAN}}(F, D_X, X, Y)$, is used for another mapping function F and discriminator D_X. $L_{\text{cyc}}(G, F)$ refers to the losses during feature mapping, including the loss between the real image and the fake image, the fake image with the reconstructed real image, and the real image with the reconstructed real image. The loss function for the adversarial loss is

$$L_{\text{MSE}} = \frac{1}{n} \sum_{t=1}^{n} [y - f(x_t)]^2.$$
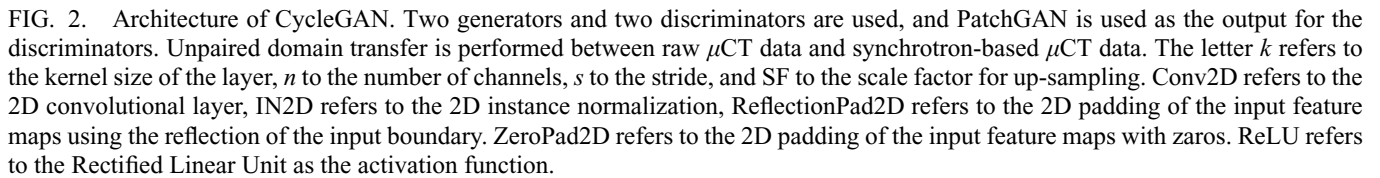
(2)

The $L1$ function used for cycle consistency loss is

$$L_1 = \frac{1}{n} \sum_{t=1}^{n} |y - f(x_t)|,$$

(3)

where y is the GT pixel value and $f(x_t)$ is the network prediction. CycleGAN is trained with an initial learning rate of 0.0001 using the ADAM solver with a batch size of $8 \times 192 \times 192$. In total, the 2D data sets comprise 4800 raw

μCT images and 4800 synchrotron-based μCT images. The data are then split into 4000 images for training and 800 images for testing.

C. Image degradation

Gaussian-Poisson noise that commonly exists in μCT images results in reduced image quality and subsequently affects segmentation results and further DRP analyses. In addition, the utilization of denoising filters may add additional artifacts and/or smooth finer details of the pore structure. Therefore, by considering the fine structure and topology as highly essential for the segmentation of sparsely occurring minerals, we simulate a scenario where there is a certain degree of Gaussian-Poisson noise in our raw μCT data and test the capability of CNN-based multimineral segmentation methods to distinguish minerals with such noise. Gaussian and Poisson noises are manually added to the raw μCT data by using a Gaussian filter and Poisson distribution in NUMPY and SCIPY packages in PYTHON after domain transfer. Random Gaussian noise is added to each image using a Gaussian filter with a standard range of 2–3. Poisson noise is added with the expectation of intervals based on the pixel values of each image. To test whether the degree of noise is realistic, no-reference image quality metrics, blind/referenceless image spatial quality evaluator (BRISQUE) and the naturalness image quality evaluator (NIQE), are calculated using the statistical features of the input image to evaluate the similarity of our data to other dynamic synchrotron-based μCT data. Overall, after domain transfer and image degradation, the raw μCT data that contain noise associated with dynamic synchrotron scanning and commonly occurring Gaussian-Poisson noise are used to mimic the realistic fast synchrotron-based scanning results. Therefore, instead of attempting to segment directly on the synchrotron-based μCT data, the CNN-based network can be trained using the domain-transferred Mt. Simon μCT image. This trained network can then be used later to segment a real synchrotron-based μCT image.

FIG. 2. Architecture of CycleGAN. Two generators and two discriminators are used, and PatchGAN is used as the output for the discriminators. Unpaired domain transfer is performed between raw μCT data and synchrotron-based μCT data. The letter k refers to the kernel size of the layer, n to the number of channels, s to the stride, and SF to the scale factor for up-sampling. Conv2D refers to the 2D convolutional layer, IN2D refers to the 2D instance normalization, ReflectionPad2D refers to the 2D padding of the input feature maps using the reflection of the input boundary. ZeroPad2D refers to the 2D padding of the input feature maps with zaros. ReLU refers to the Rectified Linear Unit as the activation function.

D. Segmentation CNN architectures and training schedules

Multimineral segmentation is performed in 2D using four CNN architectures based on the encoder-decoder structure, which exploits features from the encoding step and recovers the spatial resolution from the decoding step. We train all networks in 2D because the QEMSCAN image is generated in 2D and the domain transfer by CycleGAN is performed in 2D. As shown by Ref. [37], networks in 2D and 3D provide similar pixel-wise and physical accuracy, while networks in 2D are computationally more efficient compared to 3D because they have fewer trainable parameters. The CNN networks that contain both pretrained and nonpretrained models are U-ResNet [37,43], U-ResNet-cGAN [46], U-Net with EfficientNet-B3 as the backbone (EfficientU-Net), and EfficientU-Net-cGAN. The advantages of each network and the main differences are listed in Table II. The main reasons for selecting these networks are as follows:

(1) U-ResNet has been proven to perform better in multimineral segmentation than SegNet and U-Net, which are commonly used for semantic-segmentation CNN networks [37]. Therefore, U-ResNet is tested as a baseline to check whether other networks can perform better.

(2) With the addition of the cGAN module, both binary-cross-entropy loss and cross-entropy loss are used to regulate training, which is beneficial in preventing overfitting and edge determination.

(3) One of the state-of-the-art network architectures in image classification is EfficientNet, which has the advantage of a high-balancing network depth, width, and resolution. Therefore, we use EfficientNet for feature extraction

to improve the training efficiency. In addition, although EfficientU-Net has more trainable parameters than U-ResNet, the total network size is only 60% of that of U-ResNet.

(4) EfficientU-Net-cGAN, which is combined with the cGAN module, is proposed to regulate training with high efficiency.

A nonpretrained symmetric U-ResNet with a structure similar to that used by Ref. [37] is employed in this study instead of the U-Net with pretrained ResNet as an encoder. The main reason for this choice is that ResNet was originally designed for image classification that contains several repeated residual blocks. However, when it comes to segmentation, a particularly deep and complex encoding process for a symmetric encoder-decoder structure would require a similarly complex decoding process. This means that a large amount of feature information is lost because of the many up-sampling layers in the decoding process. Therefore, the U-ResNet used herein requires short-skip connections between each block to preserve shallow image information. Moreover, long-skip connections, which link the encoder blocks to their equivalent decoder blocks and are used in U-Net, are also used to retain the shallow features of the input image. The U-ResNet architecture is shown in Fig. 3.

Based on the U-ResNet structure, an advanced U-ResNet architecture combined with a conditional adversarial network (U-ResNet-cGAN) is introduced. Conditional adversarial networks are designed mainly for image-to-image translation tasks. For the segmentation task, output segmented images are produced after the encoder-decoder structure, and then both the GT and segmented outputs are passed through a discriminator to determine whether

TABLE II. A comparison of the tested networks.

Networks	Total parameters	Total network size (MB)	Pretrained	Loss function	Advantages
U-ResNet	8 761 858	3107	No	Cross-entropy loss	Utilization of long-skip connection and short-skip connection
U-ResNet-cGAN	19 926 211	3233	No	Cross-entropy loss + binary-cross-entropy loss	cGAN module is used to further distinguish the output image and binary-cross-entropy loss added to regulate the training
EfficientU-Net	26 063 594	2043	Yes	Cross-entropy loss	Efficient feature extraction by balancing depth, width, and resolution
EfficientU-Net-cGAN	37 225 899	2169	Yes	Cross-entropy loss + binary-cross-entropy loss	Combines the advantages of cGAN module and EfficientNet

the image is real or fake. PatchGAN is used here because of its advantage of generating a fixed-size patch instead of a single number after the discriminator. By applying conditional adversarial networks, U-ResNet works as a generator and then competes with the discriminator during training. The detailed architecture of the U-ResNet-cGAN is shown in Fig. 4.

As discussed previously, for symmetric encoder-decoder structures—particularly for deep pretrained CNNs that are designed for image classification tasks—are inefficient during the decoding step. However, an encoder-decoder asymmetric structure that contains a deep complex network structure in the encoding step and a simple network structure in the decoding step could solve this issue and further improve feature extraction and identification. Therefore, in this study, EfficientNet-B3 [44], which is one architecture of EfficientNet (Effiecint-Net contains EfficientNet-B1 to B7) and achieves high accuracy and efficiency in image-classification tasks, is used as a feature extractor in the encoding step. Several transposed convolutional layers are used directly to increase the feature shape to the same size as the input image to avoid feature loss. Several long-skip connections are used to retain the shallow features of the input image. The network architecture of the EfficientU-Net is shown in Fig. 5. The core structure contains 25 blocks that use a structure similar to MobileNet [77]. The overall design concept is based on the utilization of inverted residual structures and residual blocks. A 1×1 convolution is used before the 3×3 or 5×5 network structure to increase the dimension, and an attention mechanism that assigns a weight to each feature of the image is added after the 3×3 or 5×5 network structure.

An advanced EfficientU-Net is then proposed by introducing a conditional adversarial network called EfficientU-Net-cGAN, and its architecture is shown in Fig. 6. The discriminator that distinguishes the output segmented

image from the GT has a similar structure to that of the U-ResNet-cGAN, as shown in Fig. 4.

All networks are trained for 40 epochs with an initial learning rate of 0.0001 using the ADAM solver. The input images are cropped to 454×454 with a batch size of 5. The learning rate is reduced by a factor of 0.5 when the loss reaches a plateau for 8 epochs. The loss function used to train all networks is the cross-entropy loss,

$$\text{cross entropy} = -\sum_{k=1}^{N}(p_k \log q_k), \qquad (4)$$

where p is the GT target in scalar, and q is the prediction after the softmax function. For the U-ResNet-cGAN and EfficientU-Net-cGAN, an extra binary-cross-entropy loss is applied to regulate the training of the discriminator,

$$\text{binary cross entropy} = -\sum_{k=1}^{N} w_k[p \log q_k \qquad (5)$$

$$+ (1-p) \log(1 - q_k)]. \quad (6)$$

The training data set contains 8800 degraded μCT images split into 7200 for training and 1600 for testing. The training is implemented on PyTorch, using an Nvidia RTX 3090 graphics-processing unit.

E. Physical accuracy measurements

Aside from pixel-wise accuracy, physical accuracy is also important to measure because in DRP the objective of segmentation is to isolate each mineral phase for subsequent physical analyses. In this study, physical accuracy is first evaluated by connectivity and porosity. Other nontrivial physical parameters that reveal the behavior of fluid displacement in rock are absolute permeability and relative permeability. Absolute permeability is sensitive to the porous structure, while relative permeability is
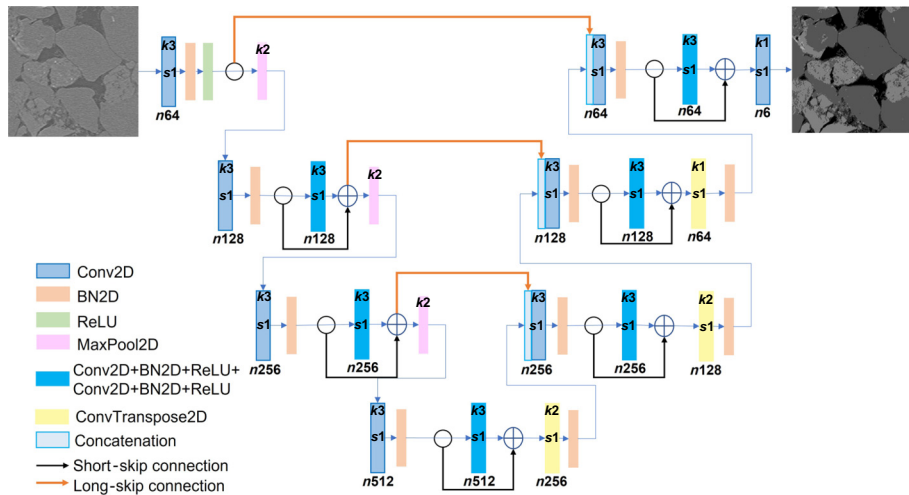

FIG. 3. Architecture of the U-ResNet, containing both a short-skip and long-skip connection. The input image contains three channels, and the output has six channels, indicating six different phases. BN2D refers to the 2D batch normalization, MaxPool2D refers to the process of selecting the maximum value of the feature map within each kernel. It is essential for useful feature extraction. ConvTranspose2D refers to the transposed 2D convolutional layer that used during decoding process.

sensitive to both the porous structure and mineral occurrence that are determined by multiphase segmentation [78]. These parameters are calculated based on a mixed-wetting condition that commonly occurs in reservoir rock and can be qualified by comparing the network output with the GT [58,79,80]. The Euler characteristic (χ) is used as an indication of connectivity, which is calculated as the difference between the number of loops and the number of disconnected pixels. The equation for calculating χ is

$$\chi = \text{objects} - \text{loops} + \text{holes}. \tag{7}$$

In addition, the volume fraction is determined for all phases, and is calculated by dividing the phase volume by the total volume; the pore-phase volume fraction is the porosity of the sample. The absolute permeability is calculated using a multirelaxation time lattice Boltzmann method (MRTLBM) preconditioned with a domain-decomposed Laplace solver [81–83], and the relative permeability is solved using the MorphLBM method [27].

MorphLBM utilizes a multiphase LBM simulation routine directly in the pore space of the image and performs morphological updates on the phases to emulate steady-state fluid configurations in an accelerated manner compared to directly simulating the coinjection of fluids, which requires significant simulation time to move through the domain. Fluid configurations are updated with small increments of erosion or dilation (depending on drainage or imbibition) to the target saturation. LBM simulation

is performed continuously as these small morphological increments are performed, and when the target saturation is reached, the LBM simulation is run until the capillary number converges. When the capillary number converges, the steady-state relative permeability point is recorded, and the morphological updates are started again. In this study, morphological imbibition is performed on the domains initialized by simulated primary drainage to residual saturation. The morphological updates are performed at a distance of 0.1 by interpolating the phase indicator value, and relaxation is performed for 1000 LBM time steps between morphs. The saturation increments are 5%, and the capillary number tolerance is set to less than 1×10^{-3} per 1000 time steps of the 50 000 time step exponential moving average, and the relative variance is less than 0.01. The system capillary number is maintained at 1×10^{-5} to emulate capillary-dominated two-phase flow. Mixed wettability is easily modeled using the LBM method (Color LBM) by simply assigning the contact angle as an affinity between -1 (water) and $+1$ (oil) for various solid voxels, where the static contact angle is equivalent to the inverse cosine of the affinity [84].

III. RESULTS AND DISCUSSION

A. Image degradation results

The raw μCT data is degraded by transferring fast-imaging-associated noise from the synchrotron-based μCT data, followed by the addition of Gaussian-Poisson noise.
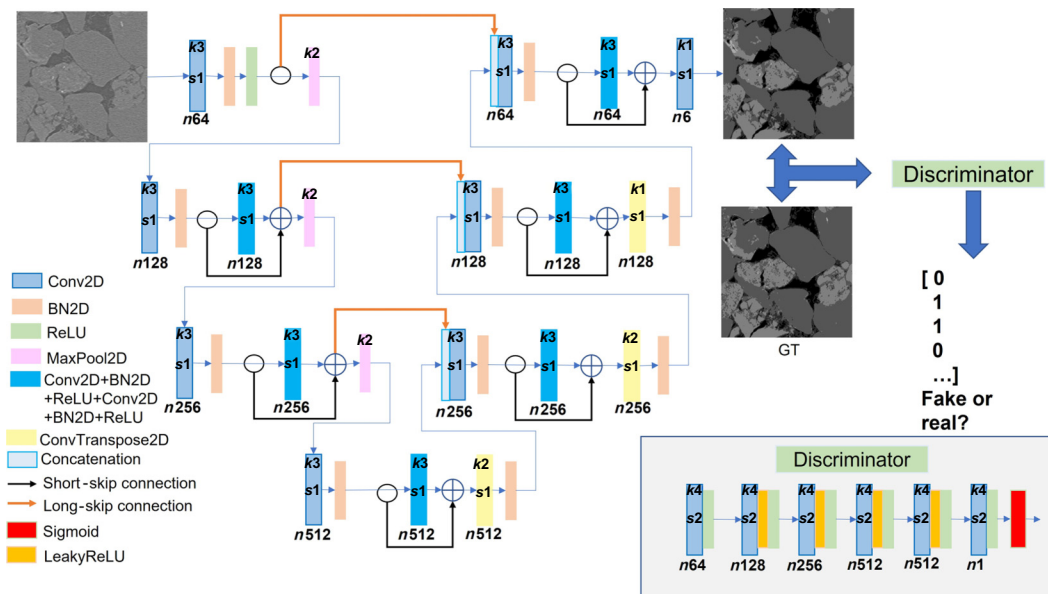

FIG. 4. Architecture of the U-ResNet-cGAN. The encoder-decoder structure is similar to the early U-ResNet structure. The output for the U-ResNet with the corresponding GT image is further passed to the discriminator. After application of the Sigmoid function, the output for the patch ranges from 0 to 1. LeakyReLU is aother activation function which is similar to the ReLU. The difference is that LeakyReLU outputs a non-zero slope for the negative input, while ReLU outputs zero forthe negative input. AveragePooling2D refers to the process of selecting the average value of the feature map as the output. It is essential for useful feature extraction.
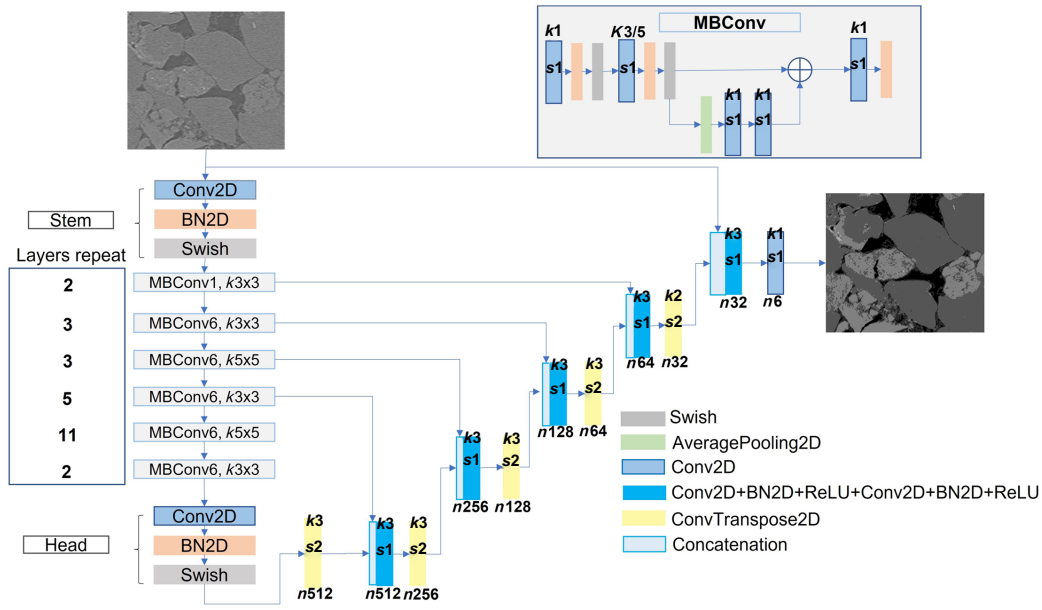
FIG. 5. Architecture of the EfficientU-Net-B3. The encoder contains a stem layer, 25 MBConv blocks, and head layer. The decoder contains five up-sampling layers with long-skip connections between encoding layers. Swish refers to swish activation function that helps alleviate the vanishing gradient during training for deeper network architecture. MBConv is the building block from MobileNet. It used as the attentional mechanism that guides the network to focus more on the important features.

This unpaired domain transfer is performed using Cycle-GAN. A sample output from CycleGAN is provided in Fig. 7, which includes the dynamic-styled μCT data and static-styled synchrotron data. Because this study mainly focuses on the segmentation of degraded images, only the domain transfer from raw μCT data to dynamic-styled μCT data is required. Furthermore, the output dynamic-styled μCT data is further degraded by adding Gaussian-Poisson noise, as shown in Fig. 7(e). Figure 7(f) shows an example of a GT image. To determine whether the degree of
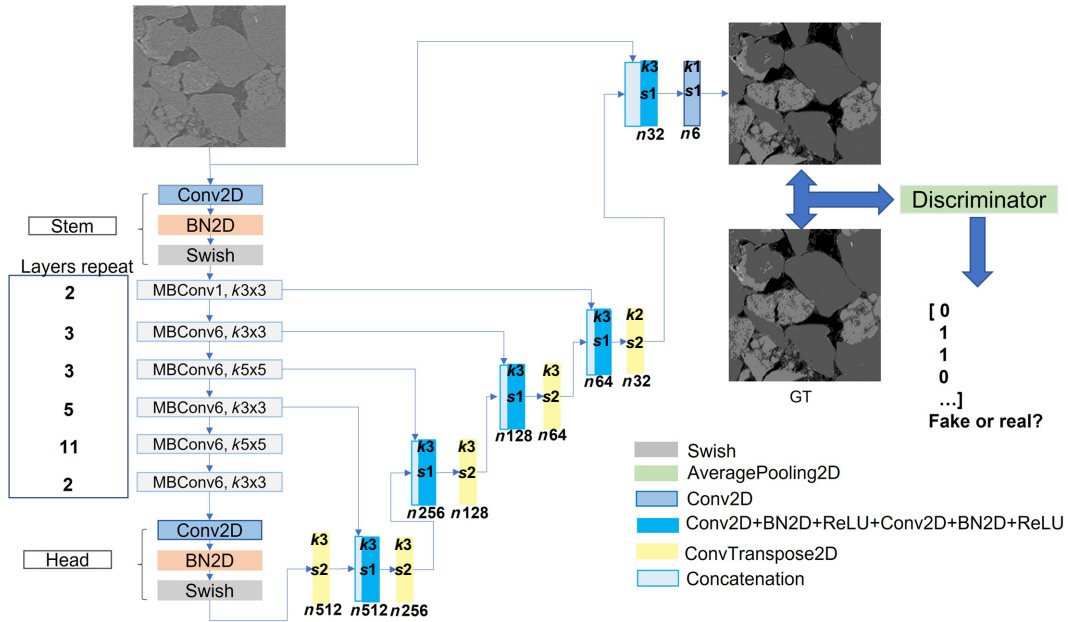

FIG. 6. Architecture of EfficientU-Net-cGAN. The encoder-decoder structure is identical to the previous EfficientU-Net-B3 structure. The addition of a discriminator would further force the output segmented image to appear similar to GT. The discriminator structure is identical to the discriminator depicted in Fig. 4.

noise added is realistic, two image-quality metrics are calculated using predictable statistical features to compute a quality score, BRISQUE and NIQE, which provide a qualitative measure of noise. These two methods do not require a paired reference image for image quality measurement, as required for the presented data, because domain transfer is performed between unpaired images. A detailed theoretical background of BRISQUE and NIQE can be found in Refs. [85,86]. The BRISQUE and NIQE metrics are compared between the raw μCT data, synchrotron-based μCT data, and degraded dynamic-styled μCT data, as well as two other existing dynamic-synchrotron-based μCT images that are accessible on the Digital Rocks Portal [87]. The first data set is the synchrotron-based μCT image of Ketton limestone [88] and the second is the synchrotron-based μCT image of Gildehauser sandstone [89]. The results of the BRISQUE and NIQE metrics for each data set are listed in Table III. For both metrics, a lower score reflects better image quality. As shown in Table III, the raw μCT data has the best image quality, while the other dynamic data sets are of lower quality.

For the degraded dynamic-styled μCT data, both metrics demonstrate that the image quality after domain transfer and degradation is within a reasonable range of commonly used dynamic synchrotron-based μCT data.

Comparing the regions in the red boxes in Figs. 7(c) and 7(e), it can be seen that the pixel value difference between feldspar and quartz is reduced after degradation, which could result in errors during segmentation. To evaluate the performance of CycleGAN, the voxel distribution of the 3D volume ($500 \times 500 \times 1200$) of raw μCT data, synchrotron-based μCT data, and dynamic-styled μCT data are compared, as shown in Fig. 8. The variances of the normalized images that show the differences in voxel distribution are then measured to validate the performance of the domain transfer. It should be noted that because both Mt. Simon sandstone and Bentheimer sandstone consist of a significant amount of quartz, the pixel distribution peaks (indicating the quartz phase) for the raw μCT data and synchrotron-based μCT data are at the same location. In addition, the results demonstrate that the raw μCT data and synchrotron-based μCT data have different
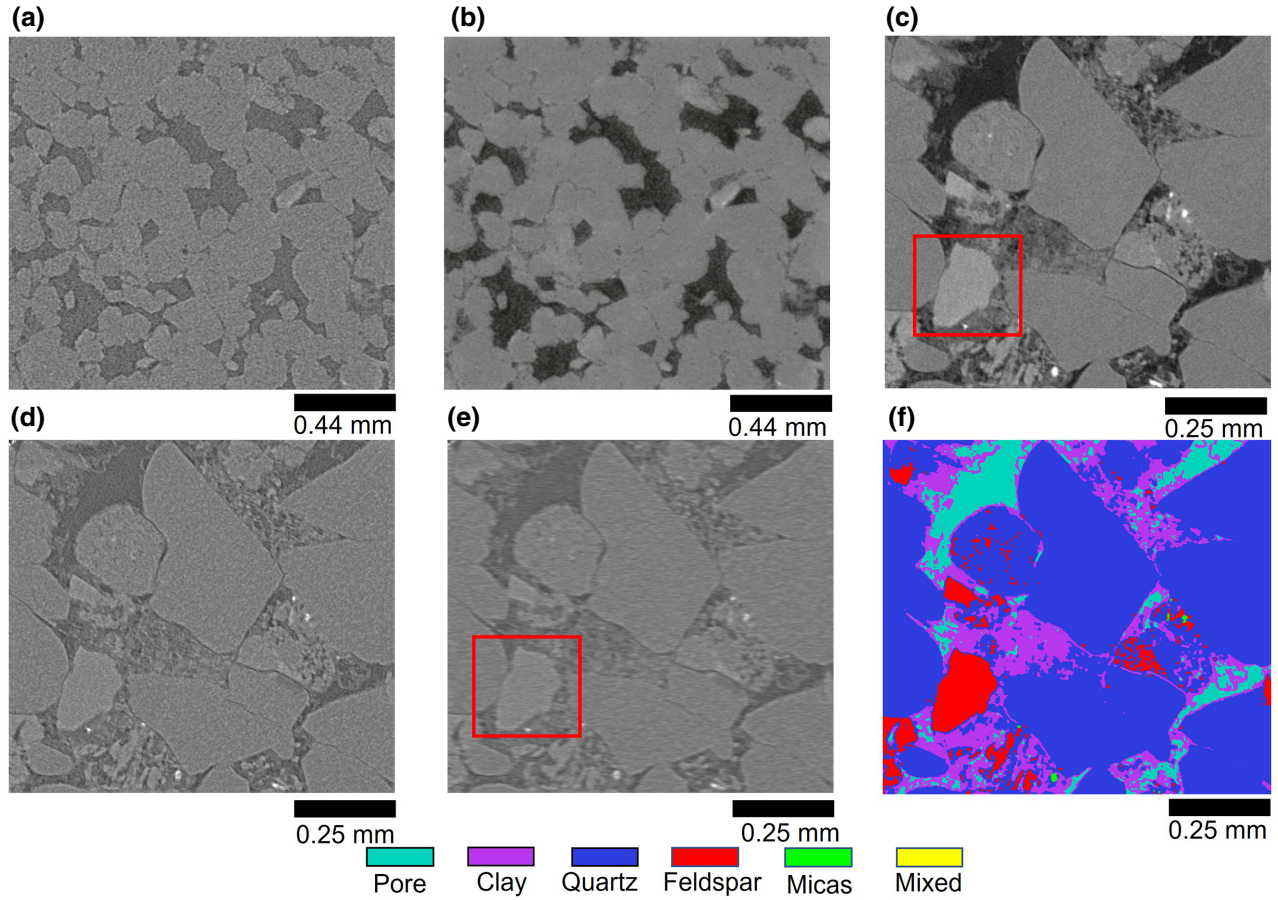

FIG. 7. (a) Sample image transferring from synchrotron-based μCT data to the (b) static-styled synchrotron data, and (c) raw μCT data to the (d) dynamic-styled μCT data. (e) Further degradation of dynamic-styled μCT data by the addition of Gaussian-Poisson noise. (f) Ground truth QEMSCAN slice of the corresponding μCT slice.

TABLE III. Image-quality measurement of five data sets. Both BRISQUE and NIQE are calculated slice by slice, and the average value of all slices is taken. The final degraded dynamic-styled μCT data have a similar image quality as the other three synchrotron-based μCT data sets.

	BRISQUE	NIQE
Raw μCT data	17.9	3.4
Synchrotron-based μCT data	42.3	11.0
Degraded dynamic-styled μCT data	43.4	8.9
Synchrotron-based Gildehauser sandstone μCT data	30.5	6.8
Synchrotron-based Ketton limestone μCT data	40.8	6.2

pixel distributions in a range lower than the pixel value of 100 and at the location where the pixel value is 255 (with a variance of 0.015). After applying domain transfer, the pixel value of the dynamic-styled μCT data is consistent with the synchrotron-based μCT data, with a variance of 0.002, indicating that CycleGAN is sufficient

for the domain transfer of synchrotron noise. The degraded dynamic-styled μCT data are subsequently used in the next step of multimineral segmentation.

B. Multimineral segmentation accuracy

All networks are trained on data containing dynamic-styled μCT data that are cropped into a domain of $454 \times 454 \times 7200$ voxels and test images of $454 \times 454 \times 1600$ voxels. The weighted accuracy, which considers the correctly labeled pixels as well as the volume fraction of each mineral, is utilized to evaluate each network. Phase accuracies are also calculated to evaluate the ability of the networks when dealing with different minerals. This is performed by averaging the phase accuracies of the last 10 epochs, where the testing accuracy curve of each network reaches a plateau. The testing accuracies are shown in Fig. 9, with the visualization of a region of interest in Fig. 11. The accuracy of all networks converges to approximately 0.94. The accuracy of EfficientU-Net and EfficientU-Net-cGAN is marginally better than that of U-ResNet and U-ResNet-cGAN, with a difference of approximately 0.1% in pixel-wise accuracy. Compared to the weighted accuracy used by Ref. [37], which included the μCT image as input, the accuracy is reduced by 3% as a result of degradation. Overall, all networks perform well, even with synchrotron and Gaussian-Poisson noises,
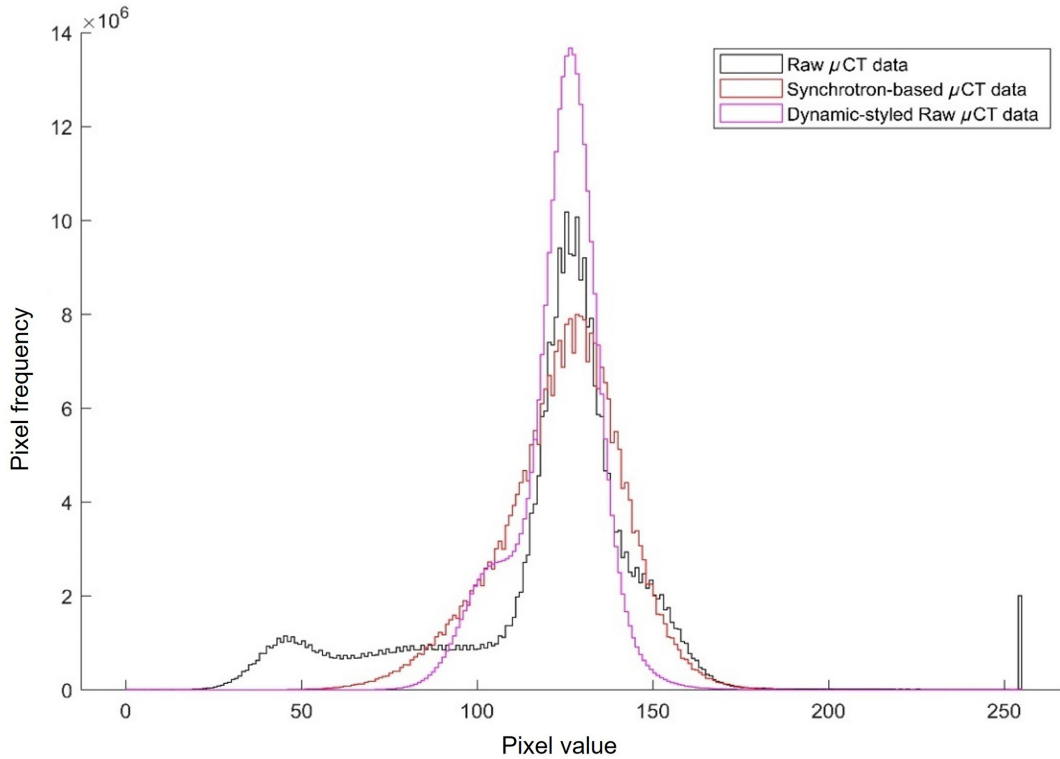

FIG. 8. Pixel value comparison between long-scan μCT images, synchrotron images, and the domain-transferred dynamic-styled Mt. Simon images before adding Gaussian-Poisson noise.
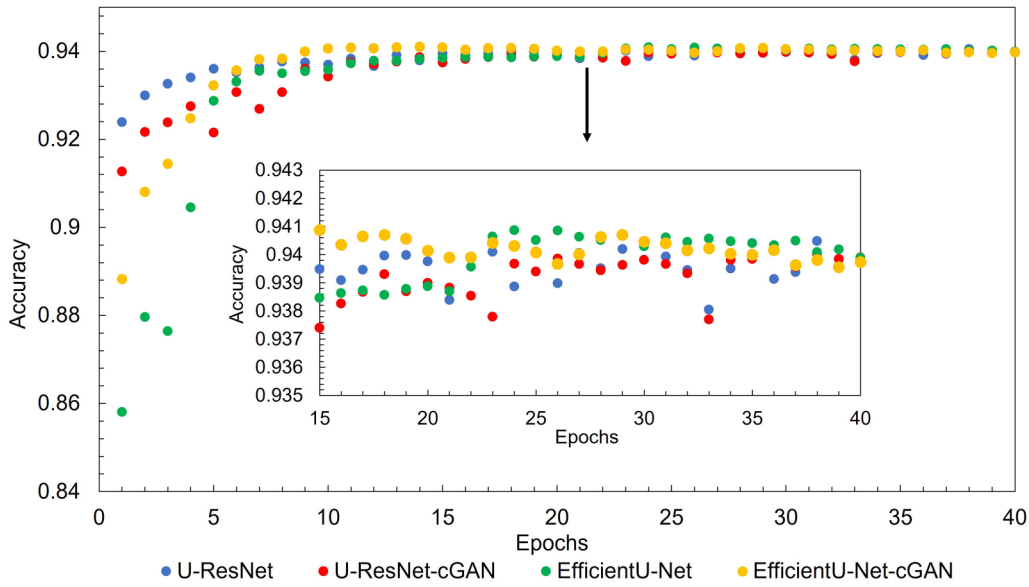
FIG. 9. Testing accuracy for four networks. The accuracy for all networks converges to around 0.94 after 15 epochs. An amplified curve shows that EfficientU-Net and EfficientU-Net-cGAN slightly outperform the other two networks.

which is also supported by the data given in Fig. 10. In this figure, all segmented outputs are visually similar to the GT slice, even for sparse minerals. For example, in the black box presented in Fig. 10, the clay and feldspar are sparsely distributed across the quartz; these sparse clay and feldspar phases are easily distinguishable as noise because the input synchrotron-styled Mt. Simon images already contain several types of noise across the entire pore region. However, the networks could correctly label these two minerals in the pore space (within the black box). Some errors occur in the feldspar (in the black box) due to the reduction of the pixel value difference between quartz and feldspar, as discussed in more detail in Sec. III A.

Because all networks perform similarly in terms of total accuracy, it is necessary to investigate how they perform for each mineral phase; therefore, the phase accuracy is calculated during each epoch. The average phase accuracy is obtained by averaging the phase accuracy of the last 10 epochs, as listed in Table IV. U-ResNet and U-ResNet-cGAN architectures achieve a higher pore-phase accuracy than the two EfficientNet-based architectures, whereas EfficientU-Net and EfficientU-Net-cGAN can better identify the clay phase. All networks perform best in the quartz phase, with an accuracy higher than 0.96, because all networks tend to learn well with a commonly occurring phase. EfficientU-Net-cGAN shows marginally higher accuracy in the feldspar and mica phases. In sparsely occurring micas and mixed-mineral phases, all networks result in low accuracy, with approximately 0.7 in micas and 0.5 in the mixed-mineral phase, which further indicates that the volume fraction of each phase is an essential factor that influences the CNN segmentation. Owing to the low phase-volume fraction, the quantity of

phase labels is imbalanced, and the networks tend to learn more from the commonly occurring phases, sacrificing accuracy in sparsely occurring phases. In general, all networks can achieve an accuracy in the 85%–97% range for the four most common phases, indicating that even for the μCT image with synchrotron-based noise and for a certain degree of degradation, CNN networks arguably provide accurate pixel-wise results for multimineral segmentation.

Furthermore, the Euclidean distance of the wrongly labeled pixels and the region-based error from an interface are calculated for each $530 \times 530 \times 1600$ voxel data set. This approach is used because the interface between phases is considered to be the most complicated region to segment [90]. It is not directly evaluated by pixel-wise accuracy because the majority of the pixels are internal to a given phase and are thus easy to segment. Considering that the objective of segmentation for DRP is to perform further physical analyses, the interface-region-based accuracy and the influence of segmentation on sequential pore characterization and pore-scale simulation should be thoroughly evaluated to show how well each network segments difficult regions, thus revealing the "true" phase structure [91,92]. Therefore, the data are prepared according to the following steps:

(1) Each phase's wrongly labeled pixel distribution is first determined by subtracting each network's output from the GT. Meanwhile, the Euclidean distance maps for each phase based on the GT images are generated.

(2) The incorrectly labeled pixel distribution is multiplied with the Euclidean distance map for the respective phase, which provides the Euclidean distance of the
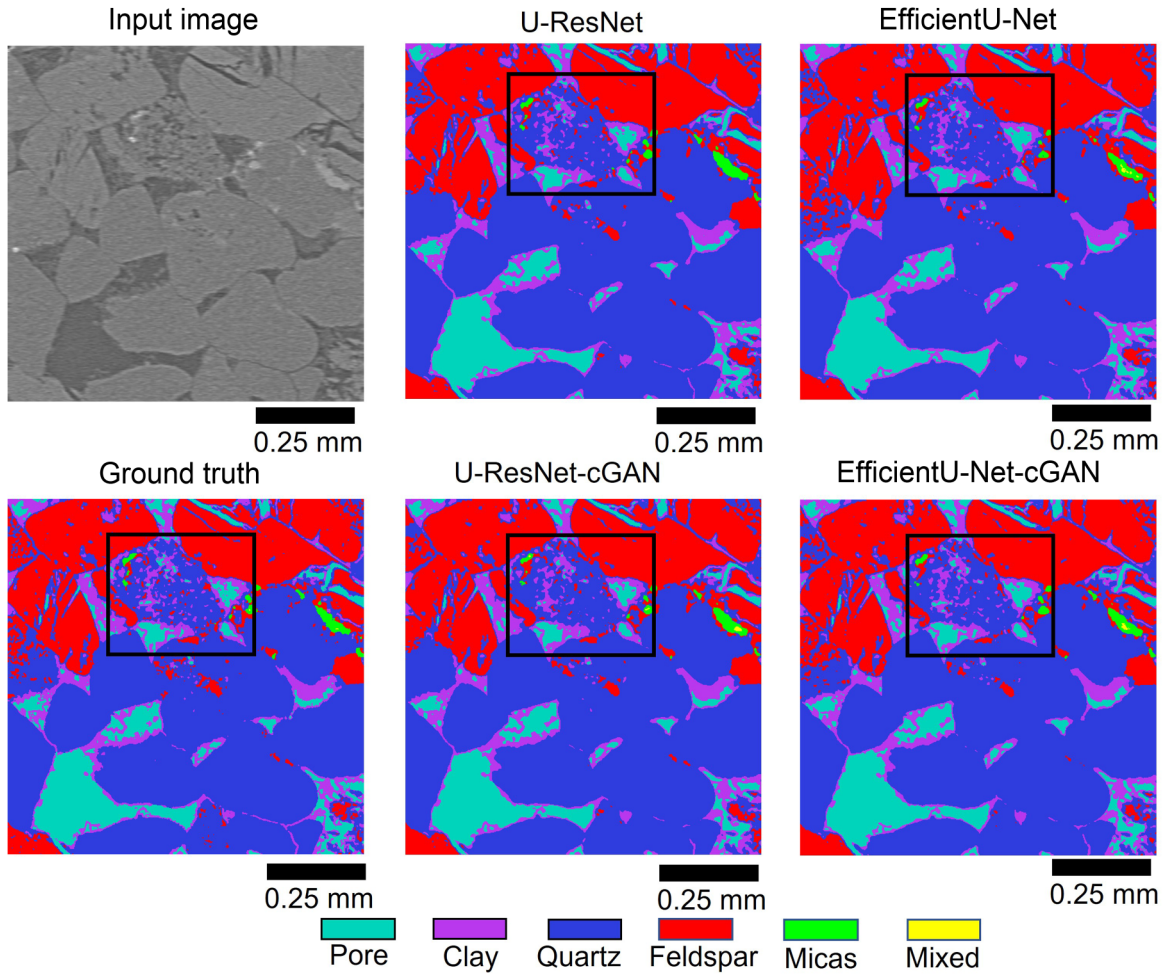
FIG. 10. A slice of the testing data set and output of four networks. All networks perform well visually; however, in the black box, all networks fail to capture the fine bodies of feldspar and clay that exist in quartz.

wrongly labeled pixels from the interface of the given phase.

(3) A histogram of the Euclidean distances for the GT data is generated to provide the total number of pixels that are located at a given distance from the interface.

(4) The region-based error is calculated by dividing the number of incorrectly labeled pixels that are located at a given distance by the total number of pixels in the GT that correspond to the given distance.

In Fig. 11, wrongly labeled pixels in all six phases mainly arise from the Euclidean distance within 1 pixel, indicating that most of the error for all networks is derived from the interface determination. EfficientU-Net-cGAN shows the lowest error near an interface in terms of pore and feldspar phases, indicating that EfficientU-Net-cGAN provides better segmentation at the boundary of these phases. Moreover, U-ResNet provides a better interface segmentation in the clay and mixed phases, whereas U-ResNet-cGAN can

TABLE IV. Average phase accuracy for identified minerals using four networks as well as the volume fraction of each phase. Overall, each network has its strength in different mineral phases.

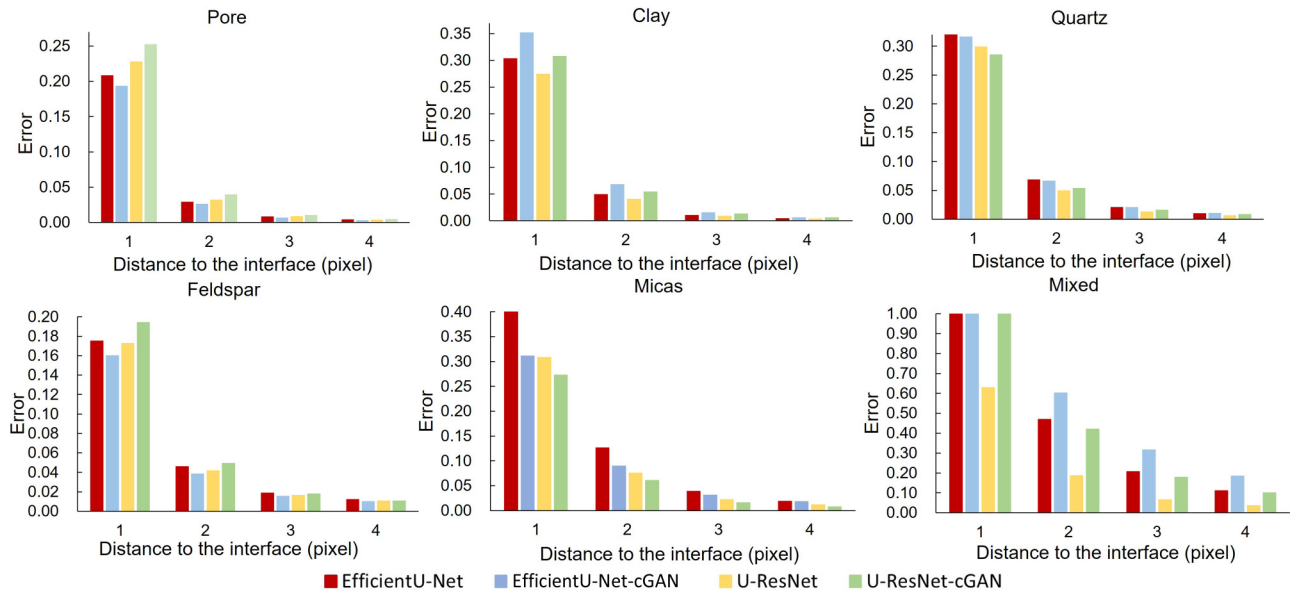|	Vol (%)	U-ResNet	U-ResNet-cGAN	EfficientU-Net	EfficientU-Net-cGAN
Pore	7.941	0.924	0.926	0.898	0.905
Clay	11.781	0.855	0.852	0.869	0.860
Quartz	63.185	0.963	0.966	0.965	0.963
Feldspar	16.460	0.901	0.891	0.895	0.904
Micas	0.293	0.694	0.708	0.675	0.713
Mixed	0.339	0.540	0.476	0.536	0.502

FIG. 11. Histograms of the region-based segmentation error of the wrongly labeled pixels in the interfaces of six phases for four networks. The main segmentation error arises from the pixels close to the phase interface.

handle the interface segmentation of the quartz and mica phases. It is noteworthy that not all networks can provide an accurate segmentation of the mixed phase near the interface, with the lowest region-based error being only 62%. This is mainly because the mixed phase has an extremely high pixel value in the gray-scale image in the raw μCT data, but in the synchrotron-based μCT data, it does not contain high-density mineral components. Therefore, after domain transfer, in the dynamic-styled μCT data, the gray-scale value between the mixed phase and other phases

decreases, making the segmentation of the mixed phase more difficult.

C. Physical accuracy measurement

In addition to the pixel-wise accuracy of each network, the physical accuracy of the segmented outputs is critical for the DRP. Connectivity, volume fraction, absolute permeability, and relative permeability are the physical parameters considered for a mixed-wet
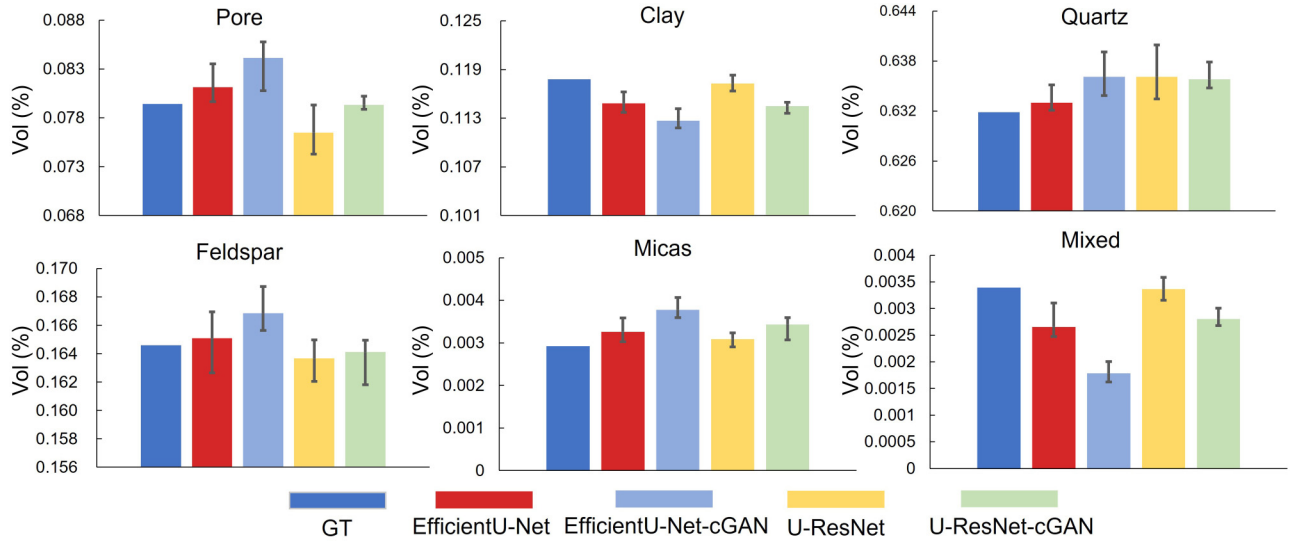

FIG. 12. Bar charts showing the volume fraction of each phase compared to the ground-truth result for four networks. Porosity is described by the pore volume fraction. Error bars represent the range of volume fraction calculated based on the top five most accurately trained epochs of each network.
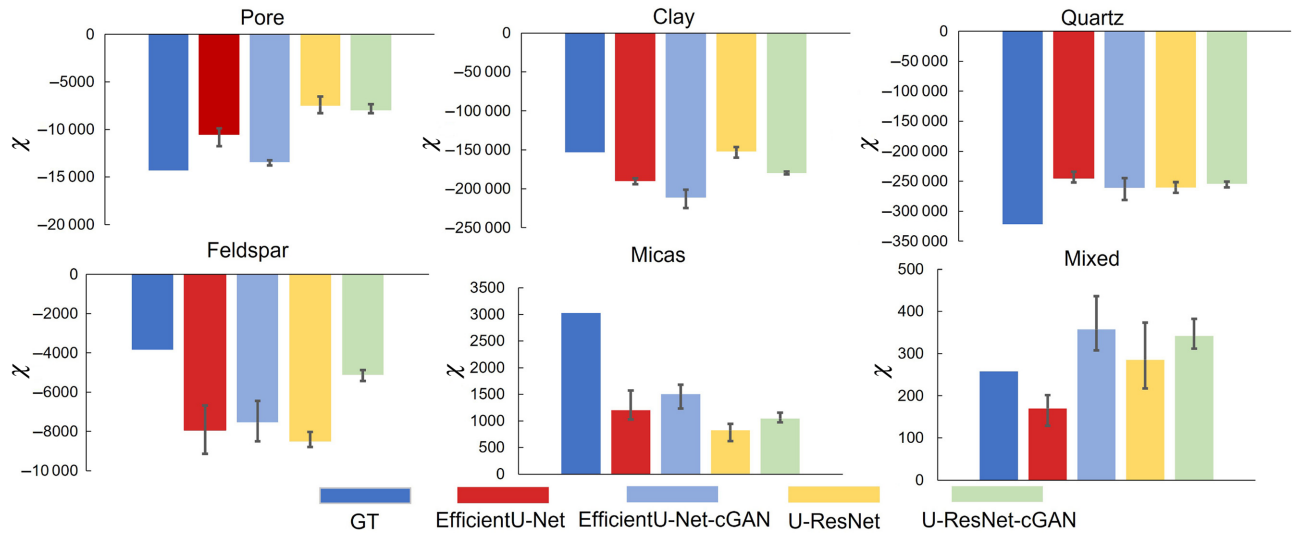
FIG. 13. Bar charts depicting the Euler characterization of each phase compared with the ground-truth result for four networks. The value of χ is calculated by averaging the χ of the network outputs with the top five pixel-wise accuracies. The error bars represent the range of χ calculated based on the top five most accurately trained epochs of each network.

condition. The physical accuracy is measured in the domain of $530 \times 530 \times 1600$ voxels. First, the volume fraction of each phase is calculated, as shown in Fig. 12. It can be observed that U-ResNet-cGAN performs best in the volume fraction of the pore phase, with a difference of less than 0.1% (i.e., porosity), whereas EfficientU-Net yields an accurate prediction for the quartz and feldspar phases, with differences of 0.2% and 0.3%, respectively. Moreover, U-ResNet performs best in clay, micas, and mixed phases, with differences of 0.4%, 0.6%, and 0.9%, respectively.

The connectivity, as described by the Euler number (χ) of each phase, is determined from the segmented images, where χ is a topological invariant defined by the number of objects, loops, and holes in a given phase [93]. The χ values for each phase are presented in Fig. 13. In terms of connectivity, it is found that none of the segmented data sets compare well with the GT data. In addition, the variability of the results across all networks is relatively high. The percentage differences of χ measured by each network in relation to the GT result are listed in Table V. Interface pixel errors and existing disconnected small bodies are likely to be the main reasons for the variations in χ.

Specifically with respect to χ, as shown in Eq. (6), a more positive value for a phase means that there are more isolated objects or fewer loops, indicating that the phase is less connected. On the other hand, a more negative value means that there are more loops than isolated objects, indicating that the phase is well connected. Therefore, it can be seen in Fig. 13 that in the pore phase, the χ value for GT is more negative than for any other network result. This suggests that all networks provide segmented images that are less connected than the GT. EfficientU-Net-cGAN provides the most accurate result in

terms of the pore phase, being only 6% less negative compared to the GT. For Mt. Simon sandstone, clay fills the pore space, and the less connected pore phase results in a more connected clay phase. Therefore, all networks yield a more negative value of χ in the clay phase. U-ResNet performs best in the clay phase, with only a 1% difference compared to the GT, while the other three networks produce a larger difference compared to GT, in the range of 17%–38%. The most accurate result for all networks is found for the χ of the quartz phase. It might be concluded that because quartz is the most abundant phase, the networks tend to train for it more often compared with the other phases, resulting in the highest physical accuracy in terms of connectivity. Moreover, U-ResNet-cGAN provides the best result in the feldspar phase, but the difference is still approximately 33% compared to GT, while all other networks provide a significantly negative value of χ compared to GT. In addition, in the mixed-mineral phase, the χ of U-ResNet shows a mere 10% difference compared to GT, while other networks yield a difference greater

TABLE V. Percentage differences of χ measured by each network compared with GT.

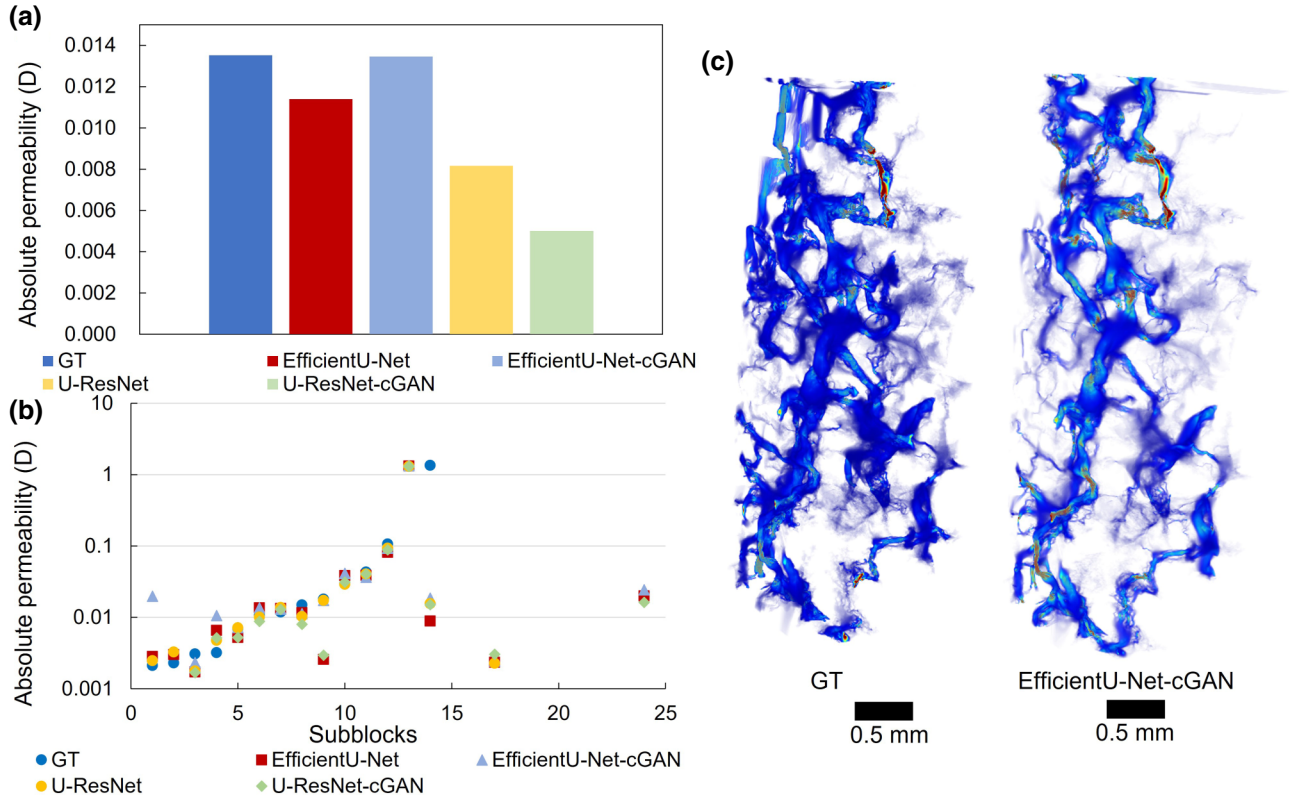|	EfficientU-Net (%)	EfficientU-Net-cGAN (%)	U-ResNet (%)	U-ResNet-cGAN (%)
Pore	27	6	48	45
Clay	24	38	1	17
Quartz	24	20	19	21
Feldspar	107	96	122	33
Micas	60	50	73	65
Mixed	34	39	11	32

FIG. 14.   (a) Absolute permeability comparison of each network's output domain, with the closest match achieved by EfficientU-Net-cGAN; EfficientU-Net yields the second most accurate prediction. (b) Absolute permeability comparison of subblocks. All networks show an accurate estimation in the subblocks. (c) Visualization of the velocity field of the GT and the best matched network output, as obtained from MRTLBM on a $530 \times 530 \times 1600$ voxel domain. The velocity field images are visually similar, but at some locations EfficientU-Net-cGAN results in a higher velocity due to the narrow flow path.

than 30%. In addition, all networks underestimate χ in the mica phase, with a difference greater than 50%. This is because the mica phase is rare and mainly consists of many small bodies; in many cases, the networks misidentify these small bodies, which results in fewer isolated objects.

The flow characteristics of segmented images are an essential measure of the physical accuracy. The absolute permeability of the pore phase is determined using single-phase flow simulation. After simulating the whole volume, it is further cropped into 24 subblocks with a domain size of 256^3 voxels. The absolute permeability is then calculated for each block, and the results are shown in Fig. 14. A close match is achieved by the EfficientU-Net-cGAN output in terms of the absolute permeability of the bulk volume, as shown in Fig. 14(a). A possible reason for this is that the EfficientU-Net-cGAN provides the closest match in terms of χ in the pore phase; the absolute permeability is known to be sensitive to the pore-phase connectivity [94]. To further test the absolute permeability, 24 subblocks are generated and simulated, as shown in Fig. 14(b). The mean square error (MSE) is reported in Table VI, using these subblocks for each network. It

is noted that the absolute permeability cannot be calculated in the case where the subblock is formed without a pore phase. All networks provide accurate absolute permeabilities in the majority of the subblocks, with EfficientU-Net-cGAN providing a marginally lower MSE value; this further confirms that the connectivity of these blocks is an essential parameter that affects the bulk absolute permeability. The accurate EfficientU-Net-cGAN result for the bulk absolute permeability is highly related to the matched pore-phase connectivity.

The same domain used for the absolute permeability simulation is used for the relative permeability simulation based on a mixed-wetting condition. The wettability of the

TABLE VI.   MSE results for the absolute permeability of these subblocks for each network. EfficientU-Net-cGAN gives a marginally lower MSE, which is consistent with the whole block absolute permeability result.

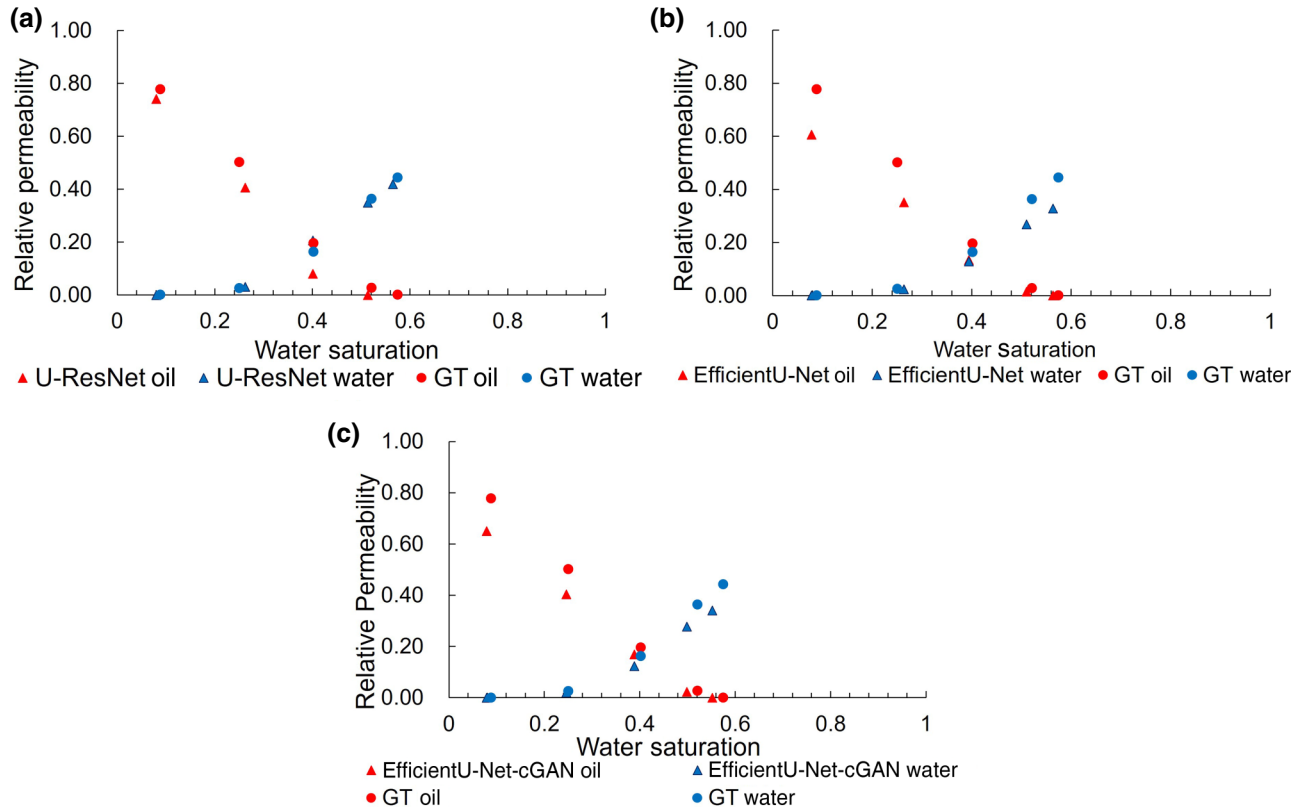|	U-ResNet	U-ResNet-cGAN	EfficientU-Net	EfficientU-Net-cGAN
MSE	0.1459	0.1460	0.1475	0.1453

FIG. 15. Relative permeability curves for (a) U-ResNet output with GT, (b) EfficientU-Net output with GT, and (c) EfficientU-Net-cGAN with GT, which have a similar shape. The U-ResNet output gives the best match for the endpoint oil and water relative permeability relative to GT, while EfficientU-Net gives the closest cross-point value and irreducible oil saturation to that of GT.

quartz phase varies from water wet (contact angle typically ranging from 0°–55°) to intermediate oil wet (contact angle typically ranging from 100°–140°) [95]. In the Mt. Simon sandstone sample, feldspar exists with quartz to form the grain, and the wettability of both the quartz and feldspar phases in the simulation is assigned as intermediate water wet, with contact angles of 45° and 60°, respectively [58]. For the clay phase, the wettability ranges from intermediate water wet to intermediate oil wet, which depends on the composition (e.g., kaolinite, illite, and montmorillonite) of the clay minerals [58,96,97]. In this simulation, we assign the clay phase as oil wet with a contact angle of 120° to provide a mixed-wetting condition. Micas and mixed phases are sparsely occurring phases, so they are set to have a contact angle of 0°. The relative permeability results for all data sets are shown in Fig. 15.

The MorphLBM method [27] first performs morphological initialization of primary drainage using a local distance maximum transform with hydraulic connectivity considered. Then, multiphase LBM simulations are performed until the system reaches a steady-state configuration for the given saturation. This is determined by tracking the system capillary number and identifying when the relative exponential moving average of relative permeability diverges by less than 1×10^{-3}, and the relative variance

is less than 0.01. Once this point is reached, the relative permeability values for the given saturation are recorded, and a negative morphological shell aggregation operation is performed until the fluid saturations reach a desired incremental change, that is, 5% saturation. Once this saturation is reached by morphological shell aggregation, LBM relaxes phase distributions and redistributes the phases among the pore space until a steady state is reestablished, as defined previously. An in-depth description of this method is available in Ref. [80]. The shape of the relative permeability of the GT domain simulated using mixed-wetting conditions resembles the curves obtained in a similar study of relative permeability in mixed-wetting Mt. Simon sandstones [58]. However, for the U-ResNet-cGAN output domain, the simulation does not converge because of the existence of an extremely narrow flow path. The extremely low absolute permeability of U-ResNet-cGAN is also due to the same reason. Therefore, only the relative permeability curves based on the output domains of the other three networks are reported.

The simulation begins with a water saturation of 0.1 and oil saturation of 0.9, ending when the two phases are hydraulically disconnected. The shape of the relative permeability curves for all simulation domains is similar to that of the GT domain. To further analyze the accuracy

TABLE VII.   Comparison of endpoint relative permeability value, irreducible oil saturation, and cross-point values between each network with the GT. The parameter k_{rwr} refers to water endpoint relative permeability, k_{ror} to oil endpoint relative permeability, s_{wcp} to the water saturation at the cross point, and s_{or} to irreducible oil saturation.

	GT	U-ResNet	EfficientU-Net	EfficientU-Net-cGAN
k_{rwr}	0.44	0.35	0.33	0.34
k_{ror}	0.78	0.74	0.61	0.65
s_{wcp}	0.40	0.34	0.39	0.39
s_{or}	0.57	0.51	0.56	0.55

of the relative permeability, endpoint relative permeability, irreducible saturation, and cross-point values are compared for each network's output with the GT results, as shown in Table VII. These are particularly important parameters for CCS and UHS applications with irreducible saturations, defining how much of a given phase is trapped in the rock and the cross point defining how saturation waves propagate through a reservoir [98]. The endpoint relative permeability for the GT domain is 0.44 for water and 0.78 for oil. The best performing network is U-ResNet, where the relative permeability value is 0.35 for the water endpoint and 0.74 for the oil endpoint. In addition, the cross-point relative permeability value for GT is approximately 0.4; the best performance is achieved by the domain of EfficientU-Net, with a cross-point value of 0.39. The most accurate result in terms of irreducible oil saturation is also produced by EfficientU-Net, with a value of 0.56 compared to 0.57 for the GT domain. It is worth noting that small clay features that are easily washed out during segmentation play a significant role in the relative permeability; therefore, this further stresses the importance of preserving fine regions, such as fine clay structures and mineral interfaces, during segmentation. Figure 16 visualizes the fluid-phase distributions, showing the fluid distributions at high and low oil saturations. The fluid distributions differ for each segmentation domain, which means that the fluid flow and displacement in the porous media is directly influenced by the multimineral phase segmentation. Overall, EfficientU-Net and U-ResNet are visually more like the GT than EfficientU-Net-cGAN.

To summarize the overall performance of the four tested networks, Table VIII includes all the metrics used to evaluate the network performance. Only four commonly
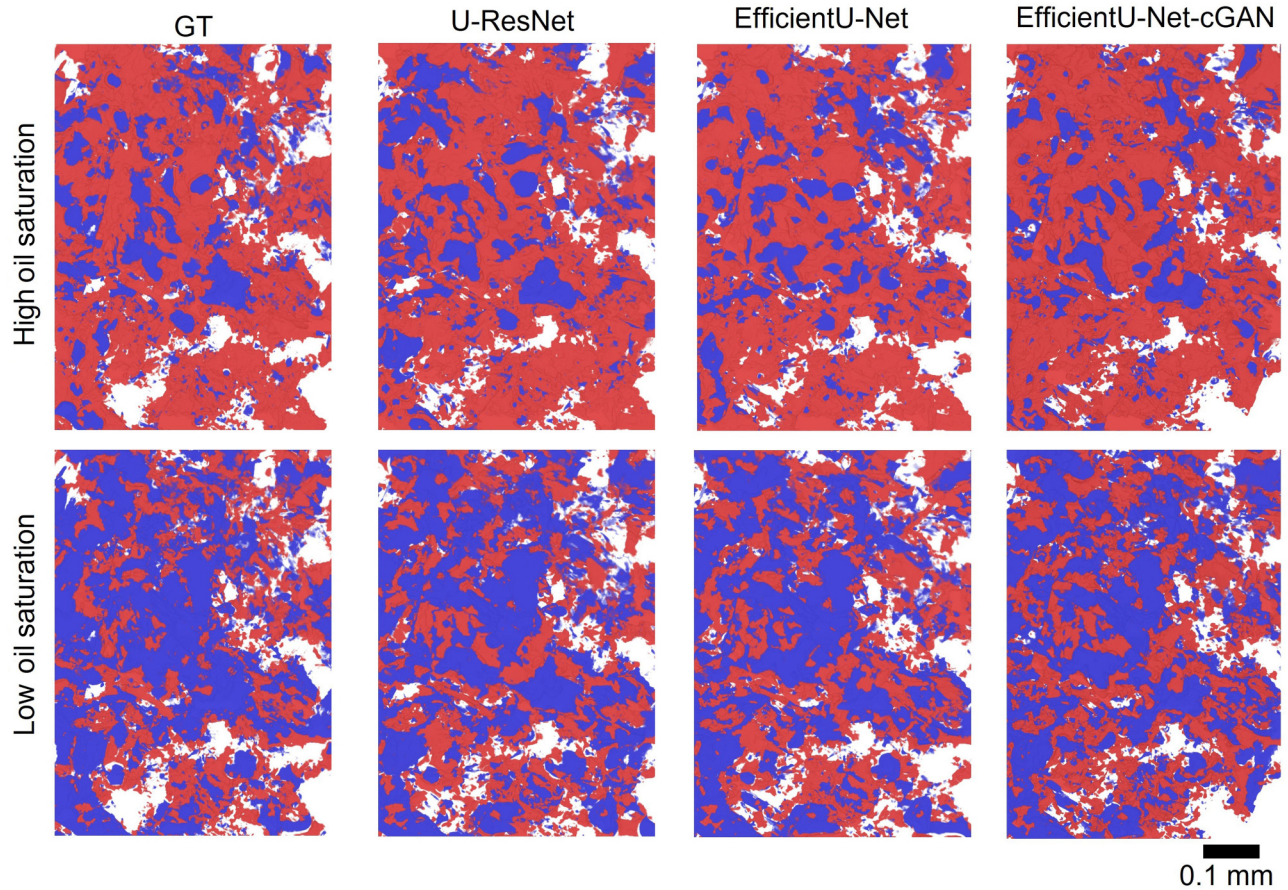

FIG. 16.   Visualization of the simulated domain from top to bottom at 1 001 000 LBM time steps (high oil saturation) and 2 484 000 LBM time steps (low oil saturation). The blue region corresponds to water and red to oil.

TABLE VIII. Summary of accuracy measurements for the four networks. The symbol ✓ refers to the best result in terms of the given metric, while ○ refers to the second-best result. For those networks that have only one acceptable result, only the best performance is ticked. EfficientU-Net-cGAN has the best overall performance, including five best performance metrics and three second-best metrics.

	Overall pixel accuracy	Phase pixel accuracy				Relative permeability
		Pore	Clay	Quartz	Feldspar	
U-ResNet		○		○	○	
U-ResNet-cGAN		✓		✓		
EfficientU-Net	✓		✓			✓
EfficientU-Net-cGAN	○			○	✓	

	Euler characteristic				Porosity
	Pore	Clay	Quartz	Feldspar	
U-ResNet		✓	✓		
U-ResNet-cGAN		○		✓	✓
EfficientU-Net	○				○
EfficientU-Net-cGAN	✓		○		

	Region-based accuracy				Absolute permeability
	Pore	Clay	Quartz	Feldspar	
U-ResNet		✓	○	○	
U-ResNet-cGAN			✓		
EfficientU-Net	○	○			
EfficientU-Net-cGAN	✓			✓	✓

occurring phases are considered here because both the mica phase and mixed phase are rare. All networks have their own best metrics, e.g., EfficientU-Net-cGAN has a higher number of accurate metrics. However, the network selection should be based on the application they are designed for. For example, to capture the flow behavior in a clay-coated sample (such as Mt. Simon sandstone), where the clay commonly exists in the pore, EfficientU-Net and U-ResNet might be more suitable; EfficientU-Net has a better clay-phase segmentation as well as relative permeability; and U-ResNet shows a better clay connectivity and clay interface determination. If the sample contains insignificant clay in the pore (Bentheimer sandstone), EfficientU-Net-cGAN is a better choice that gives a better segmentation of the pore phase. For mineral identification applications, where the flow behavior is not essential, EfficientU-Net-cGAN might be selected since it provides better segmentation results for minerals other than quartz. Additionally, if there is limited training data, to avoid overfitting, U-ResNet might be the best option, because the architecture is simpler than other three and has fewer training parameters (Table II). When considering these metrics, it should also be considered that consistency in the processing workflow is of utmost importance, which is an attribute that all networks provide.

IV. CONCLUSION

To perform multimineral segmentation on dynamic synchrotron-based images with a certain degree of noise associated with fast imaging, unpaired domain transfer is implemented using CycleGAN. It performs unpaired domain transfers of synchrotron-based μCT data into raw μCT data with ground-truth labels. Dynamic-styled μCT data are further degraded by adding Poisson and Gaussian noises, which are the common noise in μCT imaging. Four deep CNNs are used to segment the degraded synchrotron-style Mt. Simon sandstone images. The pixel-wise accuracy for all networks converges to approximately 94%. The accuracy for EfficientU-Net and EfficientU-Net-cGAN is marginally better than that of U-ResNet and U-ResNet-cGAN. Physical accuracy is also determined to further compare the segmentation results of each network. The physical measurements of connectivity show higher variance, especially in the less common phases. All networks provide an accurate prediction in the volume-fraction measurement, with a maximum difference of less than 5%. EfficientU-Net-cGAN provides an accurate measurement of absolute permeability and yields the best performance for interface segmentation, whereas EfficientU-Net provides an accurate prediction in terms of relative permeability simulations for mixed-wetting conditions. From the high pixel-wise and physical accuracy, we demonstrate that the unpaired domain transfer by Cycle-GAN can capture the semantic or style from an image and transfer into another image. It is helpful in reducing the quantity requirement of ground truth for semantic segmentation tasks.

With the integration of the dynamic-based image-processing workflow of unpaired domain transfer and CNN methods, this research presents an application of real-time imaging and DRP for pore-scale CCS and UHS investigations. More specifically, the proposed image-processing workflow performs multimineral segmentation on a real-time image by transferring dynamic information from synchrotron-based scanning of a rock sample to a long-scanned high-quality rock image without the requirement of real-time ground-truth data. DRP with multiple phases can then be implemented with the validation of multiphase-flow experimental data generated during dynamic image scanning. The workflow is generalizable to studying any type of porous media with multiphases, flow, or transport. For example, it could be applied to H_2 diffusion experiments with dynamic scanning in order to understand the relationship between diffusion and rock mineralogy, as well as to mineral dissolution and precipitation or the development of gas pockets within a hydrogen fuel cell that reduces overall transport efficiencies [99].

In addition to the specific case described in this study, the workflow serves as a semiautomatic process for the image processing of porous materials. Traditionally, the μCT image is first subjected to several preprocessing steps such as filtering and pixel matching before segmentation, which requires a great amount of human effort, including human biases. By using the proposed workflow, automatic "filtering" and "pixel matching" is implemented by unpaired domain transfer with CycleGAN, and this is the major objective and contribution of the study. Domain transfer is then followed by an automatic multiphase segmentation without any human effort or bias. The objective is that training data and learning mappings from previous works can be applied to unseen works from completely different instruments and settings. Thus, the methodology allows for applying trained algorithms to a broad range of data in a consistent and objective way. A step toward the development of a fully automatic workflow is to couple the CycleGAN with the CNN segmentation networks into a single network; this requires a more complex network structure and hyperparameter tuning to balance the loss functions along with a means to dynamically adjust the learning rate. Meanwhile, additional image information could also help to increase the segmentation performance. This could be different image modalities or the same modality with different settings, such as dual energy images or phase contrast images. The overall workflow provides a digital material platform for the study of physical processes within complex porous structures containing multiphases that can deal with the noise associated with dynamic real-time imaging.

All the data that support the findings of this study are available from the corresponding author upon reasonable request.

[1] R. S. Dimitrov, The Paris agreement on climate change: Behind closed doors, Glob. Environ. Politics **16**, 1 (2016).

[2] J. Rogelj, M. den Elzen, N. Höhne, T. Fransen, H. Fekete, H. Winkler, R. Schaeffer, F. Sha, K. Riahi, and M. Meinshausen, Paris agreement climate proposals need a boost to keep warming well below 2 °C, Nature **534**, 7609 (2016).

[3] S. N. Seo, Beyond the Paris agreement: Climate change policy negotiations and future directions, Reg. Sci. Policy Pract. **9**, 121 (2017).

[4] P. O. Carden and L. Paterson, Physical, chemical and energy aspects of underground hydrogen storage, Int. J. Hydrog. Energy **4**, 559 (1979).

[5] K. Damen, A. Faaij, and W. Turkenburg, Health, safety and environmental risks of underground Co_2 storage – overview of mechanisms and current knowledge, Clim. Change **74**, 289 (2006).

[6] S. L'Orange Seigo, S. Dohle, and M. Siegrist, Public perception of carbon capture and storage (CCS): A review, Renewable Sustainable Energy Rev. **38**, 848 (2014).

[7] R. Tarkowski, Underground hydrogen storage: Characteristics and prospects, Renewable Sustainable Energy Rev. **105**, 86 (2019).

[8] A. Alhosani, Q. Lin, A. Scanziani, E. Andrews, K. Zhang, B. Bijeljic, and M. J. Blunt, Pore-Scale characterization of carbon dioxide storage at immiscible and near-miscible conditions in altered-wettability reservoir rocks, Int. J. Greenh. Gas Control **105**, 103232 (2021).

[9] S. Krevor, M. J. Blunt, S. M. Benson, C. H. Pentland, C. Reynolds, A. Al-Menhali, and B. Niu, Capillary trapping for geologic carbon dioxide storage – from pore scale physics to field scale implications, Int. J. Greenh. Gas Control **40**, 221 (2015).

[10] S. Flesch, D. Pudlo, D. Albrecht, A. Jacob, and F. Enzmann, Hydrogen underground storage—petrographic and petrophysical variations in reservoir sandstones from laboratory experiments under simulated reservoir conditions, Int. J. Hydrog. Energy **43**, 20822 (2018).

[11] D. Zivar, S. Kumar, and J. Foroozesh, Underground hydrogen storage: A comprehensive review, Int. J. Hydrog. Energy **46**, 23436 (2021).

[12] Z. Bo, L. Zeng, Y. Chen, and Q. Xie, Geochemical reactions-induced hydrogen loss during underground hydrogen storage in sandstone reservoirs, Int. J. Hydrog. Energy **46**, 19998 (2021).

[13] M. J. Blunt, B. Bijeljic, H. Dong, O. Gharbi, S. Iglauer, P. Mostaghimi, A. Paluszny, and C. Pentland, Pore-Scale

imaging and modelling, Adv. Water Resour. **51,** 197 (2013).

[14] P. Mostaghimi, R. T. Armstrong, A. Gerami, Y. Hu, Y. Jing, F. Kamali, M. Liu, Z. Liu, X. Lu, H. L. Ramandi, A. Zamani, and Y. Zhang, Cleat-Scale characterisation of coal: An overview, J. Nat. Gas Sci. Eng. **39,** 143 (2017).

[15] Y. D. Wang, M. J. Blunt, R. T. Armstrong, and P. Mostaghimi, Deep learning in pore scale imaging and modeling, Earth Sci. Rev. **215,** 103555 (2021).

[16] A. Kääb and M. Vollmer, Surface geometry, thickness changes and flow fields on creeping mountain permafrost: Automatic extraction by digital image analysis, Permafr. Periglac. Process. **11,** 315 (2000).

[17] M. A. Knackstedt, S. Latham, M. Madadi, A. Sheppard, T. Varslot, and C. Arns, Digital rock physics: 3D imaging of core material and correlations to acoustic and flow properties, Lead. Edge **28,** 28 (2009).

[18] H. Andrä, N. Combaret, J. Dvorkin, E. Glatt, J. Han, M. Kabel, Y. Keehm, F. Krzikalla, M. Lee, C. Madonna, M. Marsh, T. Mukerji, E. H. Saenger, R. Sain, N. Saxena, S. Ricker, A. Wiegmann, and X. Zhan, Digital rock physics benchmarks—part I: Imaging and segmentation, Comput. Geosci. **50,** 25 (2013).

[19] I. Verri, A. Della Torre, G. Montenegro, A. Onorati, S. Duca, C. A. Mora, F. Radaelli, and G. Trombin, Development of a digital rock physics workflow for the analysis of sandstones and tight rocks, J. Pet. Sci. Eng. **156,** 790 (2017).

[20] S. Berg, H. Ott, S. A. Klapp, A. Schwing, R. Neiteler, N. Brussee, A. Makurat, L. Leu, F. Enzmann, J.-O. Schwarz, M. Kersten, S. Irvine, and M. Stampanoni, Real-Time 3D imaging of haines jumps in porous media flow, Proc. Natl. Acad. Sci. **110,** 3755 (2013).

[21] 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 Resour. **95,** 341 (2016).

[22] G. Dalen and M. Koster, *2D & 3D Particle Size Analysis of Micro-CT Images* (The Netherlands, 2012), p. 16.

[23] Y. Jing, R. T. Armstrong, and P. Mostaghimi, Image-Based fracture pipe network modelling for prediction of coal permeability, Fuel **270,** 117447 (2020).

[24] S. Karimpouli, P. Tahmasebi, H. L. Ramandi, P. Mostaghimi, and M. Saadatfar, Stochastic modeling of coal fracture network by direct Use of micro-computed tomography images, Int. J. Coal Geol. **179,** 153 (2017).

[25] F. Reyes, Q. Lin, O. Udoudo, C. Dodds, P. D. Lee, and S. J. Neethling, Calibrated X-Ray micro-tomography for mineral Ore quantification, Miner. Eng. **110,** 122 (2017).

[26] M. Liu, M. Shabaninejad, and P. Mostaghimi, Predictions of permeability, surface area and average dissolution rate during reactive transport in multi-mineral rocks, J. Pet. Sci. Eng. **170,** 130 (2018).

[27] 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).

[28] T. Akai, A. M. Alhammadi, M. J. Blunt, and B. Bijeljic, Modeling Oil recovery in mixed-Wet rocks: Pore-scale comparison between experiment and simulation, Transp. Porous Media **127,** 393 (2019).

[29] C. H. Arns, M. A. Knackstedt, W. V. Pinczewski, and K. R. Mecke, Euler-Poincaré characteristics of classes of disordered media, Phys. Rev. E **63,** 031112 (2001).

[30] C. Madonna, B. S. G. Almqvist, and E. H. Saenger, Digital rock physics: Numerical prediction of pressure-dependent ultrasonic velocities using micro-CT imaging, Geophys. J. Int. **189,** 1475 (2012).

[31] A. Narváez, T. Zauner, F. Raischel, R. Hilfer, and J. Harting, Quantitative analysis of numerical estimates for the permeability of porous media from lattice-boltzmann simulations, J. Stat. Mech. **2010,** P11026 (2010).

[32] Y. A. Alzahid, H. Aborshaid, M. Asali, J. McClure, C. Chen, P. Mostaghimi, Y. Da Wang, C. Sun, and R. T. Armstrong, Real-Time synchrotron-based X-Ray computed microtomography during in situ emulsification, J. Pet. Sci. Eng. **195,** 107885 (2020).

[33] D. Wildenschild and A. P. Sheppard, X-Ray imaging and analysis techniques for quantifying pore-scale structure and processes in subsurface porous medium systems, Adv. Water Resour. **51,** 217 (2013).

[34] R. T. Armstrong, N. Evseev, D. Koroteev, and S. Berg, Modeling the velocity field during haines jumps in porous media, Adv. Water Resour. **77,** 57 (2015).

[35] C. Madonna, B. Quintal, M. Frehner, B. S. G. Almqvist, N. Tisato, M. Pistone, F. Marone, and E. H. Saenger, Synchrotron-Based X-Ray tomographic microscopy for rock physics investigations, Geophysics **78,** D53 (2013).

[36] M. Heyndrickx, T. Bultreys, W. Goethals, L. Van Hoorebeke, and M. N. Boone, Improving image quality in fast, time-resolved micro-CT by weighted back projection, Sci. Rep. **10,** 1 (2020).

[37] Y. Da Wang, M. Shabaninejad, R. T. Armstrong, and P. Mostaghimi, Physical Accuracy of Deep Neural Networks for 2D and 3D Multi-Mineral Segmentation of Rock Micro-CT Images, ArXiv:2002.05322 [Cs, Eess] (2020).

[38] S. Karimpouli and P. Tahmasebi, Segmentation of digital rock images using deep convolutional autoencoder networks, Comput. Geosci. **126,** 142 (2019).

[39] Y. Niu, P. Mostaghimi, M. Shabaninejad, P. Swietojanski, and R. T. Armstrong, Digital rock segmentation for petrophysical analysis With reduced user bias using convolutional neural networks, Water Resour. Res. **56,** e2019WR026597 (2020).

[40] O. Sudakov, E. Burnaev, and D. Koroteev, Driving digital rock towards machine learning: Predicting permeability with gradient boosting and deep neural networks, Comput. Geosci. **127,** 91 (2019).

[41] D. Tang and K. Spikes, in SEG Technical Program Expanded Abstracts 2017 (Society of Exploration Geophysicists, 2017), pp. 3898–3902.

[42] R. Girshick, J. Donahue, T. Darrell, and J. Malik, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (Columbus, 2014), p. 580.

[43] K. He, X. Zhang, S. Ren, and J. Sun, in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (Boston, 2015), p. 770.

[44] M. Tan and Q. V. Le, EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks, ArXiv:1905.11946 [Cs, Stat] (2020).

[45] C. Dong, C. C. Loy, K. He, and X. Tang, Image Super-Resolution Using Deep Convolutional Networks, ArXiv:1501.00092 [Cs] (2015).

[46] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, Image-to-Image Translation with Conditional Adversarial Networks, ArXiv:1611.07004 [Cs] (2018).

[47] K. He, G. Gkioxari, P. Dollár, and R. Girshick, Mask R-CNN, ArXiv:1703.06870 [Cs] (2018).

[48] O. Ronneberger, P. Fischer, and T. Brox, U-Net: Convolutional Networks for Biomedical Image Segmentation, ArXiv:1505.04597 [Cs] (2015).

[49] H. P. Narkhede, Review of image segmentation techniques, Int. J. Sci. Mod. Eng. **1,** 54 (2013).

[50] J. Kuruvilla, D. Sukumaran, A. Sankar, and S. P. Joy, in 2016 International Conference on Data Mining and Advanced Computing (SAPIENCE) (2016), pp. 198–203.

[51] S. Rubaiya Muin, P. T. Spicer, K. Tang, Y. Niu, M. Hosseini, P. Mostaghimi, and R. T. Armstrong, Dynamic X-Ray micotomography of microfibrous cellulose liquid foams using deep learning, Chem. Eng. Sci. **248,** 117173 (2022).

[52] H. Zhu, F. Meng, J. Cai, and S. Lu, Beyond pixels: A comprehensive survey from bottom-up to semantic image segmentation and cosegmentation, J. Vis. Commun. Image Represent. **34,** 12 (2016).

[53] T. Zhou, S. Ruan, and S. Canu, A review: Deep learning for medical image segmentation using multi-modality fusion, Array **3–4,** 100004 (2019).

[54] S. Hao, Y. Zhou, and Y. Guo, A brief survey on semantic segmentation with deep learning, Neurocomputing **406,** 302 (2020).

[55] P. Gravel, G. Beaudoin, and J. A. DeGuise, A method for modeling noise in medical images, IEEE Trans. Med. Imaging **23,** 1221 (2004).

[56] I. Ar Rushood, N. Alqahtani, Y. D. Wang, M. Shabaninejad, R. Armstrong, and P. Mostaghimi, in *SPE Annual Technical Conference and Exhibition* (Society of Petroleum Engineers, Colorado, 2020).

[57] N. Srisutthiyakorn, in SEG Technical Program Expanded Abstracts 2016 (Society of Exploration Geophysicists, 2016), pp. 3042–3046.

[58] M. Fan, J. E. McClure, R. T. Armstrong, M. Shabaninejad, L. E. Dalton, D. Crandall, and C. Chen, Influence of clay wettability alteration on relative permeability, Geophys. Res. Lett. **47,** e2020GL088545 (2020).

[59] M. Liu, M. Shabaninejad, and P. Mostaghimi, Impact of mineralogical heterogeneity on reactive transport modelling, Comput. Geosci. **104,** 12 (2017).

[60] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, Focal Loss for Dense Object Detection, ArXiv:1708.02002 [Cs] (2018).

[61] V. Davidoiu, L. Hadjilucas, I. I, N. P. Smith, J. E. Schneider, and J. Lee, Evaluation of noise removal algorithms for imaging and reconstruction of vascular networks using micro-CT, Biomed. Phys. Eng. Express **2,** 045015 (2016).

[62] M. Tuller, R. Kulkarni, and W. Fink, in *Soil-Water-Root Processes: Advances in Tomography and Imaging* (2013), Vol. 61, pp.157–182.

[63] J. Wang, H. Lu, Z. Liang, D. Eremina, G. Zhang, S. Wang, J. Chen, and J. Manzione, An experimental study on the noise properties of X-Ray CT sinogram data in radon space, Phys. Med. Biol. **53,** 3327 (2008).

[64] H. Zhang, D. Zeng, H. Zhang, J. Wang, Z. Liang, and J. Ma, Applications of nonlocal means algorithm in Low-dose X-Ray CT image processing and reconstruction: A review, Med. Phys. **44,** 1168 (2017).

[65] P. I. Guntoro, Y. Ghorbani, P.-H. Koch, and J. Rosenkranz, X-Ray microcomputed tomography (μCT) for mineral characterization: A review of data analysis methods, Minerals **9,** 3 (2019).

[66] M. Balafar, Review of noise reducing algorithms for brain MRI images, Int. J. Tech. Phys. Probl. Eng. **4,** 54 (2012).

[67] Y. Wang, S. Basu, and C.-Y. Wang, Modeling Two-phase flow in PEM fuel cell channels, J. Power Sources **179,** 603 (2008).

[68] A. Anagnostopoulos, S. Knauer, Y. Ding, and Y. Grosu, Giant effect of negative compressibility in a water–porous metal–CO_2 system for sensing applications, ACS Appl. Mater. Interfaces **12,** 39756 (2020).

[69] A. Lowe, N. Tsyrin, M. Chorążewski, P. Zajdel, M. Mierzwa, J. B. Leão, M. Bleuel, T. Feng, D. Luo, M. Li, D. Li, V. Stoudenets, S. Pawlus, A. Faik, and Y. Grosu, Effect of flexibility and nanotriboelectrification on the dynamic reversibility of water intrusion into nanopores: Pressure-transmitting fluid with frequency-dependent dissipation capability, ACS Appl. Mater. Interfaces **11,** 40842 (2019).

[70] Y. Grosu, Y. Zhao, A. Giacomello, S. Meloni, J.-L. Dauvergne, A. Nikulin, E. Palomo, Y. Ding, and A. Faik, Hierarchical macro-nanoporous metals for leakage-free high-thermal conductivity shape-stabilized phase change materials, Appl. Energy **269,** 115088 (2020).

[71] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (ICCV)* (Venice, 2017), p. 2223.

[72] J. E. Elkhoury, R. Shankar, and T. S. Ramakrishnan, Resolution and limitations of X-Ray micro-CT with applications to sandstones and limestones, Transp. Porous Media **129,** 413 (2019).

[73] R. T. Armstrong, C. Sun, P. Mostaghimi, S. Berg, M. Rücker, P. Luckham, A. Georgiadis, and J. E. McClure, Multiscale characterization of wettability in porous media, Transp. Porous Media **140,** 215 (2021).

[74] R. Huang, A. L. Herring, and A. Sheppard, Effect of saturation and image resolution on representative elementary volume and topological quantification: An experimental study on bentheimer sandstone using micro-CT, Transp. Porous Media **137,** 489 (2021).

[75] Z. Wojna, V. Ferrari, S. Guadarrama, N. Silberman, L.-C. Chen, A. Fathi, and J. Uijlings, The Devil Is in the Decoder: Classification, Regression and GANs, ArXiv:1707.05847 [Cs] (2019).

[76] D. Ulyanov, A. Vedaldi, and V. Lempitsky, Instance Normalization: The Missing Ingredient for Fast Stylization, ArXiv:1607.08022 [Cs] (2017).

[77] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam, MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications, ArXiv:1704.04861 [Cs] (2017).

[78] J. Bear, *Dynamics of Fluids in Porous Media* (Courier Corporation, Chelmsford, 1988).

[79] K. Xi, Y. Cao, K. Liu, J. Jahren, R. Zhu, G. Yuan, and H. Hellevang, Authigenic minerals related to wettability and their impacts on Oil accumulation in tight sandstone reservoirs: An example from the lower cretaceous quantou formation in the southern songliao basin, China, J. Asian Earth Sci. **178,** 173 (2019).

[80] J. E. McClure, Z. Li, M. Berrill, and T. Ramstad, The LBPM Software Package for Simulating Multiphase Flow on Digital Images of Porous Rocks, ArXiv:2007.12266 [Physics] (2020).

[81] T. Chung, Y. D. Wang, R. T. Armstrong, and P. Mostaghimi, Approximating permeability of microcomputed-tomography images using elliptic flow equations, SPE J. **24,** 1154 (2019).

[82] Y. D. Wang, T. Chung, R. Armstrong, J. Mcclure, and P. Mostaghimi, Computations of permeability of large rock images by dual grid domain decomposition, Adv. Water Resour. **126,** 1 (2019).

[83] Y. D. Wang, T. Chung, A. Rabbani, R. T. Armstrong, and P. Mostaghimi, Fast direct flow simulation in porous media by coupling with pore network and laplace models, Adv. Water Resour. **150,** 103883 (2021).

[84] M. Latva-Kokko and D. H. Rothman, Scaling of Dynamic Contact Angles in a Lattice-Boltzmann Model, Phys. Rev. Lett. **98,** 254503 (2007).

[85] A. Mittal, R. Soundararajan, and A. C. Bovik, Making a "completely blind" image quality analyzer, IEEE Signal Process. Lett. **20,** 209 (2013).

[86] A. Mittal, A. K. Moorthy, and A. C. Bovik, No-Reference image quality assessment in the spatial domain, IEEE Trans. Image Process. **21,** 4695 (2012).

[87] https://www.digitalrocksportal.org.

[88] K. Singh, H. Menke, M. Andrew, C. Rau, B. Bijeljic, and M. J. Blunt, Time-Resolved synchrotron X-Ray micro-tomography datasets of drainage and imbibition in carbonate rocks, Sci. Data **5,** 180265 (2018).

[89] M. Rücker, S. Berg, R. T. Armstrong, A. Georgiadis, H. Ott, A. Schwing, R. Neiteler, N. Brussee, A. Makurat, L. Leu, M. Wolf, F. Khan, F. Enzmann, and M. Kersten, From connected pathway flow to ganglion dynamics, Geophys. Res. Lett. **42,** 3888 (2015).

[90] A. P. Sheppard, R. M. Sok, and H. Averdunk, Techniques for image enhancement and segmentation of tomographic images of porous materials, Phys. A: Stat. Mech. Appl. **339,** 145 (2004).

[91] Y. Bazaikin, B. Gurevich, S. Iglauer, T. Khachkova, D. Kolyukhin, M. Lebedev, V. Lisitsa, and G. Reshetova, Effect of CT image size and resolution on the accuracy of rock property estimates, J. Geophys. Res.: Solid Earth **122,** 3635 (2017).

[92] L. Leu, S. Berg, F. Enzmann, R. T. Armstrong, and M. Kersten, Fast X-Ray micro-tomography of multiphase flow in berea sandstone: A sensitivity study on image processing, Transp. Porous Media **105,** 451 (2014).

[93] 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).

[94] Z. Jiang, K. Wu, G. D. Couples, and J. Ma, The impact of pore size and pore connectivity on single-phase fluid flow in porous media, Adv. Eng. Mater. **13,** 208 (2011).

[95] L. E. Treiber and W. W. Owens, A laboratory evaluation of the wettability of fifty Oil-producing reservoirs, Soc. Pet. Eng. J. **12,** 531 (1972).

[96] C. A. Fauziah, E. A. Al-Khdheeawi, A. Barifcani, and S. Iglauer, in *SPE Gas & Oil Technology Showcase and Conference* (OnePetro, Dubai, 2019).

[97] B. Pan, X. Yin, and S. Iglauer, A review on clay wettability: From experimental investigations to molecular dynamics simulations, Adv. Colloid Interface Sci. **285,** 102266 (2020).

[98] L. W. Lake, *Enhanced Oil Recovery* (Prentice Hall Inc., 1989).

[99] P. P. Edwards, V. L. Kuznetsov, W. I. F. David, and N. P. Brandon, Hydrogen and fuel cells: Towards a sustainable energy future, Energy Policy **36,** 4356 (2008).