Computer Methods in Applied Mechanics and Engineering 448 (2026) 118469

Contents lists available at ScienceDirect

Comput. Methods Appl. Mech. Engrg.

journal homepage: www.elsevier.com/locate/cma

3D microstructure reconstruction of heterogeneous material from slice descriptors using explicit neural network

Xiangyun Ge [1], Liyuan Wang [1], Liam J. Garcia, Shan Zhong, Bingbing Chen, Chenfeng Li *

Zienkiewicz Institute for Modelling, Data and AI, Swansea University Bay Campus, Swansea SA1 8EN, United Kingdom

ARTICLE INFO

Keywords:
Microstructure reconstruction
Heterogeneous material
Explicit descriptor aware
Generative adversarial network

ABSTRACT

Heterogeneous materials such as rocks, concrete, and composites exhibit random microstructures that strongly influence their physical properties. These structures can be captured using imaging techniques such as micro-Computed Tomography (micro-CT) and Scanning Electron Microscopy (SEM), but such methods are costly. As an alternative, digital reconstruction offers a way to generate synthetic microstructural images. Traditional reconstruction methods provide geometric control but lack flexibility and generality. In contrast, Computer Vision (CV) approaches offer strong generative capabilities, but often lack interpretability and material-specific constraints. Conditional generation introduces descriptors as labels, yet struggles with control during training. In this work, we enhance conditional generation by decomposing the loss into two parts: an implicit CV-based component and an explicit, descriptor-driven component inspired by traditional methods. The explicit loss includes a targeted loss for individual inputs and a distribution loss for overall output quality, ensuring both accuracy and diversity. To support fast training, a rapid descriptor regression model is developed and integrated into our digital reconstruction workflow. We validate our method on Fontainebleau sandstone, Polytetrafluoroethylene (PTFE), and SOFC electrodes, demonstrating improved reconstruction performance compared to the current state-of-the-art 2D-to-3D method. Our approach accurately captures statistical descriptors, even for complex geometries, and maintains strong consistency across multiple 2D slices representing 3D structures.

1. Introduction

Heterogeneous materials such as soil, rock, composites, and alloys are widely used in engineering fields such as construction, aerospace, and automotive industries due to their unique physical and structural properties [1,2]. Understanding how macroscopic properties (e.g., strength, elasticity, durability) relate to microscopic features (e.g., grain size, phase distribution, porosity) is key to optimizing performance and designing new materials [3]. This relationship is often quantified using statistical descriptors, which capture essential microstructural details and enable the prediction of material behaviour [4,5]. Accurate microstructural representations are typically obtained through imaging techniques like micro-Computed Tomography (micro-CT) [6–10], X-ray imaging [11,12], or Scanning Electron Microscopy (SEM) [13–15]. However, these methods are costly and limited by sample size and condition. To

* Corresponding author.
E-mail address: c.f.li@swansea.ac.uk (C. Li).
[1] co-First Author

https://doi.org/10.1016/j.cma.2025.118469
Received 29 April 2025; Received in revised form 2 October 2025; Accepted 3 October 2025
Available online 14 October 2025

address these challenges, digital reconstruction algorithms offer an effective alternative. They enable the generation of realistic microstructures, evaluation of descriptor variability, and estimation of material properties, while overcoming experimental constraints related to imaging, segmentation, and sample preparation.

Traditional digital reconstruction methods include statistical methods, field transformation methods [16–18], multi-point statistics methods [19], and particle-based approaches [20,21], among others. Statistical methods reconstruct microstructures by explicitly characterizing images of real materials and stochastically generating synthetic ones based on these characteristics. Among these, Stochastic Optimization Algorithms (SOA) are the most widely used. SOA works by randomly generating candidate structures and iteratively optimizing them to match the target's statistical features. This approach relies on two key components: statistical descriptors and reconstruction algorithms. The process begins with a random structure, which is refined through optimization to reproduce the desired statistical properties. Descriptors quantify morphological features such as topological connectivity, spatial correlation, and orientation. Optimization techniques such as simulated annealing [22,23], particle swarm optimization [24], and genetic algorithms [25] are used to adjust the structure until it closely matches the target. Descriptors serve as criteria in the objective function, guiding the reconstruction. SOA is popular for its flexibility, accuracy, and controllability due to its explicit use of statistical descriptors. However, it is computationally intensive, making it inefficient for large-scale 3D reconstructions.

In recent years, Computer Vision (CV) methods have gained traction for digital reconstruction, offering several advantages over traditional statistical approaches. These methods leverage neural networks to generate images and use tailored loss functions for feedback-driven accuracy improvement. Unlike statistical methods, CV-based approaches are rooted in image processing rather than statistical descriptors, though their "black-box" nature often draws criticism for lacking interpretability. Deep learning methods include Variational Autoencoders (VAEs) [26,27], transfer learning [28,29], Recurrent Neural Networks (RNNs) [30,31], Multilayer Perceptrons (MLPs) [32–34]. Recently, with the rise of denoising diffusion probabilistic models (DDPMs) in the field of image generation, an increasing number of studies have introduced diffusion models [35–41] and flow-based approaches [42–44] into microstructure reconstruction. Among these, Generative Adversarial Networks (GANs) are the most widely used. GANs consist of two competing networks - a generator and a discriminator - that iteratively improve each other, allowing for the generation of highly realistic synthetic data. Unlike other models, GANs were first applied to 3D reconstruction using full 3D datasets [45,46]. However, due to the difficulty in acquiring 3D training data, recent studies now focus on reconstructing 3D microstructures from more readily available 2D cross-sectional images [47]. This approach has proven practical for material science applications. Several GAN variants have been employed, including VAE-GAN [48], BicycleGAN [49], StyleGAN [50], and WGAN [51]. Architectures like SliceGAN [52] offer the flexibility to model various material types without requiring changes to training parameters.

CV-based methods often lack interpretability, offering little insight into what features are learned or why certain optimizations occur. To address this, Conditional GANs (CGANs) introduce auxiliary information to guide the generation process [53]. In heterogeneous material reconstruction, two primary strategies are used to incorporate such information: CV-based and statistics-based approaches. In the CV-based approach, similarity is evaluated using deep learning networks. For example, Li et al. [54] used a fine-tuned VGG-16 model with Gram-matrix loss to capture microstructure texture features alongside standard GAN losses. Yin et al. [55] trained a network to ensure edge texture consistency between generated and reference images. The statistics-based approach uses physical descriptors as constraints. Laloy et al. [46] applied checker functions to calibrate volume fraction and two-point correlation curves, improving accuracy. Du et al. [56] trained a CNN to predict effective thermal conductivity and used it as a performance evaluator for the generator, bypassing the discriminator entirely. Some studies combine both strategies. Shams et al. [57] employed statistical descriptors in CGANs. Feng et al. [58] integrated standard GAN loss with pixel-wise, pattern distribution, and porosity losses. Cao et al. [59] used porosity as part of the latent space in a hybrid InfoGAN-StyleGAN model, enforcing distribution constraints with a classifier. Chi et al. [60] implemented a conditional VAE-GAN incorporating porosity in the encoder, decoder, and discriminator. Zhou and Wu [61] conditioned on porosity, tortuosity, surface area, and fractal dimension. Recently, there has been some reconstruction research based on diffusion generative models [35–41] and flow-based models [42–44]. In conditional diffusion models, for instance, Lyu et al. [41] embedded permeability as a feature encoding into diffusion-based generative models, enabling the generation of three-dimensional microscopic results for random porous materials with controllable permeability. Conditions are typically introduced as feature vectors during the synthesis stage to guide the generation of microstructures that align with specific physical properties [38–41]. While this emerging area shows significant promise, it remains in an early stage, and further research is required to improve model generalizability.

Despite recent efforts to improve the interpretability of CV-based reconstruction methods, several key challenges persist. Current approaches often use descriptors as labels for the discriminator, allowing only implicit optimization - unlike SOA, which uses descriptors as explicit objective functions. Moreover, these methods are typically limited to simple scalar descriptors (e.g., volume fraction) or physical properties (e.g., thermal conductivity), failing to capture the complex morphological characteristics of heterogeneous microstructures. Additionally, most conditional CV-based methods are confined to either 2D-to-2D or 3D-to-3D reconstruction, restricting practical applicability. To address these limitations, this work proposes a bi-directional descriptor-microstructure linked CGAN framework that reconstructs 3D microstructures from 2D slice descriptors. The training data consists solely of 2D microstructure images and their corresponding descriptors. We introduce the Statistical-Aware Generative Adversarial Network (SAGAN), which includes: a generator that transforms slice descriptors into 3D microstructures; a discriminator that evaluates 2D slices extracted from the generated 3D structure using CV metrics; a ResNet-like descriptor regression model that rapidly estimates descriptors from 2D slices, enabling explicit optimization during training. This bidirectional architecture allows descriptors to both guide and validate the generated microstructures, ensuring alignment between input and output and efficient convergence during training. The proposed SAGAN primarily introduces an explicit descriptor constraint and conditional input into CV methods, functioning as a plug-and-play

Table 1
The full legend of the 41 microstructure descriptors reproduced from Cui et al. [5]. These descriptors were systematically derived through fractal dimension analysis, physical interpretations and other fields across different categories. This integrated approach enables comprehensive microstructural characterization.

No.	Descriptor	No.	Descriptor
1	Volume Fraction	21	Nearest-neighbor Distribution Function (1)_ep
2	Specific Internal Surface Density	22	Nearest-neighbor Distribution Function (2)_ev
3	Grain-size Distribution Function	23	Nearest-neighbor Distribution Function (3)_hp
4	Pore-size Distribution Function	24	Nearest-neighbor Distribution Function (4)_hv
5	Core Distribution Probability	25	Nearest-neighbour Orientation Distribution Function
6	Directional Growth Probability	26	Ripley's K Function
7	Orientation Distribution Function	27	Coarseness
8	Two-point Correlation Function	28	Contiguity
9	Two-point Cluster Correlation Function	29	Point/q-particle Correlation Function
10	Lineal-path Function	30	Surface-particle Correlation Function
11	Chord-length Density Function	31	Surface Correlation Function (1)_SSC
12	Aspect Ratio Distribution Function	32	Surface Correlation Function (2)_SVC
13	Radial Distribution Function	33	Maximum Grain Radius
14	Integral Correlation Function	34	Minimum Grain Radius
15	Topological Entropic Descriptor	35	Mean
16	Microcanonical Entropic Descriptor	36	Variance
17	Local Percolation Probability	37	Standard Deviation
18	Local Porosity Distribution Function	38	Autocovariance
19	Tortuosity	39	Autocorrelation Function
20	Voronoi Polygon Area	40	Power Spectral Density Functions
		41	Margin Distribution Function

module that can be integrated with other CV-based approaches to enhance SOTA methods such as SliceGAN [52] and diffusion models [62]. Key features of SAGAN include:

1. **Explicit Training with Dual Loss Functions:** SAGAN integrates both SOA-style descriptor loss and CV-based fusion loss into the GAN training. This enables real-time monitoring and explicit optimization of descriptor accuracy. The loss function considers: individual-view loss to ensure each generated output accurately reflects its input descriptor; and distribution-view loss to ensure diversity across a batch by matching the statistical distribution of the reference dataset.
2. **Fast Descriptor Estimation via ResNet Regression:** Traditional graphics-based descriptor computation is slow. To enable fast, GPU-friendly training, SAGAN uses a ResNet-style regression network to quickly compute descriptors from generated slices, allowing real-time optimization.
3. **Descriptor-Driven Generator with Injection Mechanism:** Unlike conventional generators that rely solely on random noise, SAGAN incorporates a descriptor injection mechanism to integrate multiple statistical descriptors directly into the transpose CNN pipeline. This enables precise control over the generation process, linking 2D slice descriptors to 3D outputs.

2. Reconstruction methodology

This section presents the SAGAN algorithm and network architecture. We begin by highlighting the importance of statistical descriptors in characterizing microstructures and explaining the selection of descriptors for comprehensive representation. Next, we detail the reconstruction algorithm for isotropic materials, including the network design and loss functions. Finally, we extend the approach to accommodate anisotropic materials.

2.1. Descriptors criteria and data process

Statistical descriptors are essential for microstructure analysis, enabling explainable predictions of physical properties [63] and accurately capturing spatial patterns, randomness, and distributions. They provide a detailed statistical representation of a material's heterogeneity and serve as quantitative metrics for evaluating reconstruction quality. Incorporating the explicit descriptor-based characterization from SOA methods into CV-based approaches can improve accuracy and predictability. A central challenge, however, is selecting appropriate descriptors, as they capture different aspects - geometric, statistical, or topological - each offering unique insights into microstructural features.

A comprehensive correlation analysis of 41 statistical descriptors was conducted [5], as listed in Table 1, with their interrelationships shown in Table 2. By selecting well-suited descriptors, key morphological information can be captured with high accuracy and efficiency using a limited set. Balancing these factors, we chose five descriptor groups as input: Two-Point Correlation Function (TPCF), Pore Size Distribution (PSD), Core Distribution Probability (CDP), Lineal-Path Function (LPF), and Two-Point Cluster Correlation Function (TPCCF).

For each image in the data set, five descriptors are computed as corresponding labels to describe critical morphological features. First, the original descriptors are calculated using an in-house simulation code based on their mathematical definitions. Based on

Table 2
Correlation graph of all 41 descriptors in [5]. Cui et al. employed a clustering algorithm to analyze the correlations among 41 descriptors. Taking the fifth row in the table as an example, the descriptors were divided into five groups (A, B, C, D, and E). Within each group, any descriptor overlaps more strongly with the others, enabling a sufficiently comprehensive characterization of a material's microstructural characteristics by using just five descriptors.

	8	18	27	31	33	34	35	36	37	38	39	40	41	1	16	28	14	19	29	9	13	20	21	15	23	12	22	24	26	30	32	11	7	3	5	25	4	6	2	10	17
2-Groups	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B	B	B	B	B
3-Groups	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	C	C	C	C	C	C	C	C	C	C	C	C	C	B	B	B	B	B
4-Groups	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	C	C	C	C	C	C	C	C	C	C	C	C	C	B	B	B	D	D
5-Groups	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	E	E	E	E	E	E	E	E	E	E	E	C	C	C	C	C	B	B	B	D	D
6-Groups	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	F	E	E	E	E	E	E	E	E	E	F	C	C	C	C	C	B	B	B	D	D
7-Groups	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	F	E	E	E	E	E	E	E	E	E	F	G	C	C	C	C	B	B	B	D	D
8-Groups	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	H	H	H	F	H	H	H	H	H	E	E	E	E	E	G	C	C	C	C	B	B	B	D	D
9-Groups	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	H	H	H	F	H	H	H	H	H	E	E	E	E	E	I	G	C	C	C	B	B	B	D	D
10-Groups	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	J	H	H	H	F	H	H	H	H	H	E	E	E	E	E	I	G	C	C	C	B	B	B	D	D
11-Groups	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	J	H	H	H	F	H	H	H	H	H	E	E	E	E	E	I	G	K	C	C	B	B	B	D	D
12-Groups	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	J	H	H	H	F	H	H	H	H	H	E	E	E	E	E	I	G	K	L	C	C	B	B	B	D	D
13-Groups	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	J	H	H	H	F	H	H	H	H	H	E	E	E	E	E	I	G	K	L	C	C	B	B	M	D	D
14-Groups	A	A	A	A	A	A	A	A	A	A	A	A	A	N	N	J	N	N	N	F	H	H	H	H	H	E	E	E	E	E	I	G	K	L	C	C	B	B	M	D	D
15-Groups	A	A	A	A	A	A	A	A	A	A	A	A	A	N	N	J	N	N	N	F	H	H	H	H	H	E	E	E	E	E	I	G	K	L	C	C	B	B	M	D	O

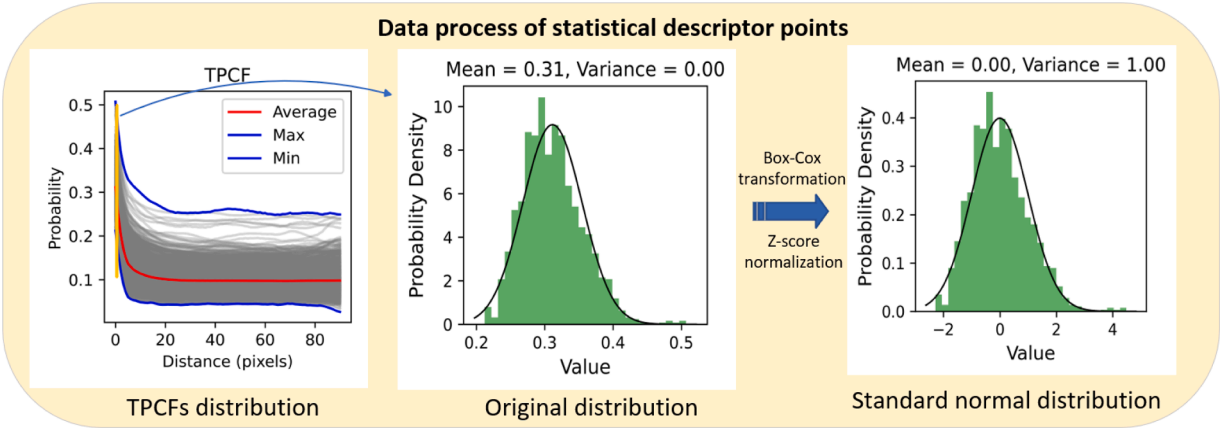

Fig. 1. Data processing for slice statistical descriptor points. Take TPCF as an example. The selected descriptor key point is transformed from normal distribution into a standard normal distribution using Box-Cox transformation followed by Z-score normalization.

the Central Limit Theorem, descriptors exhibit fundamental bell-shaped distribution characteristics when the number of random samples is sufficiently large. However, since each descriptor contains distinct geometric and/or physical information, they do not conform to a standard normal distribution. Therefore, data processing is required to transform them into a tractable standard normal distribution (zero mean and unit variance) suitable for training purposes. Box-Cox transformation and Z-score normalization are applied to transform each key point set into a standard normal distribution. This not only enables the use of Kullback-Leibler (KL) divergence in the loss function, but also accelerates network convergence and improves the robustness of the regression model. The data processing procedure is illustrated in Fig. 1. These points with a standard normal distribution are used for descriptor loss calculation.

After data processing, each slice descriptor is represented by its mean and standard deviation, which are incorporated into the generator as conditional labels alongside random noise. This design is inspired by the SOA-like method and reparameterization process from VAE, where descriptors guide the generation process to ensure that the statistical properties of the generated microstructures closely match those of real samples. To reinforce this explicit descriptor control, we introduce a dual-component descriptor loss into the generator: one component evaluates the numerical similarity of individual descriptor values, while the other assesses the overall statistical distribution of descriptors, effectively measuring the statistical descriptor discrepancy between generated and real microstructures. By integrating slice descriptors as conditional inputs and custom descriptor-based losses, our model achieves explicit control over the GAN generation process. This not only enhances the fidelity of generated structures but also accelerates training convergence and improves overall generation quality.

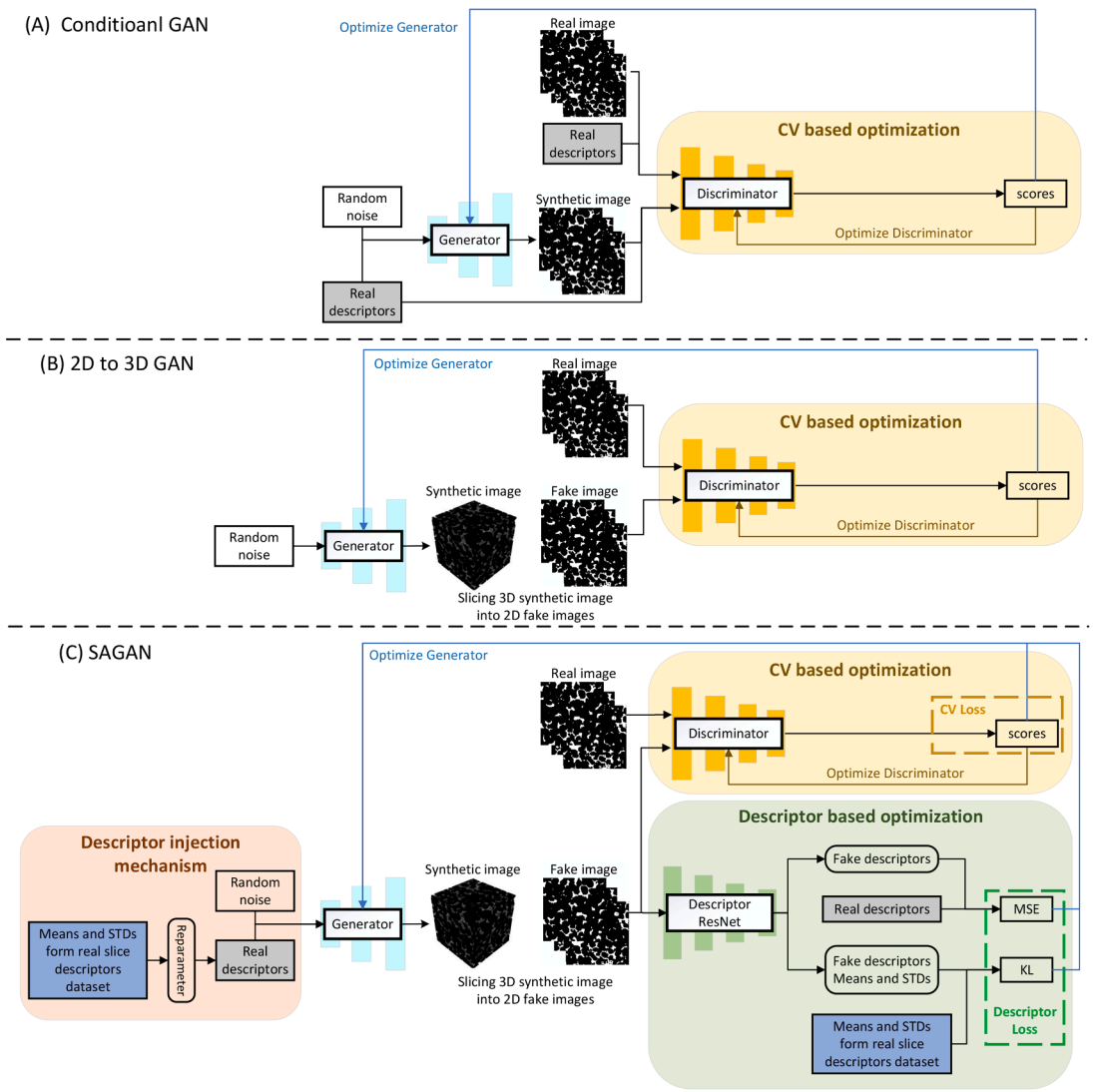

Fig. 2. Workflow of isotropic material reconstruction using conditional generative adversarial network (CGAN), two-dimension to three-dimension generative adversarial network (GAN) and statistical-aware generative adversarial network (SAGAN). (A) Two-dimension to two-dimension isotropic reconstruction using the standard CGAN method. (B) Two-dimension to three-dimension isotropic reconstruction using GAN method(SliceGAN). (C) Two-dimension to three-dimension isotropic reconstruction using SAGAN.

2.2. Isotropic reconstruction algorithm

The isotropic reconstruction algorithms for CGAN, 2D-to-3D GAN, and SAGAN are illustrated in Fig. 2. The standard CGAN algorithm is shown in Fig. 2(A), where 2D or 3D microstructure images and their corresponding descriptor or physical property labels are used as a supervised task. The generator takes real descriptors and noise to produce 2D or 3D images, which are then fed into the discriminator along with the descriptors used in the generator for training. However, during training, the descriptors are implicit and cannot be used to solve for the descriptors of the generated images. The descriptors inputted into the generator are only used as labels for the discriminator. Additionally, extending this approach from 2D to 3D reconstruction presents significant challenges. Since cross-section micrographs can contain sufficient information to statistically reconstruct 3D samples [52], the standard 2D-to-3D GAN takes noise as generator input and slices the 3D synthetic microstructure into 2D cross-sectional slices in three principal directions. The discriminator takes all fake 2D cross-sectional slices without distinguishing directions and is trained with 2D references, as shown in Fig. 2(B). Despite this, there remains a gap between slice descriptors and 3D microstructure images. To directly address this in our framework, we have implemented an explicit descriptor-constrained training paradigm that explicitly anchors morphological feature learning to quantifiable microstructure descriptors. Rather than relying on opaque latent representations, our approach imposes dual descriptor loss mechanisms: (1) an individual descriptor loss that penalizes deviations between generated image descriptors and

target labels during each training step, ensuring precise replication of specified quantitative features; and (2) a distributed descriptor loss that statistically matches the feature distributions of generated microstructures (sampled from an image buffer) to the original dataset, guaranteeing holistic morphological fidelity. This design fundamentally shifts optimization from implicit "black-box" learning to explicit descriptor-driven control - where decreasing loss values directly correlate with measurable improvements in capturing predefined characteristics (five descriptors mentioned in Section 2.1. Consequently, the model's interpretability derives not from deciphering network internals, but from verifiable convergence of physically meaningful descriptors, as illustrated in our quantitative analyses. Fig. 2(C) depicts the isotropic SAGAN algorithm. SAGAN consists of a descriptor-aware generator with a descriptor injection mechanism, a purely CV-based discriminator for optimization, and an explicit descriptor-based optimizer.

The reconstruction process for isotropic materials is as follows:

- Data preparation: The training input data for SAGAN should consist of 2D microstructure images and their corresponding slice descriptors. Since isotropic heterogeneous materials do not exhibit geometric pattern differences among different directions, the dataset can be sampled from large 2D scan slices. Random sampling is used to obtain a large number of 2D slice images, and five statistical descriptors for each slice are computed. The resulting microstructure images and descriptors then form the database.
- Descriptor ResNet training: The input consists of 2D microstructure images and the corresponding descriptor points as labels. These label points are first processed to follow the standard normal distribution. After the feature extraction and regression by the ResNet-style network, the predicted descriptor results are compared with the ground truth labels through the Mean Square Error (MSE) loss function to guide the optimization and convergence of the model. The well-trained descriptor model closely matches the input descriptor points, where the individual prediction error is controlled within $2\,\%$. This is sufficiently accurate for the comparison of descriptor loss on Generator, enabling rapid and accurate characterization with statistical descriptors. This descriptor training is completed before the training of generator and discriminator.
- SAGAN training: The training procedure of SAGAN follows a conventional CGAN framework, alternating between generator training with a fixed discriminator and discriminator training with a fixed generator. The key distinctions lie in the reparameterization of conditional descriptor variables and the incorporation of descriptor loss. The specific workflow proceeds as follows:
  (a) Discriminator training: Fix the generator parameters and train the purely CV-based discriminator. For 2D microstructure images read from the database, encourage the discriminator to give higher scores to cross-sectional slice images from the reference dataset. Reparameterize synthetic descriptors $\tilde{y}$ based on the mean μ and standard deviation σ of each descriptor in the database and input these into the generator to obtain 3D synthetic images. The reparameterization process can be expressed as
  $$\tilde{y} = \mu + \sigma \times \epsilon, \tag{1}$$
  where ϵ is the independent auxiliary random variable. Slice the 3D images into 2D images and ensure that the discriminator gives lower scores to these synthetic 2D slice images.
  (b) Generator training: Fix the discriminator parameters and train the generator. The generator produces 3D synthetic images. Slice the 3D images into 2D images and input them into the discriminator to obtain the purely CV-based score. Simultaneously, input the slices into the Descriptor ResNet to obtain the corresponding descriptors, and compare these with the input reparameterized descriptors and the database distribution to obtain the explicit descriptor loss. Sum these two loss functions to optimize the generator, resulting in outputs that are more realistic in terms of CV and more closely aligned with the descriptors.
- Synthetic image inference: Once all the loss of generator and discriminator converged to a sufficiently small value, all the parameters of generator is fixed and saved for image inference. In reconstruction process, descriptor variables are reparameterized and fed to the saved generator to get synthetic images.

2.3. Network architecture

This section introduces the network structure parameters for each component, starting with the generator that includes the Descriptor Injection Mechanism and the CV-based discriminator, followed by the regression model of the Descriptor ResNet.

2.3.1. Descriptor-aware generator and discriminator

The generator in SAGAN is designed to synthesize 3D microstructures from a combination of a random noise vector and 2D statistical descriptor variables. Unlike conventional conditional GANs, which inject conditioning information only at the input layer, our generator employs a descriptor injection mechanism that integrates descriptor features into both the first and second convolutional layers via learned fully connected mappings. To ensure computational efficiency while maintaining expressiveness, the injection is limited to these two early layers, which has been empirically found sufficient to guide the generation process. As illustrated in Fig. 3(A), the generator consists of a sequence of transposed 3D convolutional layers with progressively decreasing feature dimensions (from 1024 to 32 channels), each with a kernel size of 4, stride of 2, and padding of 2. Batch normalization and ReLU activation are applied after each transposed convolutional layer, except for the final one. Following the convolutional stack, the feature map is upsampled via trilinear interpolation and passed through a final 3D convolutional layer with a kernel size of 3 and no padding. A voxel-wise softmax activation is then applied to produce a multi-phase probability volume with the same spatial resolution as the target microstructure.

The discriminator adopts a conventional 2D convolutional architecture. It takes as input either real- or generated- 2D slices of microstructures and produces a scalar output indicating their authenticity. The network consists of five 2D convolutional layers with progressively increasing channel sizes (from 32 to 512), each with a kernel size of 4, stride of 2, and padding of 1, enabling gradual spatial downsampling. Each layer is followed by a ReLU activation function. The final layer outputs a single scalar

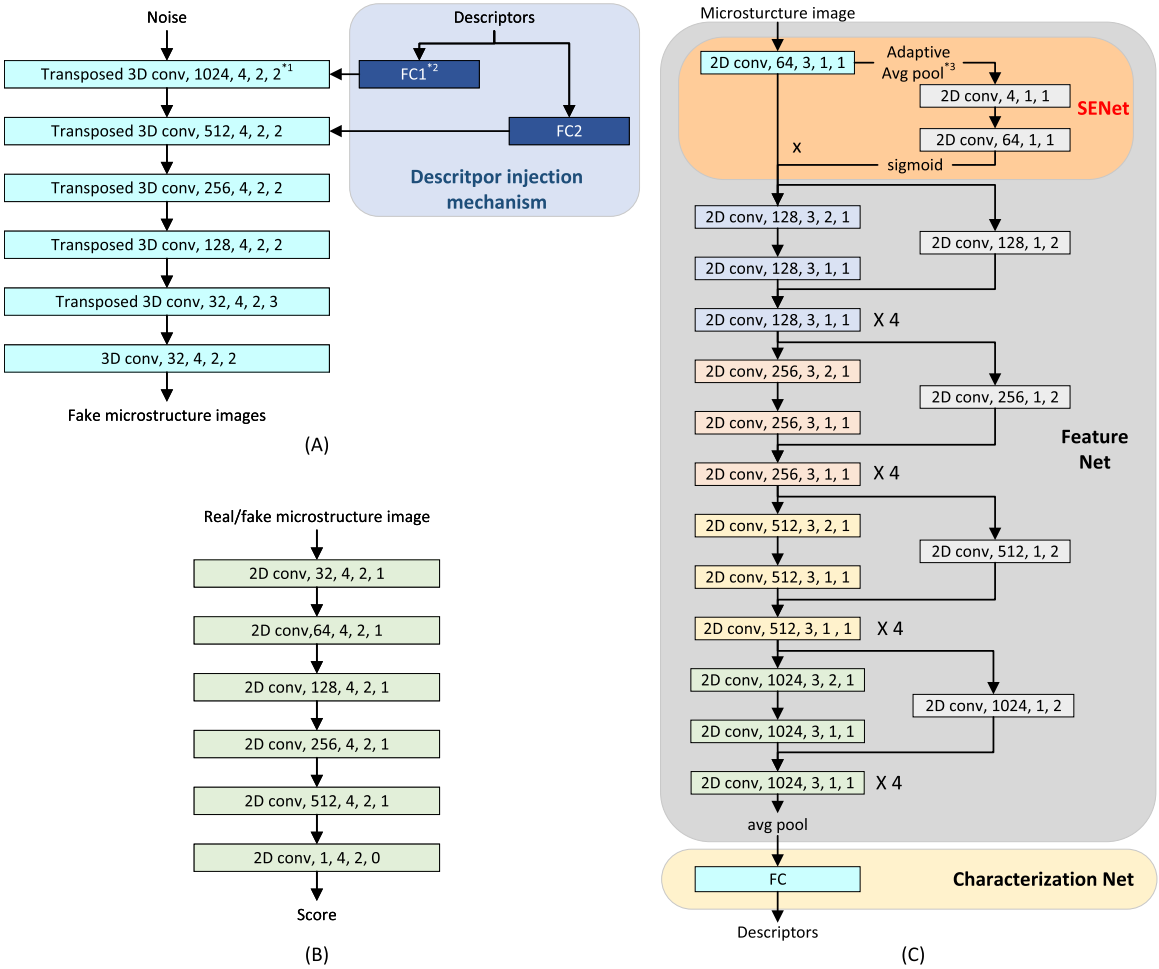

Fig. 3. Overview of the proposed statistical-aware generative adversarial network (SAGAN). (A) Descriptor-aware generator that synthesizes 3D microstructure images from noise and statistical descriptors. The generator uses a stack of transposed 3D convolutions with descriptor injection after the first two layers. (B) 2D discriminator that receives real or generated microstructure slices and outputs a probability score. (C) Descriptor ResNet-based characterization network, composed of an SE (Squeeze-and-Excitation) module and a multi-scale ResNet backbone, which predicts the statistical descriptor values from 2D slices. 1.Convolutional layers are annotated as "conv type, channels, kernel size, stride, padding". 2. FC: Fully connected layer 3. Avg pool: Average pooling.

score that reflects the likelihood of the input slice being real. The Wasserstein GAN with Gradient Penalty (WGAN-GP) framework [64] is adopted to stabilize training and enforce the Lipschitz continuity constraint. The gradient penalty term effectively prevents mode collapse and enhances convergence behavior, leading to more interpretable loss metrics. The full architecture is provided in Fig. 3(B).

2.3.2. ResNet-based microstructure characterization

To enhance the accuracy of descriptor estimation, a characterization network is constructed based on a 2D ResNet architecture [65]. The network takes a 2D microstructure slice as input and first applies a Squeeze-and-Excitation (SE) module [66] to capture channel-wise attention, thereby improving sensitivity to descriptor-relevant features. The SE module consists of a global average pooling layer followed by two 1×1 convolutional layers and a sigmoid activation.

After the SE block, the input is processed through four residual stages, each comprising four residual blocks with increasing channel dimensions: 128, 256, 512, and 1024, respectively. Each block adopts a bottleneck structure, and spatial downsampling is performed via convolutional layers with stride 2. A global average pooling layer is applied to the final feature map, and the resulting feature vector is passed through a fully connected layer to regress the target descriptor values. The details are illustrated in Fig. 3(C).

In contrast to the generator and discriminator, which are jointly optimized during adversarial training, the descriptor ResNet is pretrained on a paired dataset of 2D microstructure slices and their corresponding descriptors. Once trained, its weights are frozen and the model is used as a fixed feature evaluator in SAGAN. As for the anisotropic microstrucutre reconstruction, three separate descriptor ResNets are trained to characterize statistical features along the X, Y, and Z directions, respectively.

2.4. Loss functions of isotropic reconstruction

Two types of loss functions are employed: the CV loss, which is a generator-conditional variation of WGAN-GP, and the descriptor loss, which incorporates both individual and distribution perspectives.

As discussed earlier, we utilize WGAN-GP due to its ability to address the mode collapse issue common in traditional GANs. It provides a more stable training process and ensures meaningful gradients during training. The gradient penalty component enforces a Lipschitz constraint, mitigating vanishing or exploding gradient problems and enhancing the overall robustness and quality of the generated samples. To integrate descriptors as labels in the descriptor injection mechanism, we use the CGAN method. The generator loss for this part is defined as follows:

$$L_{G-CV} = -\mathbb{E}_{\tilde{x}\sim\mathbb{P}_g}[D(\tilde{x})|\tilde{y}], \tag{2}$$

where $\tilde{x}$ is the synthetic images, $\mathbb{P}_g$ is the generator's distribution. The discriminator is not conditional because we employ a descriptor-based optimization method as an alternative to the conditional part. The CV loss of the discriminator is defined as follows:

$$L_D = \mathbb{E}_{\tilde{x}\sim\mathbb{P}_g}[D(\tilde{x})|\tilde{y}] - \mathbb{E}_{x\sim\mathbb{P}_{data}}[D(x)] + \lambda\mathbb{E}_{\hat{x}\sim\mathbb{P}_{\hat{x}}}[(\|\bigtriangledown_{\hat{x}} D(\hat{x})\|_2 - 1)^2], \tag{3}$$

where $\hat{x}$ represents points sampled uniformly along straight lines between pairs of points from the real data distribution $\mathbb{P}_{data}$ and the generator's distribution, and λ is the weight of gradient penalty.

The descriptor-based optimization process begins by taking 2D cross-section slices from 3D synthetic images and computing descriptors using the descriptor ResNet. The foundation for using descriptor-based optimization in generative model training is twofold: individual evaluation and dataset-wide evaluation. From an individual perspective, the MSE measures the difference between the input and realization descriptors. During each training iteration, the slice descriptors of each 3D synthetic image are compared with the input real descriptors. The individual descriptor loss is defined as:

$$L_{individual} = MSE(\tilde{y}, y) = \frac{1}{n}\sum_{i=1}^{n}(\tilde{y}_i - y_i)^2, \tag{4}$$

where $\tilde{y}$ is the descriptor from synthetic images. Secondly, from a distribution perspective, a pool of generated images is created during training. This pool is progressively updated, and the average and variance of the descriptors in the pool are compared with the distribution of the whole input database using Kullback-Leibler (KL) divergence. Since the reference descriptors are consistent with the standard normal distribution, the holistic descriptor loss function can be defined as:

$$L_{distribution} = D_{KL}(P(y)\|P(\tilde{y})) = \log\left(\frac{\sigma_{\tilde{y}}}{\sigma_y}\right) + \frac{\sigma_y^2 + (\mu_y - \mu_{\tilde{y}})^2}{2\sigma_{\tilde{y}}^2} - \frac{1}{2}, \tag{5}$$

where $P(y)$ and $P(\tilde{y})$ are distributions of reparameterized descriptors and descriptors from synthetic images, respectively. Here μ and σ are the distributions' mean value and standard deviation.

Therefore, the loss function of the SAGAN generator can be written as:

$$L_{total} = L_{G-CV}(G, D) + \lambda \times L_{individual} + \lambda \times L_{distribution}, \tag{6}$$

where λ is the weight of descriptor loss. A dynamic weight scheme helps the generator balance the loss from the discriminator and the descriptors. In the early epochs of the training process, the generator is more influenced by the discriminator's direction than by the descriptor losses. This is because the descriptor ResNet is trained on real microstructure datasets and can only effectively contribute when the generator is capable of synthesizing analogous microstructure images. Since we supplemented the CV loss with a descriptor loss that has dynamic weights, we can achieve explicit control over the optimization process, ensuring a faster optimization of the generated results in terms of descriptors. Based on parameter-tuning experience across different materials, taking the reconstruction of Fontainebleau sandstone (Case 1) as an example, the weight of descriptor loss λ is set near 0 during the first 1000 batches of the initial epoch. After this stage, λ is immediately set to 1. For all subsequent epochs, λ increases by 0.2 per epoch according to:

$$\lambda^{(e)} = \lambda^{(e-1)} + 0.2, \tag{7}$$

where e denotes epoch index ($e \geq 1$). This specific configuration is not universally applicable and requires adjustment of hyperparameters for different materials and descriptors.

2.5. Anisotropic reconstruction algorithm

Unlike isotropic material reconstruction, directional consistency is not maintained in anisotropic materials. Consistency in a fixed direction for isotropic materials requires more consecutive cross-section slices in that direction to provide sufficient spatial geometric characteristics. Instead of simulating descriptors and physical properties from 2D images without distinguishing direction, we took slices from the X, Y, and Z directions and simulated the corresponding descriptors for each direction. Descriptors from the three directions are merged and injected into the generator along with noise. Consequently, three discriminators and three descriptor ResNets are applied to cross-section slices in each of the three directions. The loss function is a composite of three scores from the three discriminators, three MSEs, and three KL divergences from the three descriptor ResNets. The algorithm is illustrated in Fig. 4.

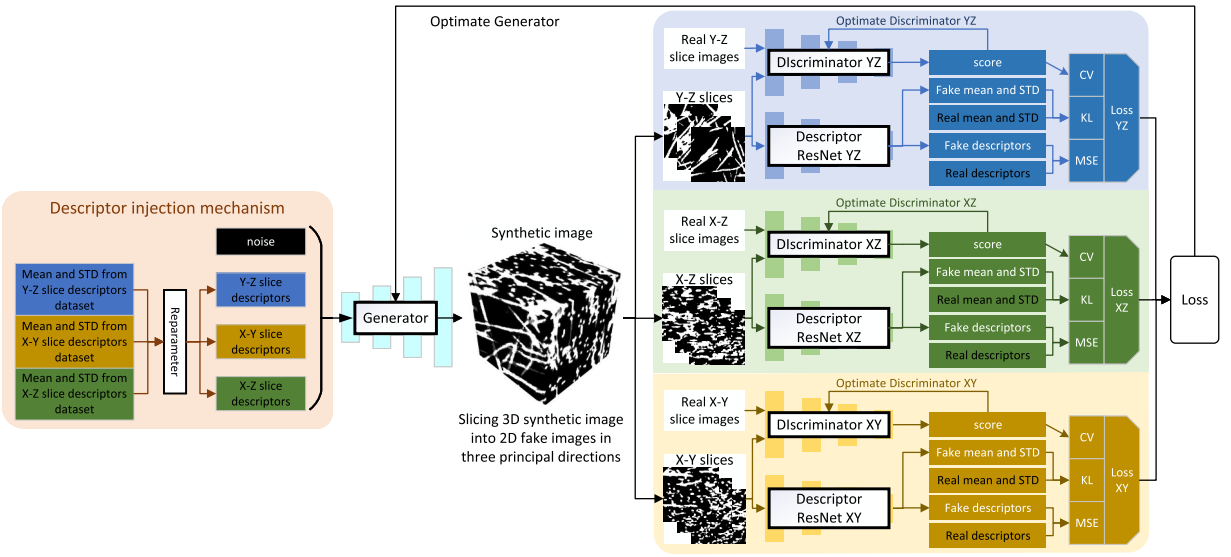

Fig. 4. Workflow of anisotropic material reconstruction.

The loss function for anisotropic reconstruction is a combination of three directional loss functions from the isotropic reconstruction algorithm and can be expressed as follows:

$$L_{total} = L_x + L_y + L_z = \sum_{x,y,z} [L_{G-CV}(G, D) + \lambda_{individual} \times L_{individual} + \lambda_{distribution} \times L_{distribution}] \tag{8}$$

Similar to isotropic materials, the addition of descriptor loss to the CV loss constrains the slice descriptors in all three directions during the reconstruction process. This explicit optimization of the generated results leads to higher optimization efficiency.

3. Results and discussion

The performance of the proposed digital reconstruction method is examined in this section by applying to different types of heterogeneous materials and comparing with recent machine learning approaches, including the SliceGAN and the Statistics-Informed Neural Network (SINN) methods. SliceGAN is a purely CV-based method where the generator improves solely based on the discriminator scores. SINN is a hierarchical characterization approach and a multi-level reconstruction approach suitable for two-phase materials [33]. We designed three distinct experiments to comparatively validate from different aspects:

- **Case 1:** An isotropic two-phase material Fontainebleau sandstone is considered. In this scenario, we compare the accuracy evolution of SliceGAN and SAGAN across multiple training epochs to validate the stable convergence property of SAGAN's accuracy improvement with increasing training volume, attributable to its explicit optimization strategy. Additionally, comparative evaluation against SINN demonstrates the superiority of our proposed method.
- **Case 2:** An anisotropic two-phase material polytetrafluoroethylene (PTFE) is considered. The PTFE sample features multiple straight fibers aligned in the yz-plane. The spatial anisotropy inherent in three-dimensional domains poses significant challenges for both SAGAN and SliceGAN, which rely solely on two-dimensional planar information as input. In this scenario, we systematically assess the impact of varying statistical descriptors on SAGAN's reconstruction accuracy through controlling the input variations of descriptor number.
- **Case 3:** A complex ternary material, Solid Oxide Fuel Cell (SOFC) cathode, is considered, which contains both large tortuous connected particles and very small isolated particles. This complexity imposes significant challenges on the 2D discriminator. Through comparative analysis of SAGAN and SliceGAN reconstruction outcomes, we demonstrate that SAGAN's explicit descriptor control mechanism achieves enhanced accuracy while effectively mitigating mode collapse and phase-specific over-generation phenomena on multiple-phase materials.

To ensure the accuracy of comparison for each method, all three experiments generated 3D microstructure volumes of size $128 \times 128 \times 128$. SliceGAN, as an earlier method, was originally trained on smaller images of size 64×64. Leveraging the enhanced memory and computational power of modern GPUs, we conducted a fair comparison between the two reconstruction methods by using identical generator architectures and increasing the training image size to 128×128. The discriminator architecture was also kept consistent. For SINN, the three-level reconstruction sizes were modified to $32 \times 32 \times 32$, $64 \times 64 \times 64$, and $128 \times 128 \times 128$. For the isotropic materials in Case 1 and Case 3, the training set comprises 28,800 images. For the anisotropic material in Case 2, 28,800 images are extracted from each of the three orthogonal directions (X, Y, and Z). Each image has a resolution of 128×128 as input

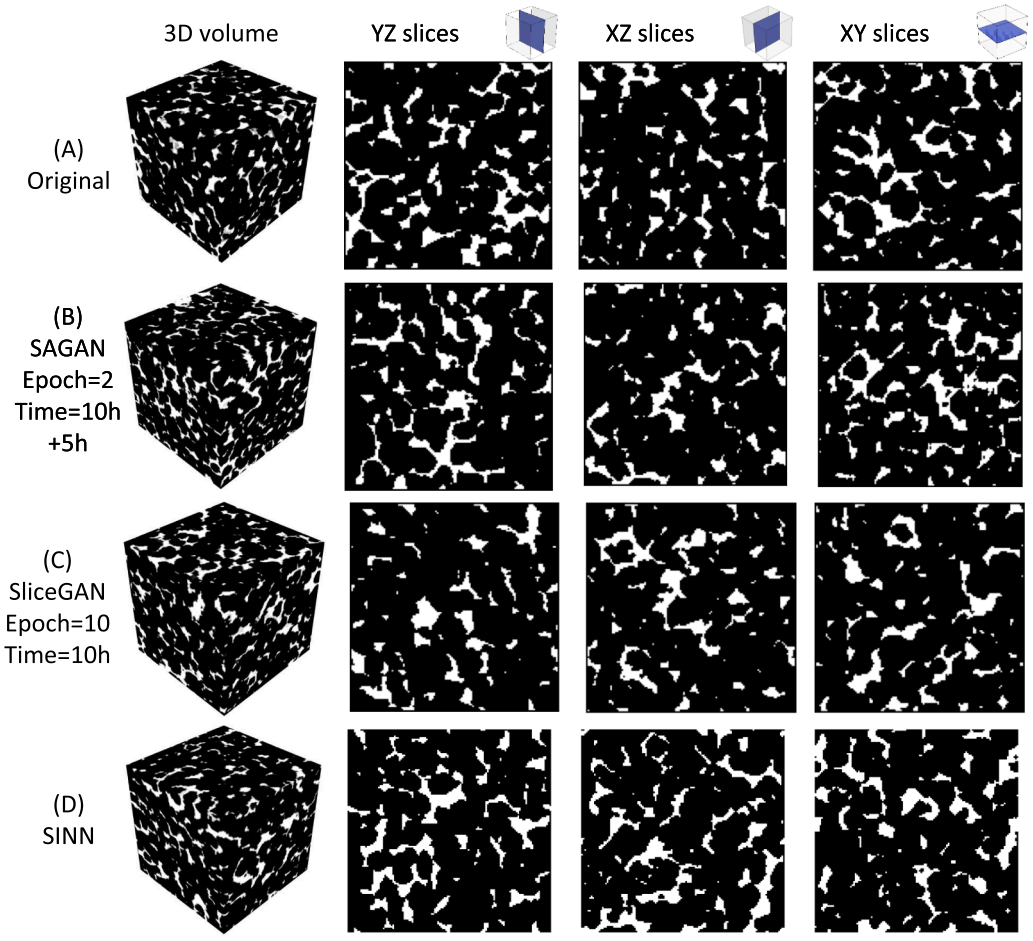

Fig. 5. Three-dimensional microstructure and cross-sections along the three principal axes of Fontainebleau sandstone reference and reconstruction results. (A) Reference. (B) Synthetic result from SAGAN on epoch 2. (C) Synthetic result from SliceGAN on epoch 10. (C4) The best synthetic result from SINN at 17h.

for SAGAN and SliceGAN. The number of training samples for three levels of SINN are 90000, 60000 and 30000. All experiments were performed on an NVIDIA A5000 GPU with 16 GB of RAM.

3.1. Case 1 - Fontainebleau sandstone

Fontainebleau sandstone is a common porous rock whose microstructure is generally considered as isotropic. The reference is segmented into two phases: 0 for solid and 1 for pore. Following the algorithm explained in Section 2.2, SAGAN reconstructed this case using one discriminator and one descriptor ResNet for all three directions.

SAGAN was trained for 1 to 3 epochs, with wall-clock training times of 5, 10, and 15 h, respectively. SliceGAN was trained for 1 to 50 epochs, with results extracted at epochs 5, 10, 15, and 50, corresponding to training times of 5, 10, 15, and 50 h. It should be noted that the training of the accompanying Descriptor ResNet required an additional 5 h, which should be accounted for in the total reconstruction time. For SINN, we compared its optimal 17-h training result. Here we present: (1) SAGAN at epoch 2 (10 h), (2) SliceGAN at epoch 10 (10 h), and (3) SINN's benchmark. Quantitative comparisons of relative error, KL divergence, and variance across training stages demonstrate the superiority of our explicit descriptor control paradigm. The three-dimensional structures and cross-sectional slices along the three principal directions are shown in Fig. 5. Visually comparing the images, both SAGAN and SliceGAN effectively capture the geometric characteristics of the structures from the early training process. We provide a quantitative evaluation of accuracy based on descriptors, specifically focusing on the slice descriptors used as inputs for SAGAN and the voxel descriptors that were not considered. Furthermore, physical simulations were performed to examine the percolation properties of Fontainebleau sandstone samples in Case 1.

Although the purely CV-based SliceGAN can effectively learn the geometric patterns of microstructures, this learning process is implicit and cannot meaningfully control descriptors like SOA methods. The optimization process does not explicitly converge

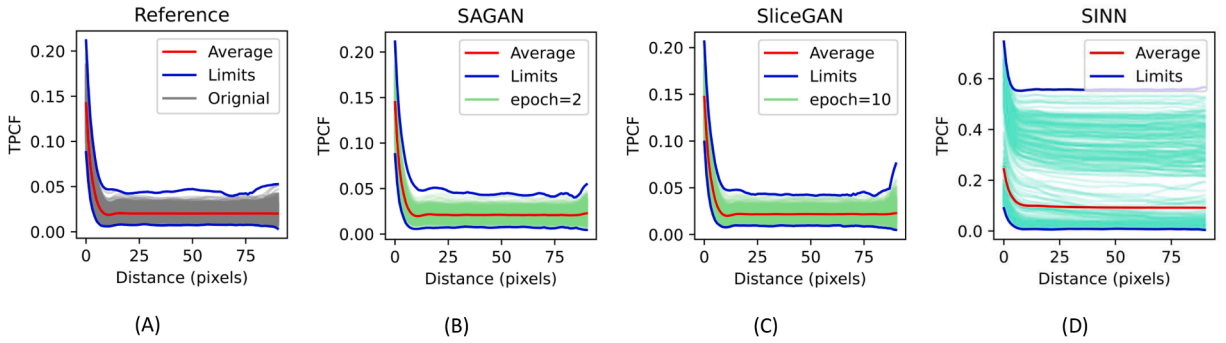

Fig. 6. Fontainebleau sandstone slice Two-Point Correlation Function (TPCF) results. (A) Reference. (B) SAGAN, epoch 2. (C) SliceGAN, epoch 10. (D) SINN.

towards the descriptors, resulting in potentially prolonged training epochs without necessarily improving accuracy beyond the initial epochs. Our proposed SAGAN explicitly controls the descriptor loss by solving descriptors, constraining the generation of planar microstructures during training. Additionally, by incorporating KL divergence, the diversity of generated samples is ensured. Since the accuracy of the 3D reconstruction can be controlled by the quality of consecutive 2D slices, SAGAN achieves optimization of 3D descriptors even without considering 3D statistical descriptors as constraints, thanks to the explicit optimization algorithm.

3.1.1. Slice descriptor verify

As discussed in Section 2.1, we compare five kinds of slice descriptors and divide them into three groups for analysis: Two-Point Correlation Function (TPCF), short-range pore information using Pore Size Distribution (PSD) and Lineal-Path Function (LPF), and long-range solid information using Core Distribution Probability (CDP) and Two-Point Cluster Correlation function (TPCCF). Due to the uncertainty of SINN's generated results, the descriptor results of its generated images are much worse than those of SAGAN and SliceGAN, so they are not used to measure various relative error with SAGAN and SliceGAN.

TPCF is a commonly used statistical measure to describe the spatial distribution of points within a given space. It quantifies the probability of two separate points belong to the same phase, thereby capturing the material's length scale and clustering arrangement. Here we present results for SAGAN at epoch 2, SliceGAN at epoch 10, and SINN, as shown in Fig. 6. Both SAGAN and SliceGAN successfully capture the planar characteristics of the original training data, whereas SINN exhibits significant deviations. In practical applications, SINN outputs require post-reconstruction filtering after extended computation. Comparatively, GAN-based reconstruction methods demonstrate superior accuracy over SINN. Due to SINN's low accuracy, it is excluded from subsequent relative error comparisons.

The explicit optimization of descriptors across training durations is shown in Fig. 7. In terms of relative error, SAGAN maintains stable error reduction over time due to descriptor loss constraints. Although SliceGAN shows error reduction between epochs 5–10, its errors increase at epochs 15 and 50, indicating that pure CV losses cannot guarantee stable convergence of slice descriptors. Similar patterns are observed in KL divergence and variance metrics representing distribution characteristics: SAGAN achieves stable convergence in accuracy throughout training owing to its distribution loss mechanism.

For short-range void phase information, the comparative results of different metrics in terms of relative error, variance and KL divergence were analysed on Pore Size Distribution (PSD) and Lineal Path Function (LPF). The results are shown in Figs. 8 and 9. The PSD measures the frequency and size of pores within a porous material, providing insights into the material's porosity and permeability. The LPF describes the probability that a randomly chosen line segment of a given length lies entirely within the pore space, characterizing connectivity and transport properties. The relationship between PSD and LPF lies in their complementary roles: PSD quantifies the size of the pores, while LPF captures the spatial arrangement and connectivity of these pores within the material. Both reconstruction methods demonstrate good accuracy at short-range scales, with relative errors below 2 % and KL divergence/variance values closely aligned with the original training data distribution. However, when examining the training timeline, SAGAN shows stable accuracy improvement, whereas SliceGAN exhibits accuracy divergence after epoch 10 in slice PSD results and accuracy degradation after epoch 15 in slice LPF results. The slice PSD results and the slice LPF results of each method are shown in Figs. A.1 and A.2 in the Appendix.

For long-range information, we estimate CDP and TPCCF. CDP refers to the likelihood of a solid phase point being distributed within a given space in a porous medium. It is used to quantify the spatial distribution and connectivity of the solid matrix within the material. TPCCF measures the probability that two points, separated by a certain distance, belong to the same cluster of connected pores within a porous medium. It is related to the TPCF but focuses specifically on connected regions rather than the overall spatial distribution. Although 2D slices may fail to fully reflect 3D cluster continuity, since connected regions in 3D may appear fragmented in 2D, we mitigate this by extracting sequential slices along three orthogonal directions. These multi-view slices provide complementary structural information. When the Sliced-TPCCF distributions of generated and target structures closely match, we infer that the 3D connectivity has been effectively preserved. CDP results are shown in Fig. 10. Both methods can achieve accurate representations

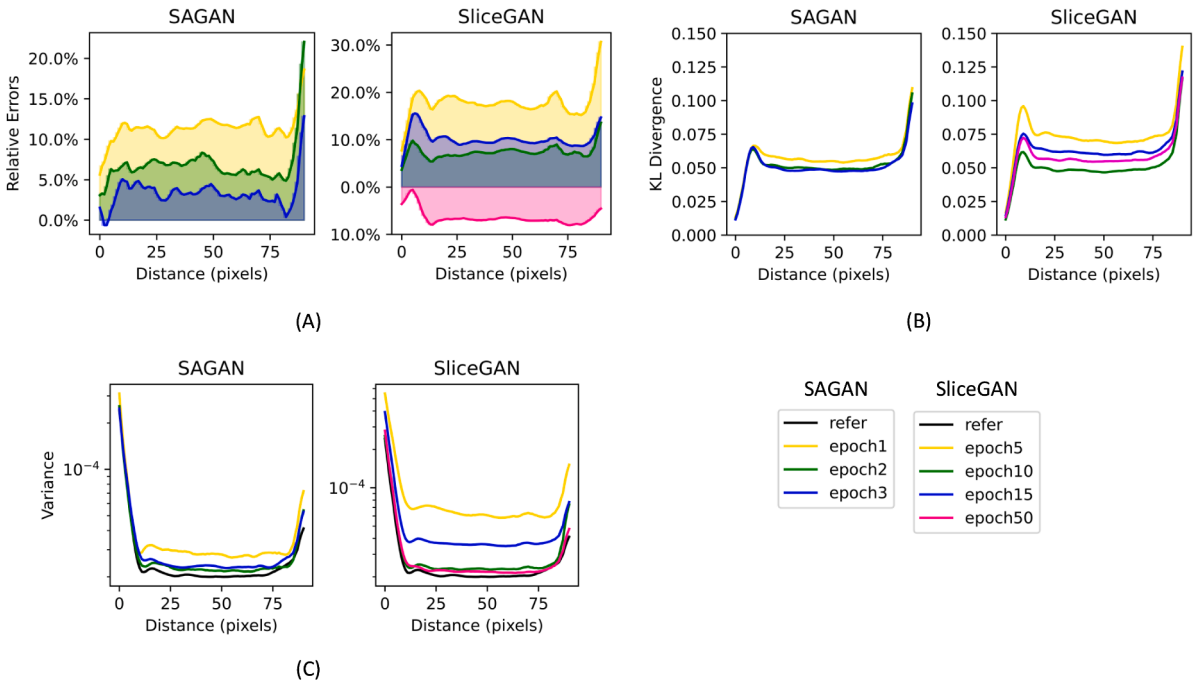

Fig. 7. Fontainebleau sandstone comparison results of different metrics on slice TPCF. (A) Relative errors for SAGAN and SliceGAN at different epochs. (B) Kullback-Leibler (KL) divergence for SAGAN and SliceGAN at different epochs. (C) Variance for SAGAN and SliceGAN at different epochs.

during training, where SAGAN converges in training accuracy and SliceGAN exhibits fluctuations, resulting in greater distribution differences by the 50th epoch. The TPCCF comparison results of different metrics are similar to the TPCF results, with SAGAN demonstrating a faster learning speed. Which is shown in Fig. 11. In terms of relative error, SAGAN's highest error in the first epoch is only 80 %, significantly lower than SliceGAN's in the fifth epoch. Throughout the training process, SAGAN's slice TPCCF explicitly and steadily approaches the reference, while SliceGAN shows fluctuations. SliceGAN achieves its optimal solution at the 10th epoch. However, at the 15th and 50th epochs, the relative error increases and continues to diverge. In terms of variance and KL divergence, SliceGAN shows fluctuations around the optimal result after the 10th epoch, but it does not significantly worsen.

Compared to SliceGAN, which uses a purely CV loss function, SAGAN achieves better accuracy early in the training process due to the consideration of planar descriptors. While SliceGAN can capture the size and connectivity information of the pore phase after 10 h, the lack of descriptor constraints leads to poorer results in the later stages of training.

3.1.2. Voxel descriptor verify

Although introducing slice descriptors as additional constraints does not directly constrain voxel descriptors, the geometric information in three-dimensional space can be accurately described by multiple consecutive two-dimensional slices in various directions. The slice descriptor loss leads to the generation of more accurate two-dimensional microstructures, and through the integration of multiple directions, the resulting three-dimensional microstructures are also more accurate. This accuracy is reflected in the voxel descriptors. To validate the reconstruction quality of voxel descriptors, we used TPCF to measure global void phase information, PSD to assess short-range void phase size information, and the Surface Void Correlation Function (SVCF) to evaluate the relationship between the two phases.

The voxel TPCF results are shown in Fig. 12. Similar to the slice descriptors mentioned above, SAGAN captures 3D geometric patterns and statistical information faster than SliceGAN. In Fig. 12(C), SliceGAN again overestimates the spatial density and exhibits more regular distribution characteristics in the void phase. Comparing Fig. 12(E) and (F), the relative error of SliceGAN is about three times larger than that of SAGAN. Throughout the training process, SAGAN shows a steady convergence toward the reference, while SliceGAN reaches its best solution at epoch 10 but fluctuates in the later epochs, becoming over-optimized by epoch 50.

The results for short-range pore size information are shown in Fig. 13. Similar to the slice descriptor results, both methods capture the void phase geometry pattern from the early stage of training. However, the pore phase in SliceGAN's 10-h reconstruction results exhibits a greater quantity and larger sizes compared to SAGAN outputs with equivalent training duration. Temporally, SAGAN maintains its characteristic accuracy convergence in voxel descriptors throughout training, while SliceGAN demonstrates significant deviation divergence at epoch 15. SINN's results remain substantially different from both methods and are consequently excluded from comparative metrics analysis.

The SVCF results are shown in Fig. 14 for the interactive information between the void and solid phases. This function provides statistical information about the surface structure and uniformity of void distribution. Distribution analysis of reconstruction results

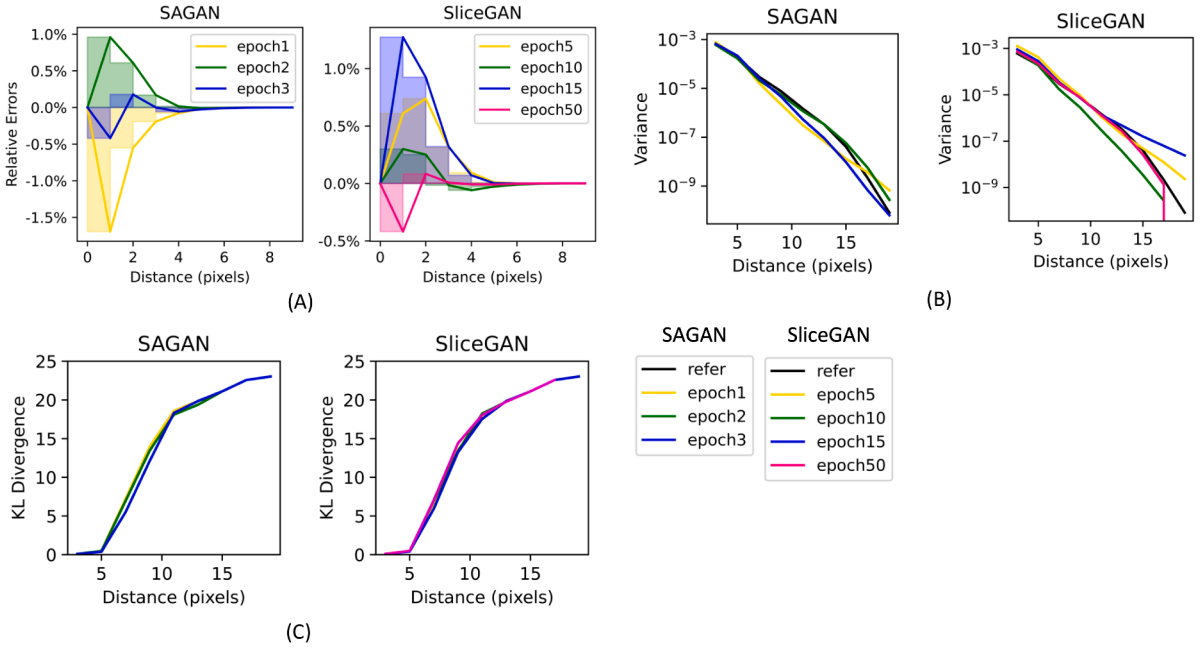

Fig. 8. Fontainebleau sandstone comparison results of different metrics on slice PSD. (A) Relative errors for SAGAN and SliceGAN at different epochs. (B) Variance for SAGAN and SliceGAN at different epochs. (C) Kullback-Leibler (KL) divergence for SAGAN and SliceGAN at different epochs.

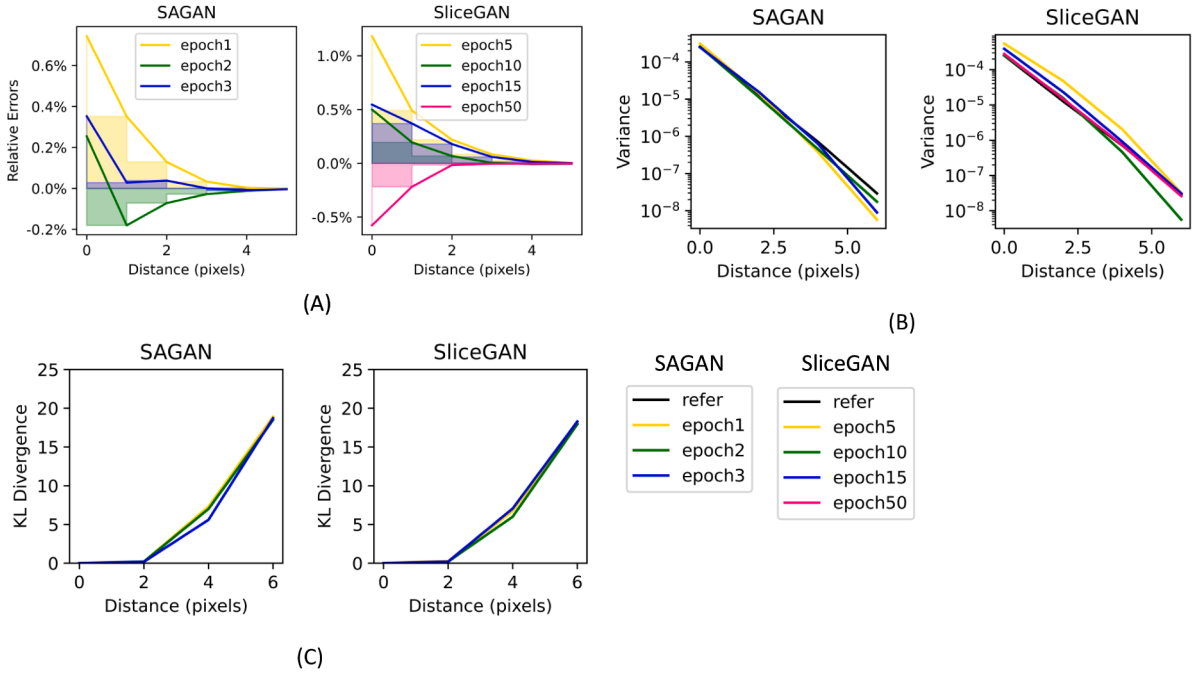

Fig. 9. Fontainebleau sandstone comparison results of different metrics on slice LPF. (A) Relative errors for SAGAN and SliceGAN at different epochs. (B) Variance for SAGAN and SliceGAN at different epochs. (C) Kullback-Leibler (KL) divergence for SAGAN and SliceGAN at different epochs.

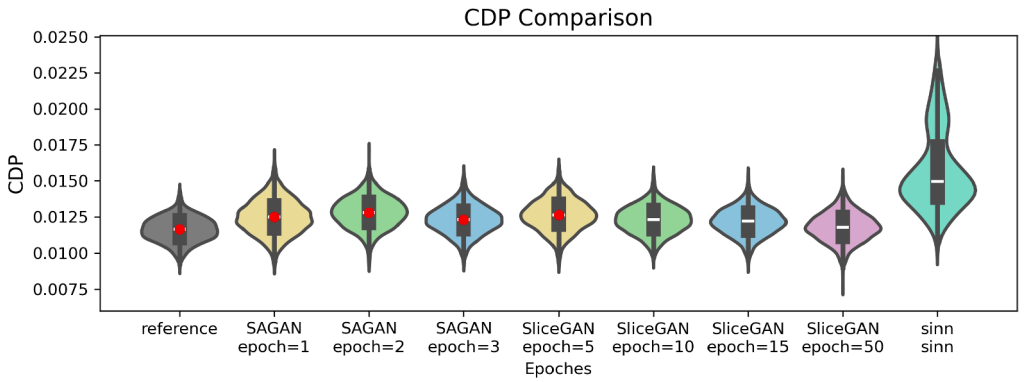

Fig. 10. Fontainebleau sandstone CDP results. The training accuracy of both methods is similar except for SINN, but SliceGAN exhibits larger errors towards the end of the training period.

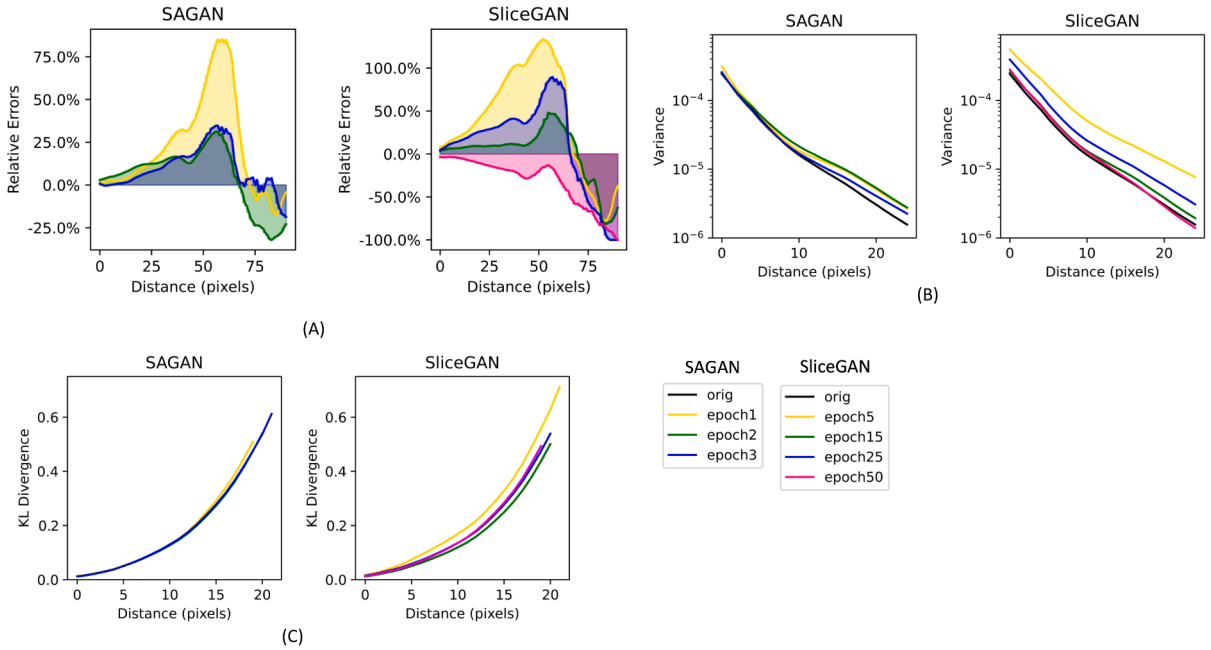

Fig. 11. Fontainebleau sandstone comparison results of different metrics on slice two-point cluster correlation function (TPCCF) results. (A) Relative errors for SAGAN and SliceGAN at different epochs. (B) Variance for SAGAN and SliceGAN at different epochs. (C) Kullback-Leibler (KL) divergence for SAGAN and SliceGAN at different epochs.

reveals that SINN still requires post-processing filtration due to significant accuracy discrepancies. While both SAGAN and SliceGAN outputs conform to the training images' distribution, SAGAN demonstrates progressive accuracy convergence with increased training duration through its descriptor loss mechanism. In contrast, although SliceGAN achieves relatively high accuracy at epochs 10 and 15, it exhibits a large error at epoch 50, indicating insufficient constraint capability in objectives when using pure CV loss functions.

The microstructure determines the physical properties of heterogeneous materials. At the macroscopic scale, permeability is a critical mechanical property that dictates the fluid transport efficiency of Fontainebleau sandstone as a reservoir rock, playing a key role in hydrocarbon and groundwater extraction. At the microscopic scale, the permeability of Fontainebleau sandstone is intrinsically linked to the uniform cementation of quartz grains and a highly interconnected pore network, which collectively contribute to the rock's superior mechanical strength and stability.

To examine the impact on physical properties, we employed OpenFOAM as the simulation solver for the images produced by the three methods, with tetrahedral meshes count approximately 2 million. The $x = 0$ surface was uniformly set as the fluid inlet, and the $x = 128$ surface as the outlet, with a pressure difference of 1000 Pa between them. Water was used as the fluid medium for steady-state permeability analysis. The absolute permeability of the original sample (based on simulation and volume fraction correction for using only the inlet-outlet interconnected region) was 1.00972E+5 milliDarcy. In comparison, SliceGAN yielded an absolute permeability

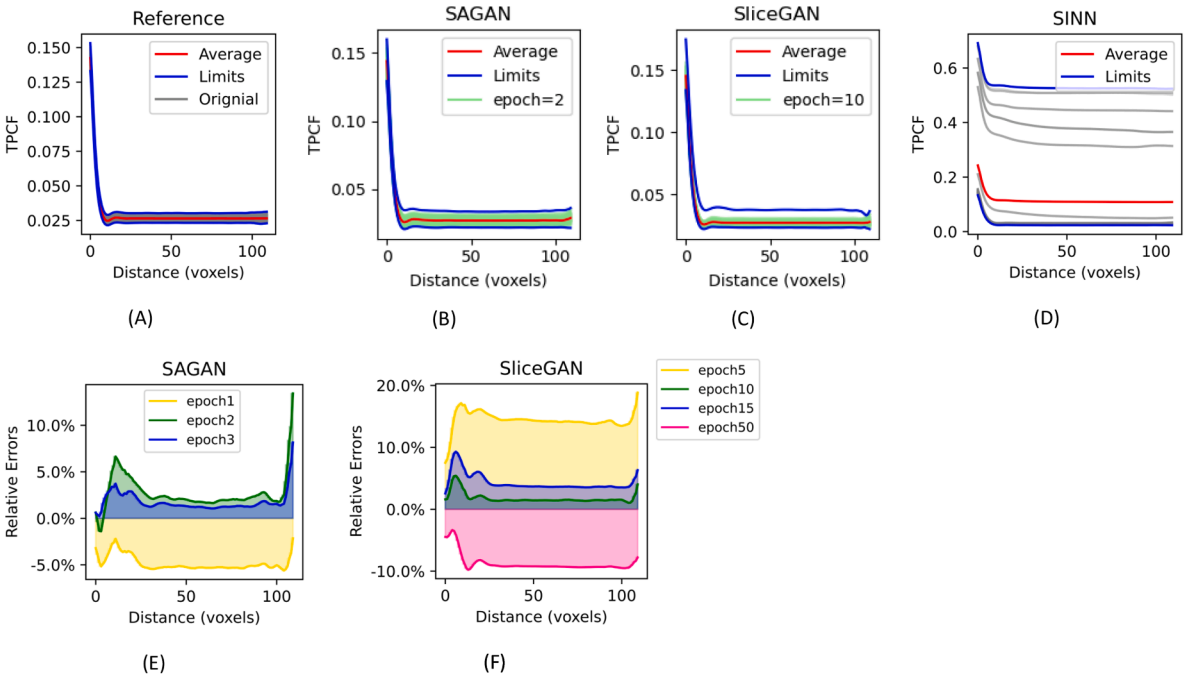

Fig. 12. Fontainebleau sandstone voxel Two-Point Correlation Function (TPCF) results. During the training process of SAGAN, the accuracy of the reconstructed 3D descriptors progressively improves. SliceGAN, lacking explicit constraints, starts to diverge in its results after 10 h of training. (A) Reference. (B) SAGAN, epoch 2. (C) SliceGAN, epoch 10. (D) SINN. (E) Relative errors of SAGAN at different epochs. (F) Relative errors of SliceGAN at different epochs.

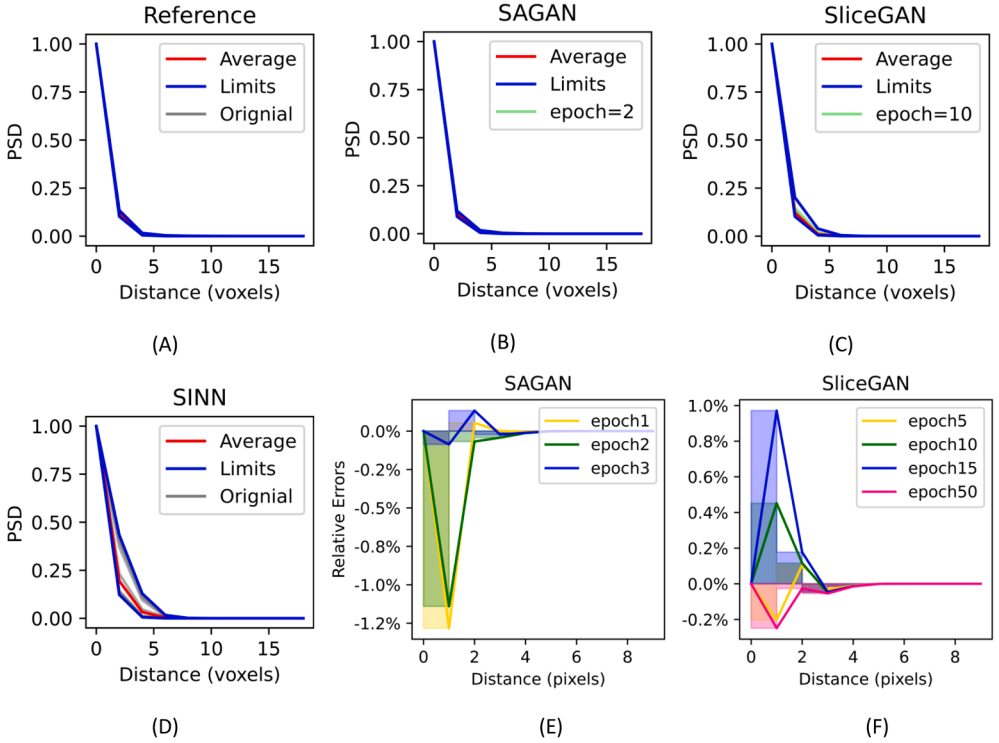

Fig. 13. Fontainebleau sandstone voxel Pore Size Distribution (PSD) results. SAGAN achieves good accuracy in the first epoch, whereas SliceGAN is less efficient, achieving good accuracy only by the 50th epoch. (A) Reference. (B) SAGAN, epoch 2. (C) SliceGAN, epoch 10. (D) SINN. (E) Relative errors of SAGAN at different epochs. (F) Relative errors of SliceGAN at different epochs.

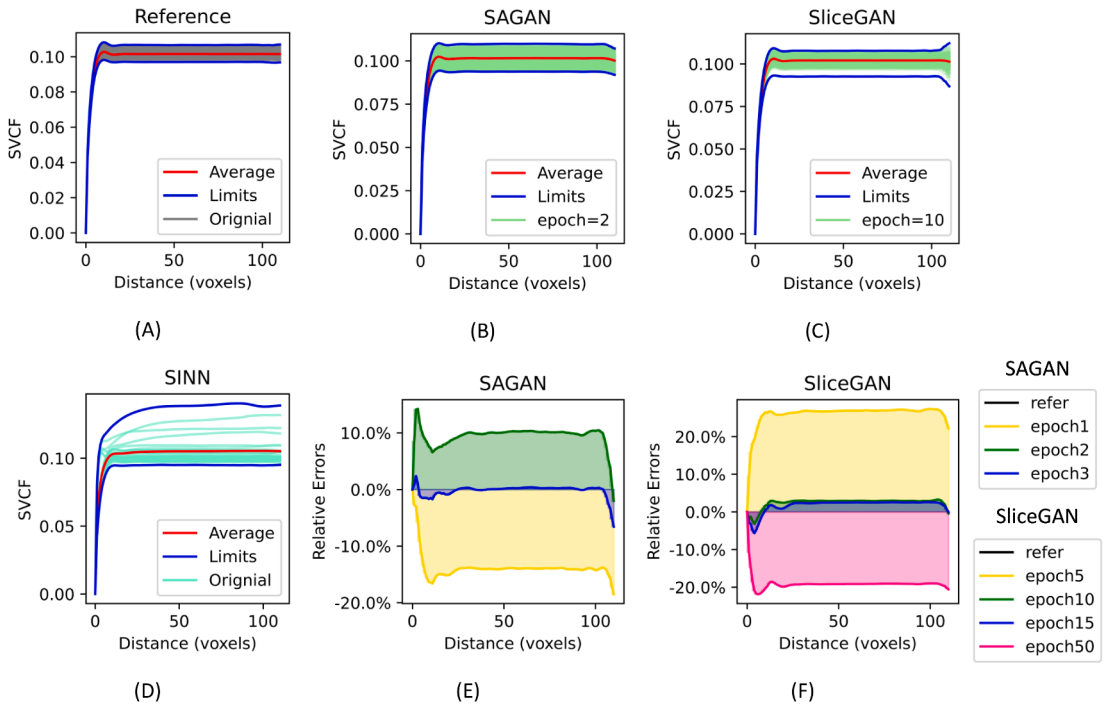

Fig. 14. Fontainebleau sandstone voxel Surface Void Correlation Function (SVCF) results. SAGAN achieves good accuracy in the first epoch, with training results progressively optimizing. In contrast, SliceGAN starts to diverge after the 10th epoch. (A) Reference. (B) SAGAN, epoch 2. (C) SliceGAN, epoch 10. (D) SINN. (E) Relative errors of SAGAN at different epochs. (F) Relative errors of SliceGAN at different epochs.

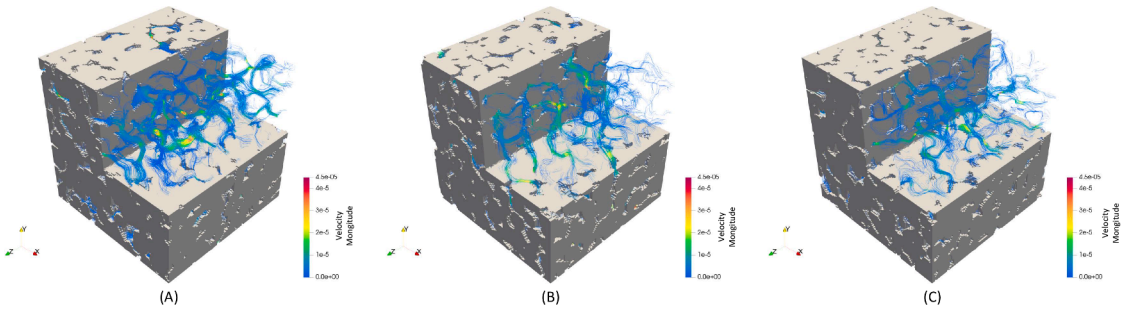

Fig. 15. The fluid velocity fields at steady state obtained from OpenFOAM simulations: (a) Original microstructure; (b) The representative microstructure from SAGAN; (c) The representative microstructure from SliceGAN.

of 1.5007E+5 milliDarcy (relative error: 48.63 %), while SAGAN resulted in 8.9377E+4 milliDarcy (relative error: −11.48 %). The results are recorded in Fig. 15.

For the voxel descriptor verification, since the descriptor loss processes consecutive two-dimensional slices of the three-dimensional microstructure along the three principal axes, the geometric structure of the two-dimensional slices is closer to the two-dimensional reference. Therefore, compared to SliceGAN, SAGAN demonstrates faster learning and more explicit, stable descriptor accuracy, especially for long-range information. This stable accuracy improvement with training duration holds significant practical value in engineering applications. SAGAN ensures progressive accuracy enhancement without mode collapse as training progresses, while SliceGAN lacks deterministic predictability regarding which training epoch yields optimal reconstruction accuracy.

3.2. Case 2 - Polytetrafluoroethylene

An anisotropic two-phase microstructure is considered in the second case. PTFE is a synthetic fluoropolymer of tetrafluoroethylene, widely used due to its chemical inertness. As shown in Fig. 16, it consists of fibre (the white phase) and air (the black phase) in a two-phase segmented anisotropic microstructure image. Generally speaking, reconstructing anisotropic 3D microstructures from 2D slices is challenging. We follow the anisotropic reconstruction algorithm introduced in Section 2.5, using three discriminators and three descriptor ResNets from three orthogonal cross-sections. YZ cross-section slices appear as multiple long, connected fibre lines on

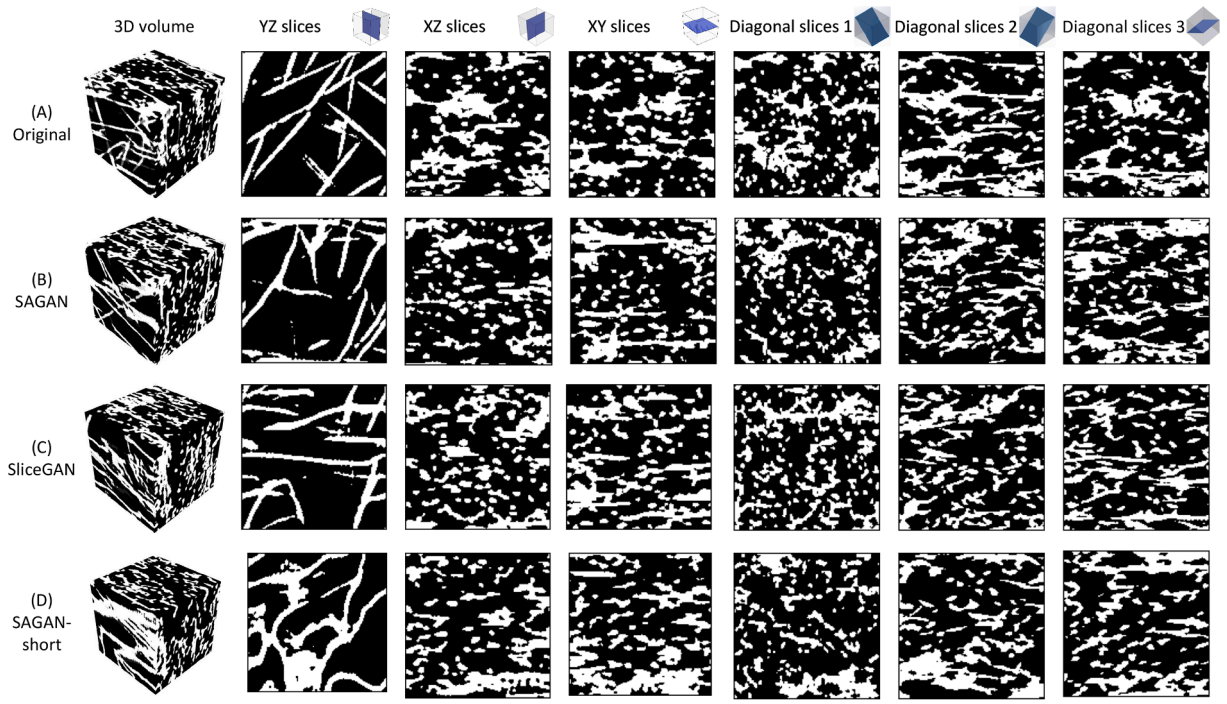

Fig. 16. Three-dimensional microstructure and cross-sections along the three principal axes and three diagonal slices of Polytetrafluoroethylene (PTFE) reference and reconstruction results. (A) Reference. (B) Synthetic results from SAGAN on epoch 2. (B) Synthetic results from SliceGAN on epoch 50. (D) Synthetic results from SAGAN-short on epoch 2.

Table 3
Comparison of the Mean Absolute Error (MAE) between the average, maximum, and minimum curves of the slice two-point correlation function (TPCF) for different reconstruction methods and the corresponding curves from the training dataset in Case 2.

		MAE		
		Min	Mean	Max
TPCF in Y-Z cross-section	SAGAN	6.10×10^{-3}	9.23×10^{-3}	4.37×10^{-2}
	SliceGAN	5.34×10^{-3}	1.82×10^{-2}	1.65×10^{-1}
	SAGAN_short	5.92×10^{-3}	1.27×10^{-2}	7.46×10^{-2}
TPCF in X-Z cross-section	SAGAN	7.70×10^{-3}	2.79×10^{-3}	4.40×10^{-2}
	SliceGAN	1.12×10^{-2}	1.01×10^{-2}	1.69×10^{-1}
	SAGAN_short	7.92×10^{-3}	3.64×10^{-3}	6.69×10^{-2}
TPCF in X-Y cross-section	SAGAN	6.10×10^{-3}	3.86×10^{-3}	2.77×10^{-2}
	SliceGAN	9.16×10^{-4}	1.31×10^{-3}	5.51×10^{-2}
	SAGAN_short	5.87×10^{-3}	2.56×10^{-3}	4.87×10^{-2}

the plane, while XZ and XY cross-section slices show the same fibre cross-sections as multiple small, discrete fibre dots or connected sections of fibre bundles. The geometric features of these two cross-sections are significantly different, posing a substantial challenge for reconstruction. In Case 1, we have already verified that the training process efficency of SAGAN is faster than that of SliceGAN. In this test, we compare the SAGAN results from epoch 2 with the SliceGAN results from epoch 50. The training time for epoch 2 of SAGAN is 27 h in total, which includes 12 h for the core SAGAN training and an additional 15 h for the descriptor ResNet training (5 h per orthogonal direction). In comparison, training epoch 50 with SliceGAN requires 48 h. There is no significant difference in visual comparison between the three-dimensional microstructure and the two-dimensional slices. Additionally, to examine the superiority of the five descriptors selected in this paper, we reduce them to three descriptors as conditional inputs to the SAGAN network for training and compared with SAGAN and SliceGAN. This method is named SAGAN-short.

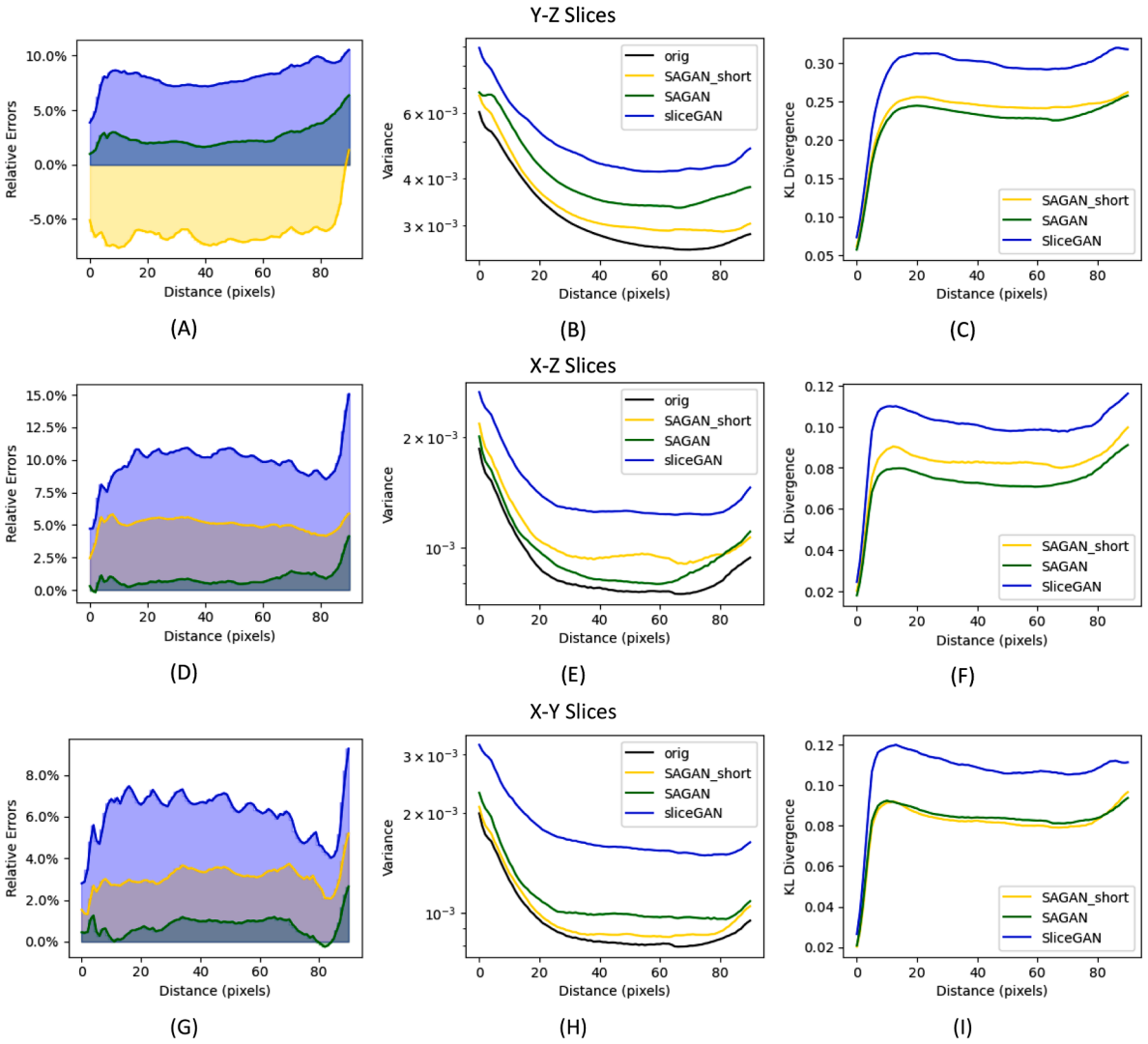

Fig. 17. Polytetrafluoroethylene (PTFE) slice Two-Point Correlation Function (TPCF) comparison results on different metrics. (A) Relative errors of TPCF for the Y-Z cross-section slices. (B) Variance of TPCF for the Y-Z cross-section slices. (C) KL divergence of TPCF for the Y-Z cross-section slices. (D) Relative errors of TPCF for the X-Z cross-section slices. (E) Variance of TPCF for the X-Z cross-section slices. (F) KL divergence of TPCF for the X-Z cross-section slices. (G) Relative errors of TPCF for the X-Y cross-section slices. (H) Variance of TPCF for the X-Y cross-section slices. (I) KL divergence of TPCF for the X-Y cross-section slices.

Introducing three Descriptor ResNets for explicit control in the reconstruction of anisotropic materials still results in better accuracy compared to the purely CV-based SliceGAN. By controlling the KL divergence of the generated distribution, the results more closely match the real distribution of the training database. Due to the explicit training process, SAGAN demonstrates higher time-to-accuracy performance than SliceGAN within the 10-h training period. For SAGAN-short with only three descriptor inputs (TPCF, PSD, LPF), its 2D slice results are similar to SAGAN results on the descriptors used as input conditions, and the results on unused descriptors are not as good as SAGAN. It is noted that both SAGAN results are better than the SliceGAN result. In the comparison of 3D voxel descriptors, SAGAN has a better overall effect than SAGAN-short, confirming the superiority of the five descriptors selected in this paper.

3.2.1. Slice descriptor verify

In this case, the TPCF comparison results analyze the fibre phase with different metrics and are shown in Fig. 17 and the Mean Absolute Error (MAE) of verage, maximum, and minimum curves are compared in Table 3. Regarding Y-Z slice comparisons, both SAGAN variants demonstrate superior accuracy in relative error, variance, and KL divergence metrics compared to SliceGAN, thanks to their explicit descriptor constraints during training. Between the two SAGAN implementations, the version incorporating more descriptor groups achieves higher accuracy. In limit curve analysis, SliceGAN shows closer alignment only with minimum value

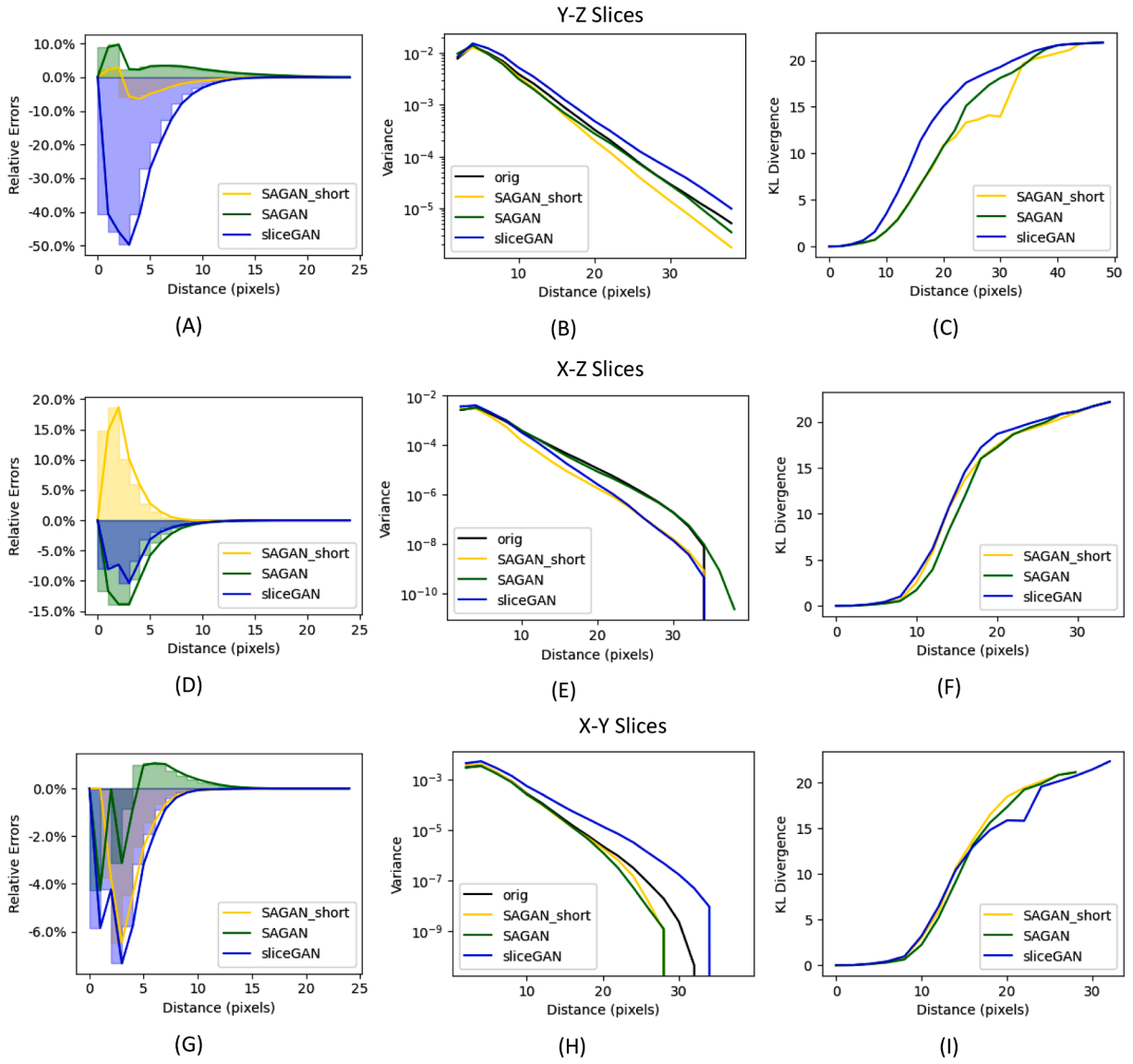

Fig. 18. Polytetrafluoroethylene (PTFE) comparison metric results on the cumulative density function of Pore Size Distribution (PSD). (A) Relative errors of PSD for the Y-Z cross-section slices. (B) Variance of PSD for the Y-Z cross-section slices. (C) KL divergence of PSD for the Y-Z cross-section slices. (D) Relative errors of PSD for the X-Z cross-section slices. (E) Variance of PSD for the X-Z cross-section slices. (F) KL divergence of PSD for the X-Z cross-section slices. (G) Relative errors of PSD for the X-Y cross-section slices. (H) Variance of PSD for the X-Y cross-section slices. (I) KL divergence of PSD for the X-Y cross-section slices.

curves, though all three methods exhibit satisfactory accuracy on average curves. This performance hierarchy (SAGAN > SAGAN-Short > SliceGAN) persists in X-Z and X-Y slices, with a notable exception: SAGAN-Short marginally outperforms standard SAGAN in variance metrics for X-Z slices and KL divergence for X-Y slices, albeit by negligible margins. All methods maintain good accuracy in curve MAE comparisons. Additionally, the slice TPCF results of each method are shown in Fig. A.3 at Appendix.

In terms of the connectivity and size of the fibre phase, both methods demonstrate better accuracy for the short-range geometric characteristics of the fibre phase's PSD and LPF compared to the long-range geometric characteristics. The comparative results measured by PSD are shown in Fig. 18 and the comparison in terms fo LPF are shown in Fig. 19. Firstly, in the results of the Y-Z cross-section slices, both SAGAN and SAGAN-short perform better than SliceGAN, and SAGAN-short achieves the best results in relative error and KL divergence. Secondly, for the X-Z cross-section slices, SliceGAN has a smaller relative error, SAGAN and SAGAN-short get similar performance on variance and KL divergence. On the X-Y cross-section, all methods have the same level of KL divergence, but the sliceGAN shows worse performance on relative error and variance. This is because SliceGAN exhibited greater variability, with anomalous generated results showing very large pore sizes, as reflected in the maximum limit, which can be found in Figs. A.4

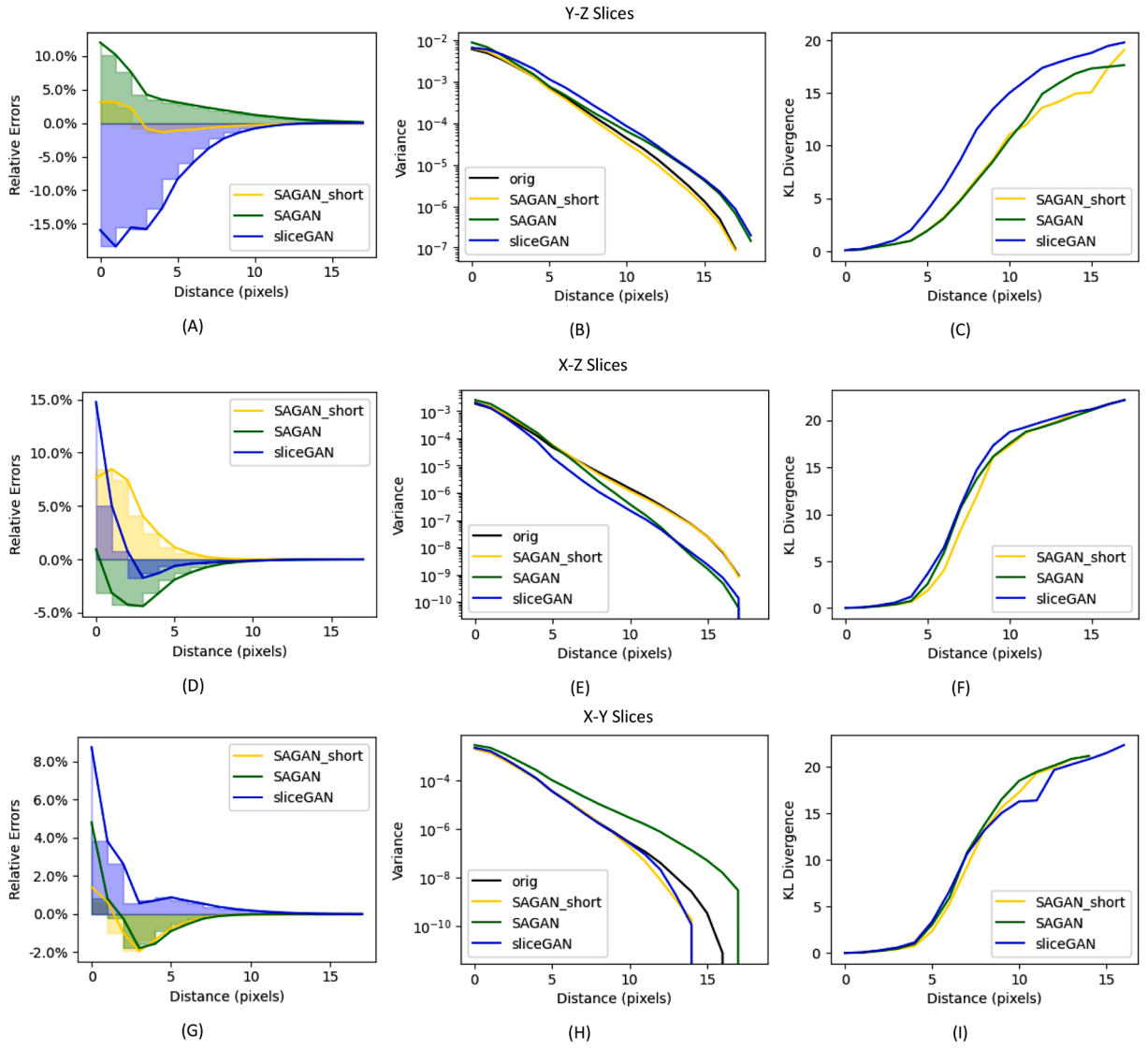

Fig. 19. Polytetrafluoroethylene (PTFE) comparison metric results on the probability density function of Lineal Path Function (LPF). (A) Relative errors of LPF for the Y-Z cross-section slices. (B) Variance of LPF for the Y-Z cross-section slices. (C) KL divergence of LPF for the Y-Z cross-section slices. (D) Relative errors of LPF for the X-Z cross-section slices. (E) Variance of LPF for the X-Z cross-section slices. (F) KL divergence of LPF for the X-Z cross-section slices. (G) Relative errors of LPF for the X-Y cross-section slices. (H) Variance of LPF for the X-Y cross-section slices. (I) KL divergence of LPF for the X-Y cross-section slices.

and A.5 in the Appendix. Although the SAGAN variants incorporating short-range descriptors show diminished advantage in X-Z slice results compared to previous benchmarks, they demonstrate marked accuracy improvements in both X-Y and Y-Z slices, yielding an overall accuracy enhancement. Comparative analysis confirms that expanding the descriptor set - utilizing more independent descriptor groups as inputs and within the loss function - significantly elevates reconstruction accuracy in the SAGAN framework.

For the global CDP information shown in Fig. 20, both methods produce accurate results except for SliceGAN in the Y-Z cross-section slices, where the probability of fibre generation is relatively high. In the TPCCF results shown in Fig. 21, SAGAN and SAGAN-short demonstrate significantly better accuracy than SliceGAN in terms of relative error, and show slightly better variance and KL divergence accuracy than SliceGAN. SAGAN performs better than SAGAN-short because TPCCF is not an input condition of SAGAN-short.

For the slice descriptor verification of the anisotropic material, SliceGAN tends to generate more fibre sections in all three directions, with greater distribution variability. Both methods achieve good results in short-range size and connectivity of the fibre phase, but SliceGAN tends to produce smaller and less connected fibres. However, neither method can replicate the high clustering observed in the reference. In summary, the purely CV-based SliceGAN tends to capture finer details, but it lacks the ability to learn long-range and global information as effectively as SAGAN and SAGAN-short, which is constrained by slice descriptors. Regarding descriptor

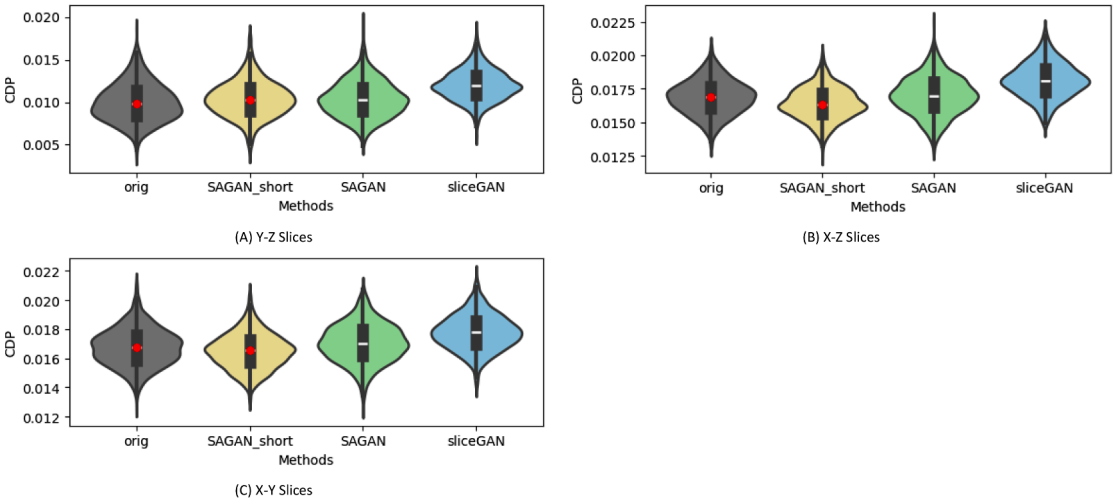

Fig. 20. Polytetrafluoroethylene (PTFE) comparison metric results on the slice Core Distribution Probability (CDP). (A) Y-Z cross-section slices. (B) X-Z cross-section slices. (C) X-Y cross-section slices.

Table 4
Mean Absolute Error (MAE) of voxel Two-Point Correlation Function (TPCF) Curves for Polytetrafluoroethylene (PTFE).

		TPCF		
		SAGAN	SliceGAN	SAGAN_short
MAE	Min	1.38×10^{-2}	3.20×10^{-2}	1.90×10^{-2}
	Mean	3.14×10^{-3}	1.54×10^{-3}	3.70×10^{-2}
	Max	1.80×10^{-2}	9.75×10^{-2}	2.10×10^{-2}

group configurations, the five-group variant demonstrates superior precision across most reconstruction criteria, with particularly evident advantages in global feature characterization as manifested by Two-Point Correlation Function (TPCF) metrics.

3.2.2. *Voxel descriptor verify*

For anisotropic materials, the geometric shapes in three-dimensional space are more complex than those of isotropic materials, making it more challenging to constrain them through a large number of two-dimensional plane descriptors. Theoretically, slice descriptors in the three principal directions can sufficiently constrain the spatial geometric patterns, leading to better accuracy in voxel descriptors. In the voxel descriptor verification part, we use TPCF to evaluate global information and LPF to assess short-range information.

As shown in Fig. 22, both SAGAN and SAGAN-short perform better than SliceGAN on voxel TPCF. Among them, SAGAN performs best. There are significant differences in the distribution: SAGAN's results have smaller variance, while SliceGAN tends to generate results with larger variance. This is also reflected in the relative error; although all methods produced more fiber phase, SAGAN has a smaller relative error and higher accuracy in KL divergence. Additionally, we quantified the differences between the three reconstruction methods and the reference curves using the Mean Absolute Error (MAE) for the minimum, mean, and maximum value curves of the TPCF. The results are shown in Table 4.

For voxel LPF results shown in Fig. 23, all three reconstruction methods exhibit primary error concentrations in short-range regions (voxel < 8). Regarding relative error, the incorporation of descriptors significantly enhances reconstruction accuracy, with the five-group comprehensive descriptor configuration demonstrating superior constraint effectiveness compared to the three-group variant. This advantage extends consistently to variance and KL divergence metrics as well.

Thermal conductivity is a critical macroscopic parameter determining the insulation and heat resistance of PTFE, directly influencing its thermal stability and energy efficiency in high-temperature environments. At the microscopic scale, the low thermal conductivity is intrinsically linked to PTFE's weak crystallinity and disordered molecular structure, which restricts heat transfer while enhancing its dielectric strength and chemical inertness. Here, we validate the accuracy of the reconstruction methods through thermal conduction simulation. Comparing SliceGAN and SAGAN with five descriptor inputs, the results are presented in Fig. 24. In our simulations, heat transfer occurs solely through the solid phase with a thermal conductivity of $\lambda_0 = 1.0W/(m \cdot K)$, while the pore space filled with air is treated as thermally insulating. Tetrahedral meshes with approximately 2 million elements were generated for the solid phase of the three methods. Using OpenFOAM, steady-state thermal conduction simulations were performed with an inlet temperature of 300 K at $x = 0$ and an outlet temperature of 350 K at $x = 128$. The effective thermal conductivity was evaluated

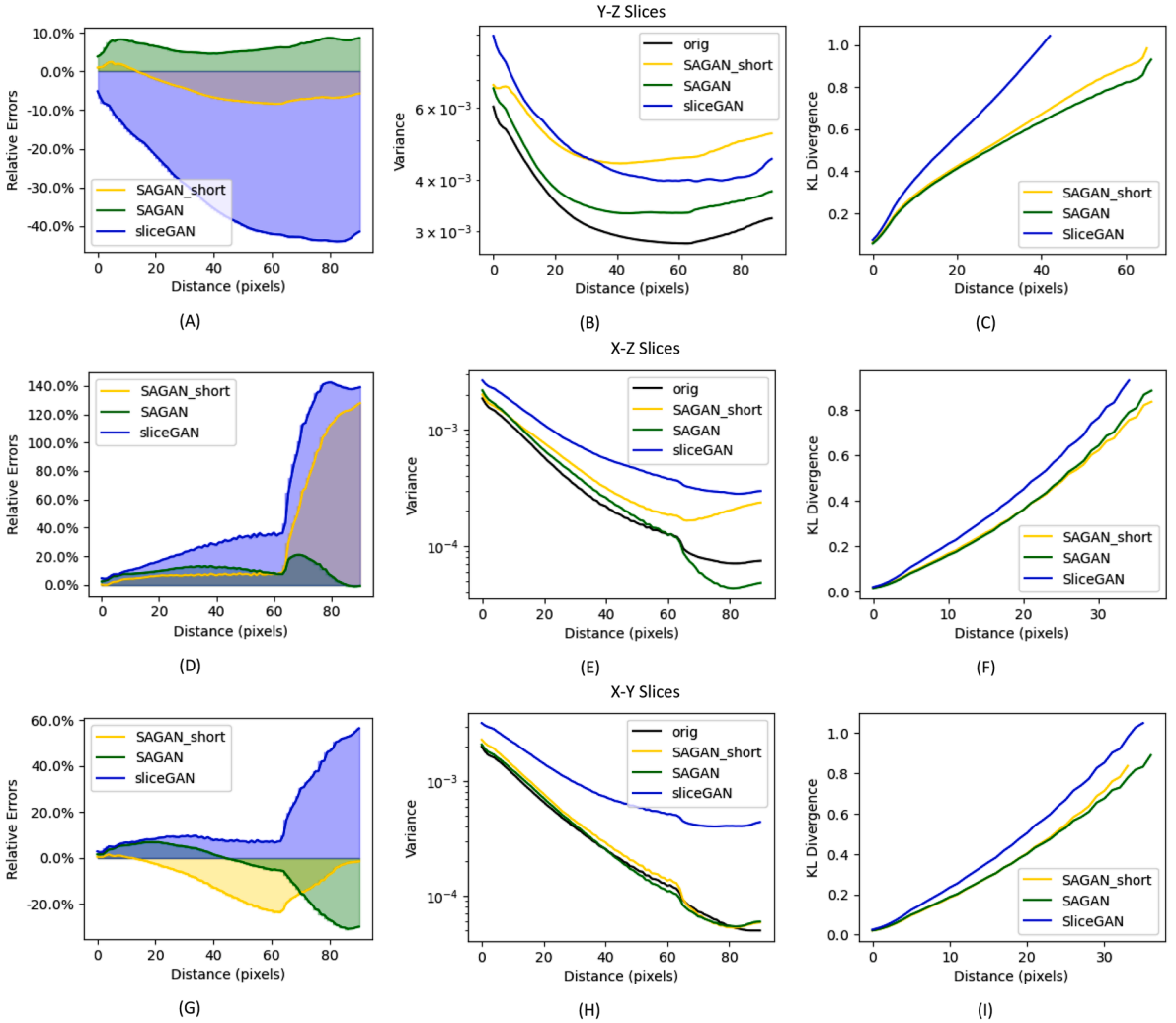

Fig. 21. Polytetrafluoroethylene (PTFE) slice Two-Point Cluster Correlation Function (TPCCF) results. (A) Relative errors of TPCCF for the Y-Z cross-section slices. (B) Variance of TPCCF for the Y-Z cross-section slices. (C) KL divergence of TPCCF for the Y-Z cross-section slices. (D) Relative errors of TPCCF for the X-Z cross-section slices. (E) Variance of TPCCF for the X-Z cross-section slices. (F) KL divergence of TPCCF for the X-Z cross-section slices. (G) Relative errors of TPCCF for the X-Y cross-section slices. (H) Variance of TPCCF for the X-Y cross-section slices. (I) KL divergence of TPCCF for the X-Y cross-section slices.

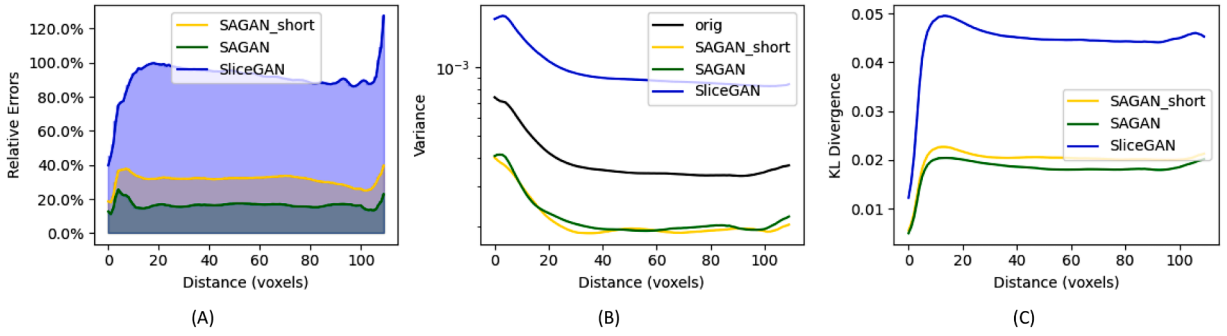

Fig. 22. Polytetrafluoroethylene (PTFE) voxel Two-Point Correlation Function (TPCF) results on different metrics. (A) Relative errors. (B) Variance. (C) KL divergence.

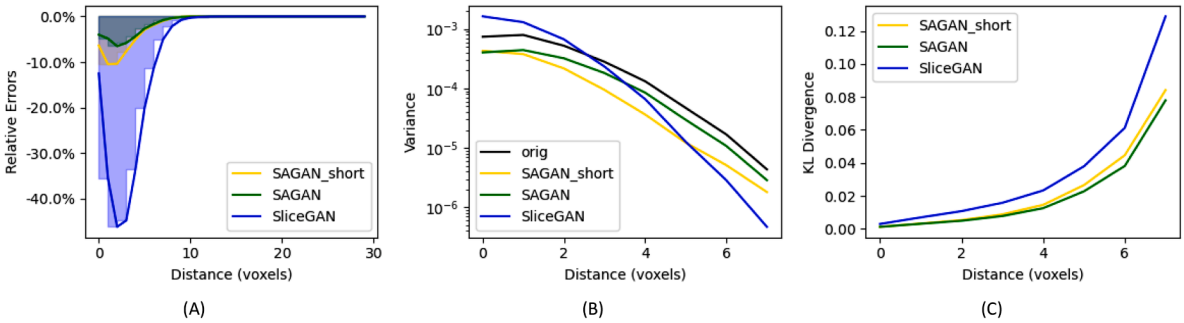

Fig. 23. Polytetrafluoroethylene (PTFE) voxel Lineal Path Function (LPF) results on different metrics. (A) Relative errors. (B) Variance. (C) KL divergence.

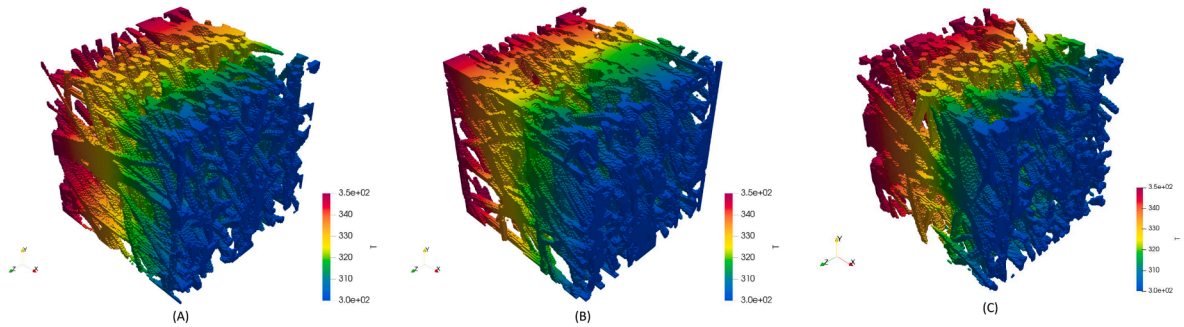

Fig. 24. The temperature fields of thermal conductivity from OpenFOAM simulation: (A) Original microstructure; (B) The representative microstructure from SAGAN; (C) The representative microstructure from SliceGAN.

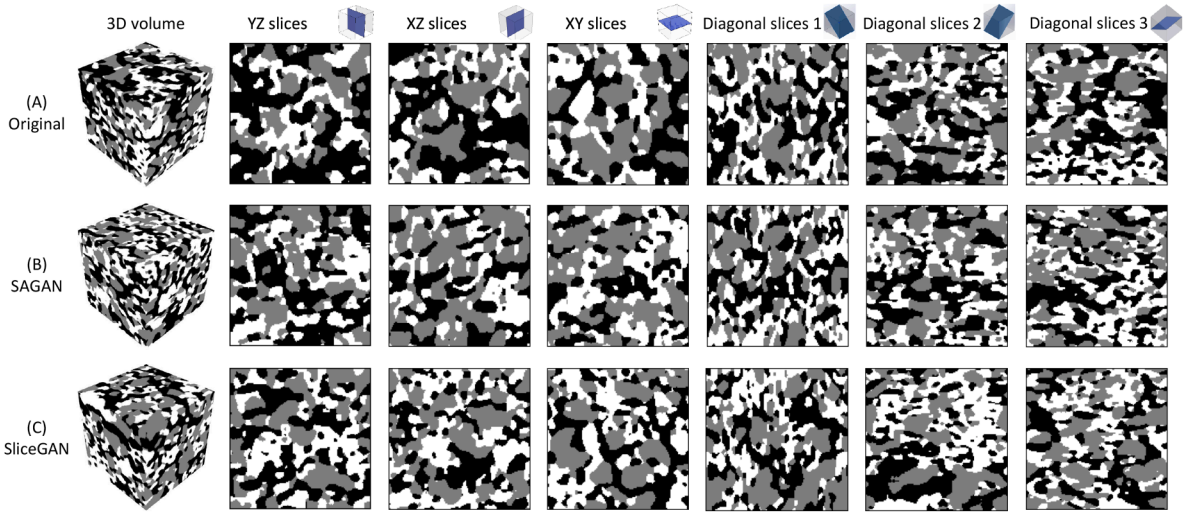

Fig. 25. Three-dimensional microstructure and cross-sections along the three principal axes and three diagonal slices of cathode materials for Solid Oxide Fuel Cell (SOFC) reference and reconstruction results. (A) Reference. (B) synthetic results from SAGAN on epoch 2. (C) synthetic results from SliceGAN on epoch 50.

based on Fourier's law. The effective thermal conductivity of the original sample was 8.285E-2 $W/(m \cdot K)$. In comparison, SliceGAN yielded an effective thermal conductivity of 9.831E-2 $W/(m \cdot K)$ (relative error: 18.64 %), while SAGAN resulted in 7.812E-2 $W/(m \cdot K)$ (relative error: −5.72 %).

For the voxel descriptor verification results, introducing slice descriptors provides better constraints on geometric information, and overcomes the overgeneration and undergeneration issues present in purely CV-based methods. A purely CV-based discriminator cannot comprehensively account for both details and global information. SAGAN achieves better overall accuracy by constraining the geometric patterns with planar descriptors. However, this constraint also results in slightly less variability in the generated results

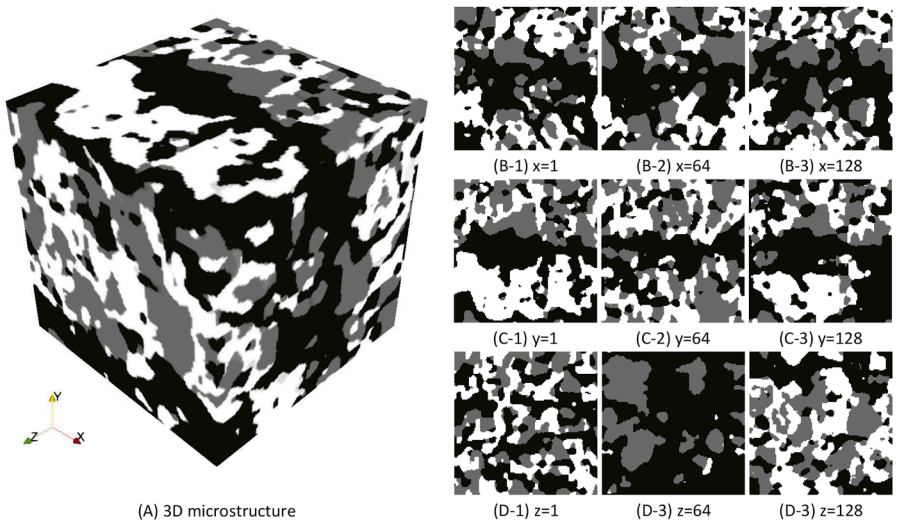

Fig. 26. First synthetic of epoch 50 from SliceGAN. The black phase exhibits an overgeneration problem, with an excessively large amount of connected black phase present in the middle part of the Z direction.

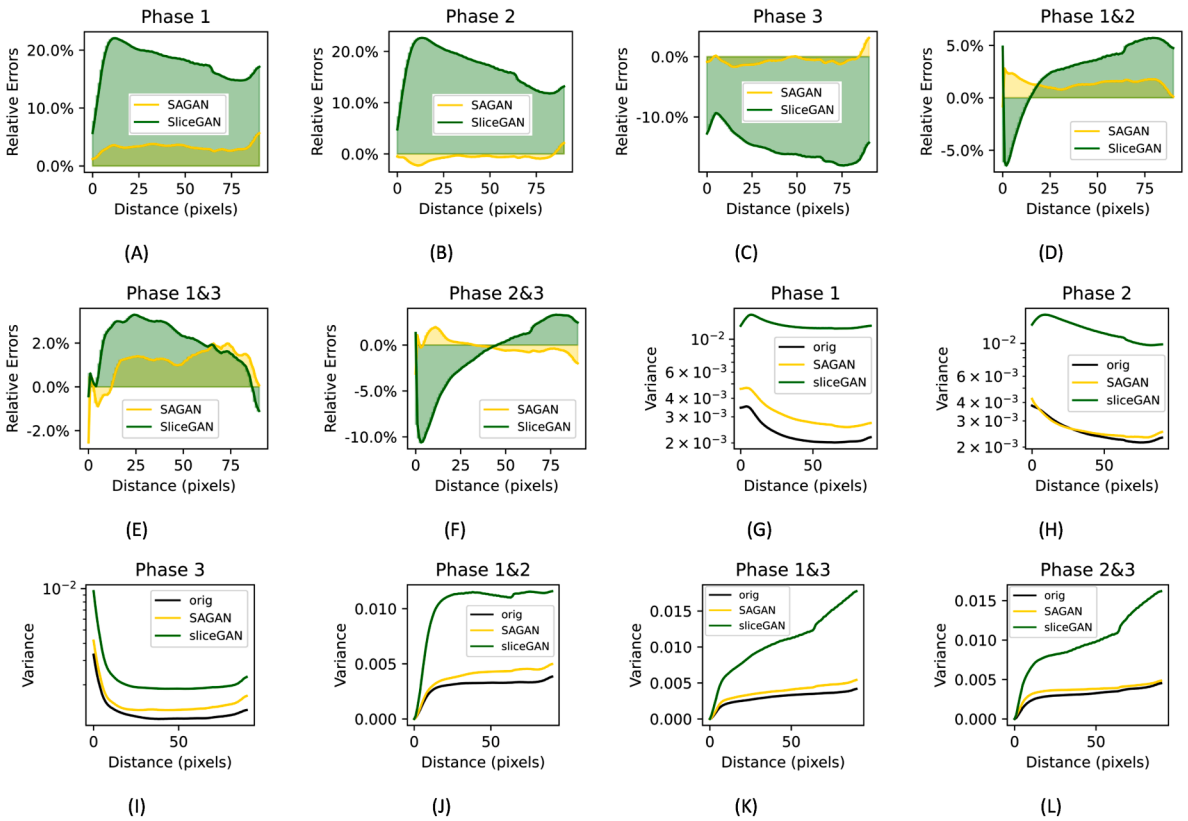

Fig. 27. Relative error and variance of the TPCF on cathode material for SOFC. (A) Relative error of phase 1. (B) Relative error of phase 2. (C) Relative error of phase 3. (D) Relative error of phase 1 and phase 2. (E) Relative error of phase 1 and phase 3. (F) Relative error of phase 2 and phase 3. (G) Variance of phase 1. (H) Variance of phase 2. (I) Variance of phase 3. (J) Variance of phase 1 and phase 2. (K) Variance of phase 1 and phase 3. (L) Variance of phase 2 and phase 3.

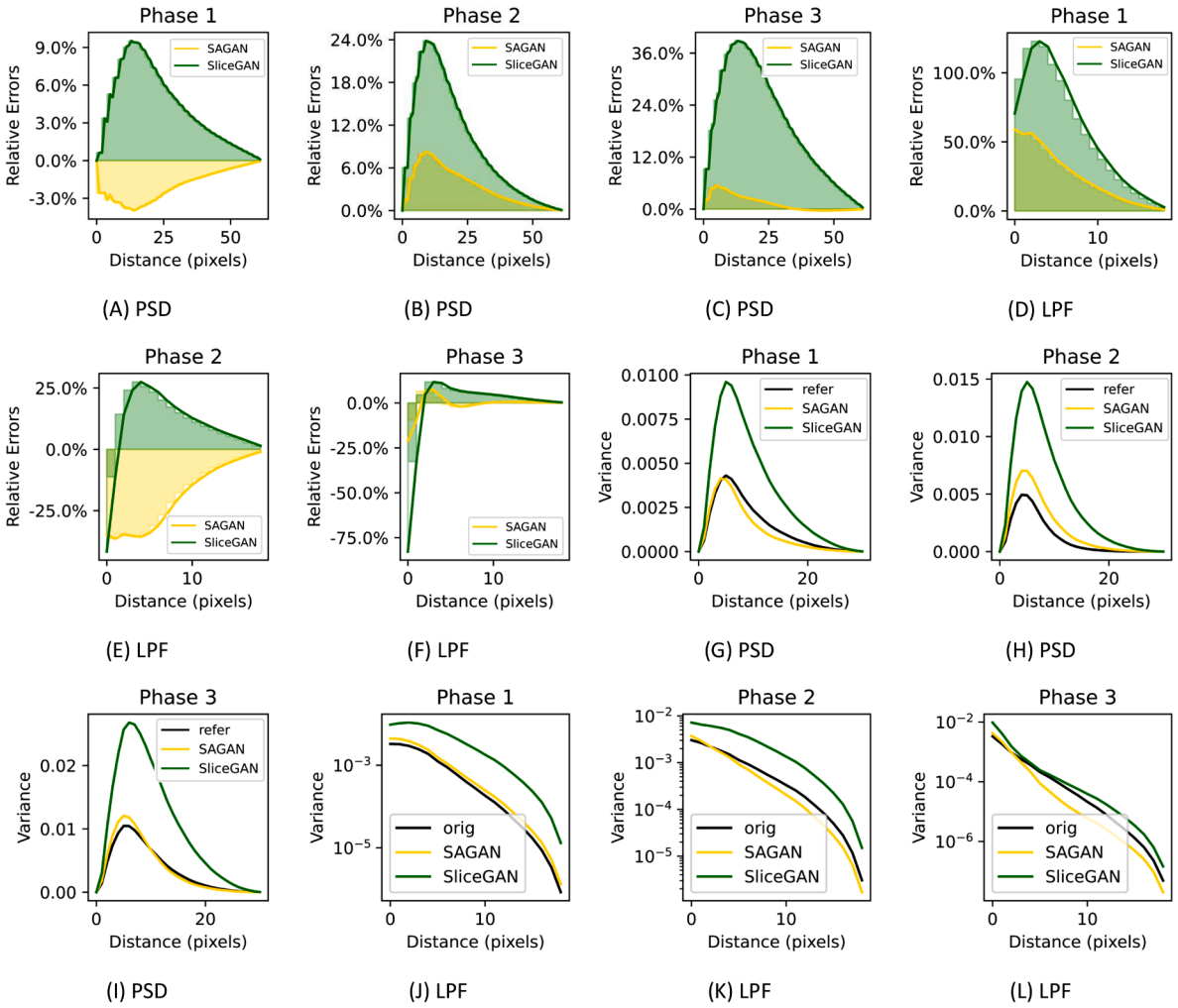

Fig. 28. Relative error and variance of slice Pore Size Distribution (PSD) and Lineal-Path Function (LPF) on cathode material for SOFC. (A) Relative error of PSD in phase 1. (B) Relative error of PSD in phase 2. (C) Relative error of PSD in phase 3. (D) Relative error of LPF in phase 1. (E) Relative error of LPF in phase 2. (F) Relative error of LPF in phase 3. (G) Variance of PSD in phase 1. (H) Variance of PSD in phase 2. (I) Variance of PSD in phase 3. (J) Variance of LPF in phase 1. (K) Variance of LPF in phase 2. (L) Variance of LPF in phase 3.

compared to the reference. At the same time, the results of SAGAN-short demonstrate the advantages of the five selected descriptors. When the number of descriptors is reduced, the overall effect of the generation will be compromised. Similar to the isotropic results in Case 1, both methods demonstrate good learning capabilities for short-range information. However, considering that SAGAN's training time is only 27 h(12 h SAGAN training and 15 h training for 3 ResNets), it has an unparalleled advantage in time-to-accuracy performance.

3.3. Case 3 - Cathode materials for solid oxide fuel cell(SOFC)

3.3.1. Slice descriptor verify

The third dataset is an isotropic three-phase material, the SOFC cathode. In this test, the microstructure consists of three phases: (La, Sr) MnO3 (LSM), yttria-stabilized zirconia (YSZ), and the gas pore phase. We compared the original image with the second epoch results from SAGAN, which takes around 15 h to train plus another 15 h ResNet training, and the 50th epoch results from SliceGAN, which takes around 50 h. A visual inspection can be seen in Fig. 25. From the representative 2D slices, all three phases exhibit very complex geometric characteristics. Even within the same slice, large areas of complex interconnected phases, medium-sized blocks, and very small particles exist. The randomness of this case makes it challenging for the pure CV-based method to reconcile different geometric patterns in one reconstruction result. Though the 3D synthetic images from SliceGAN are similar to the reference, some reconstructed results severely overgenerate certain features. This issue is even more pronounced in the 2D slices, where some features are overgenerated while others are undergenerated. For example, the first synthetic images of epoch 50 from SliceGAN

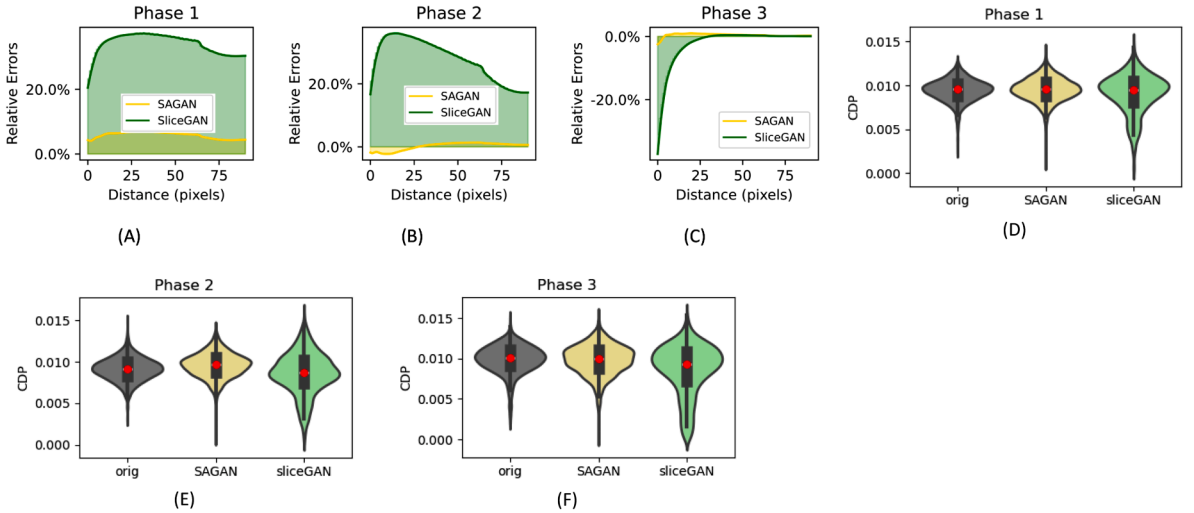

Fig. 29. CDP and the relative error on TPCCF of cathode material for SOFC. (A) Relative error of TPCCF in phase 1. (B) Relative error of TPCCF in phase 2. (C) TPCCF in phase 3. (D) CDP in phase 1. (E) CDP in phase 2. (F) CDP in phase 3.

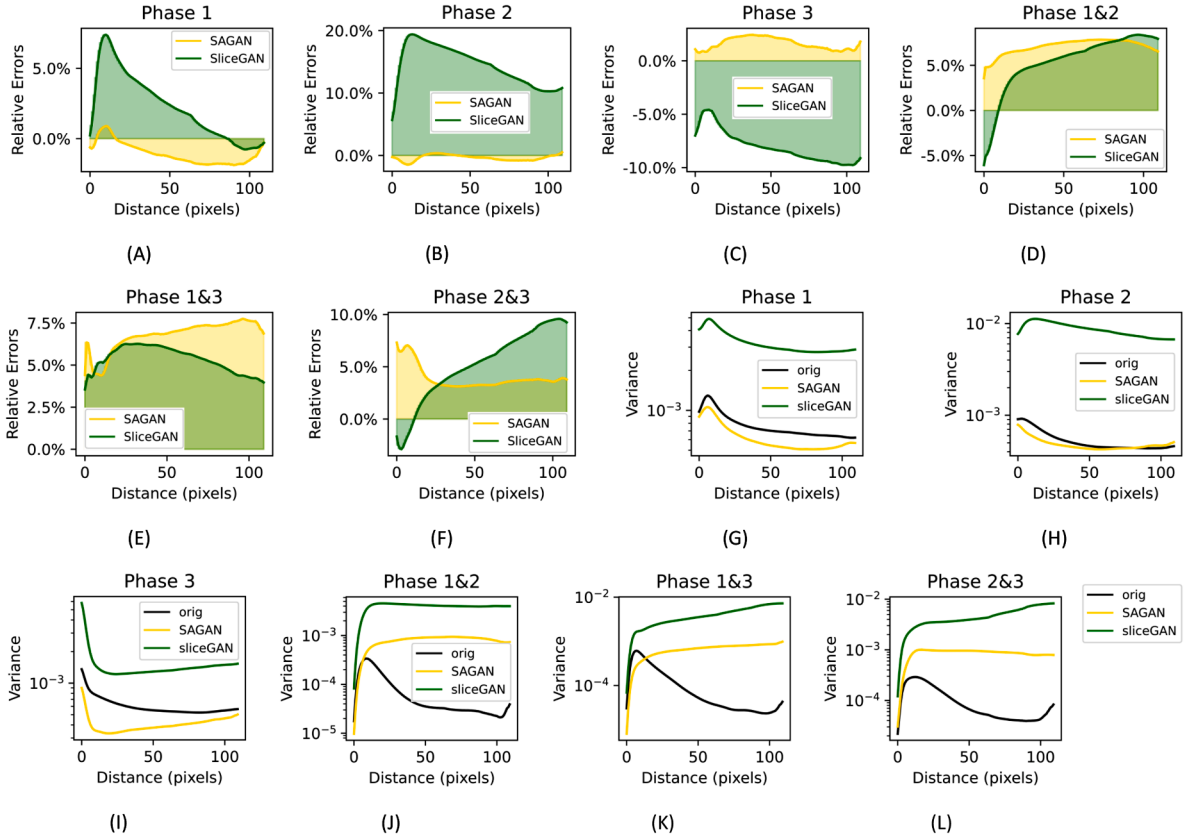

Fig. 30. Relative error and variance of the voxel Two-Point Correlation Function (TPCF) of Cathode materials for SOFC. (A) Relative error of TPCF in phase 1. (B) Relative error of TPCF in phase 2. (C) Relative error of TPCF in phase 3. (D) Relative error of TPCF in phase 1 and phase 2. (E) Relative error of TPCF in phase 1 and phase 3. (F) Relative error of TPCF in phase 2 and phase 3. (G) Variance of TPCF in phase 1. (H) Variance of TPCF in phase 2. (I) Variance of TPCF in phase 3. (J) Variance of TPCF in phase 1 and phase 2. (K) Variance of TPCF in phase 1 and phase 3. (L) Variance of TPCF in phase 2 and phase 3.

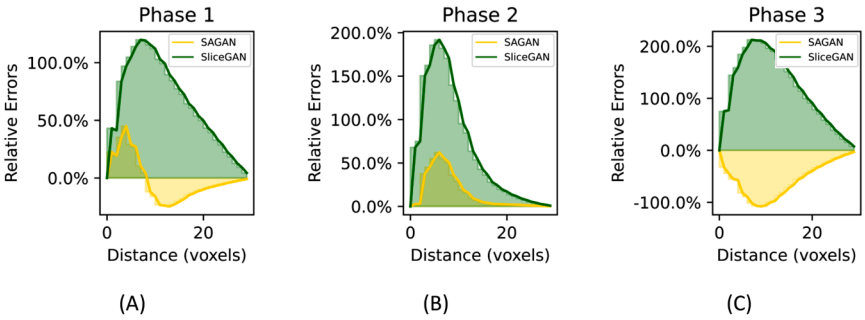

Fig. 31. The relative errors and variances on voxel Pore Size Distribution (PSD) on cathode material for SOFC. SAGAN's accuracy is superior to SliceGAN's. (A) Phase 1. (B) Phase 2. (C) Phase 3.

overemphasized the connectivity in the x and y directions in the central part, which is a typical mode collapse problem from cross-section slices in GANs and can be seen in Fig. 26. However, most of the synthetic results match the reference, as shown in the quality verification.

Since the material analyzed in this case is three-phase, the slice TPCF will be divided into results for each single phases and results between two different phases. TPCF between two phases in a multiphase material measures the probability of finding a point in one phase at a certain distance from a point in the other phase. It provides insights into the spatial distribution and interaction between the two phases. The distribution analysis of relative error and variance is shown in Fig. 27. In the single-phase results, SAGAN's results have significantly lower errors in terms of relative error than SliceGAN and are closer to the reference in variance. For the TPCF between phases, SAGAN is more accurate for the inter-phase, both for relative error and variance. And the slice TPCF results of each method are shown in Fig. A.6 in the Appendix.

For the PSD and LPF results of the multiphase material, the distribution analysis of relative error and variance is shown in Fig. 28. In terms of relative error and variance on PSD, SAGAN is significantly more accurate than SliceGAN. As for the relative error of LPF, SliceGAN has a smaller error for the second phase. But in terms of the connectivity of the other two phases, SAGAN is more accurate. The slice PSD and slice LPF results are shown in Fig. A.7 in the Appendix.

The CDP results and the relative errors of TPCCF are shown in Fig. 29. Similar to the TPCF results, SAGAN's TPCCF results have smaller relative errors compared with SliceGAN. In the CDP results, SAGAN's outcomes are more accurate, while SliceGAN exhibits greater variability. Particularly in the third phase, SliceGAN shows a lower distribution and weaker connectivity.

For the slice descriptor section, the purely CV-based SliceGAN shows severe overgeneration issues in the first and second phases of the cross-section slices, resulting in large areas of highly connected regions. However, the generation results for the third phase are relatively accurate. In contrast, the SAGAN method, which incorporates slice descriptors as constraints, produces results that more closely match the reference due to the explicit constraints of the five descriptors. Although some descriptor distributions have smaller variances compared to the reference, the accuracy is already very good. Additionally, SAGAN's training time is much shorter than SliceGAN's, making it highly efficient. (Fig. 30)

3.3.2. Voxel descriptor verify

In the voxel descriptor section, we compared the TPCF relative errors and variance results. In terms of relative error and variance in single phase, SAGAN's accuracy is significantly better than SliceGAN's. However, in the inter-phase results, the advantage of SAGAN becomes lost, especially when the distance is longer for the relative error, but the variance is still smaller than SliceGAN.

For the voxel PSD relative error in Fig. 31, although the results are not ideal, SAGAN's accuracy is significantly better than SliceGAN's.

For the voxel descriptor section, the purely CV-based SliceGAN exhibits both overgeneration and undergeneration of extreme samples in the first phase, overgeneration in the second phase, and undergeneration in the third phase. It tends to generate large contiguous areas in terms of phase generation size. Compared to the mode collapse observed in slice descriptor verification, the issue is not as severe, but localized over-connectivity still exists. In contrast, SAGAN, which incorporates descriptor control, can manage the geometric distribution on the plane through descriptors on the cross-section slices, thereby influencing the 3D microstructure's geometric structure. This explicit control intuitively and significantly improves reconstruction accuracy. Additionally, considering the training time, SAGAN's training-time effectiveness is high. To summarize, the purely CV-based method tends to focus excessively on the size and local connectivity of certain phases when generating such complex geometric distributions in multiphase materials. In contrast, the introduction of multiple independent descriptors can control planar generation and result in more accurate spatial outcomes.

4. Conclusion

Unlike traditional algorithms, CV-based reconstruction algorithms have struggled to achieve explicit and controllable reconstructions informed by morphological descriptors. In this study, we introduce SAGAN that explicitly incorporate hybrid descriptor loss

criteria during training to address complex geometric pattern reconstruction problems. Reconstruction experiments on three heterogeneous materials have demonstrated SAGAN's generality, high-fidelity, and time-effective training performance. SAGAN represents a robust and descriptor-driven framework for efficient and accurate microstructure reconstruction, offering clear advantages in both reconstruction quality and training-time performance for heterogeneous material systems.

First, SAGAN has a smart descriptor loss mechanism that includes both individual and distribution views. The individual descriptor loss function ensures consistency between each generated sample and the input descriptor, while the distribution descriptor loss function ensures high accuracy in the overall distribution of the slice descriptors. Secondly, a ResNet-like microstructure descriptor regression neural network is used to replace the traditional label judgment used in CGAN methods. This network facilitates explicit optimization of the generated results during training by quickly generating slice descriptors. Finally, a new descriptor injection mechanism is designed for the generator that allows the injection of reparameterized descriptors.

These innovations lead to substantial improvements in both reconstruction quality and training-time efficiency. Specifically, although SAGAN incurs higher computational cost per epoch due to the use of deep descriptor networks and multiple loss components, it achieves high-quality reconstructions in significantly less wall-clock training time compared to existing SOTA methods such as SliceGAN. The explicit descriptor loss also ensures more stable convergence and avoids the descriptor accuracy fluctuations observed in traditional CV-based approaches. Moreover, SAGAN yields superior accuracy. The explicit slice descriptor constraint leads to higher fidelity in 2D cross-sectional descriptor validation and, by extension, improved accuracy in reconstructed 3D voxel descriptors. Since the geometric properties of 3D microstructures can be characterized through multiple directional 2D slices, this approach enhances the overall 3D structure fidelity. Finally, the use of multiple descriptor types enables SAGAN to better capture complex geometric features. In contrast to conventional CV-based methods that may overemphasize certain patterns and suffer from overgeneration, our method maintains a more balanced and interpretable learning process.

However, our research is not without limitations. Due to the restriction that the training data's microstructure must be two-dimensional images, the descriptor loss can only consider slice descriptors. In three-dimensional reconstruction, voxel descriptors are naturally more important. Although multi-directional consecutive two-dimensional geometric information can reflect three-dimensional geometric information, we currently cannot achieve the mapping from slice descriptors to voxel descriptors, as it is not possible to simultaneously impose explicit constraints on voxel descriptors. Additionally, the descriptor ResNet used by SAGAN is pre-trained for each specific material and cannot be trained simultaneously with the discriminator like in traditional CGANs. In future work, we plan to investigate an descriptor ResNet that maps slice descriptors to voxel descriptors and reduces the pre-training component. This will improve SAGAN by enabling reconstruction methods that simultaneously consider slice descriptors and voxel descriptors for dimension enhancement.

CRediT authorship contribution statement

Xiangyun Ge: Writing – original draft, Methodology, Investigation, Formal analysis, Conceptualization; **Liyuan Wang:** Validation, Methodology, Investigation, Conceptualization; **Liam J. Garcia:** Investigation, Formal analysis; **Shan Zhong:** Validation, Investigation; **Bingbing Chen:** Validation, Investigation; **Chenfeng Li:** Writing – review & editing, Supervision, Resources, Project administration, Funding acquisition, Conceptualization.

Data availability

No data was used for the research described in the article.

Declaration of interests

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Chenfeng Li reports financial support was provided by The Royal Society. Chenfeng Li reports financial support was provided by Engineering and Physical Sciences Research Council. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

The authors would like to thank the supports from Chinese Scholarship Council, Swansea University, EPSRC (EP\X035026 \1, EP\W524694 \1) and the Royal Society (IF\R2\2320099).

Appendix A.

A.1. The descriptor results of case 1 - Fontainebleau sandstone

Fig. A.1 shows the individual PSD distribution for each method, and Fig. A.2 shows the individual LPF distribution. Both SAGAN and SliceGAN achieved good results within similar training time, among which the SAGAN model is closer to the reference value, and the divergence of the SINN model is the largest, which is caused by the uncertainty of its generative model.

A.2. The descriptor results of case 2 - Polytetrafluoroethylene

For the results of PSD and LPF in differernt methods, SliceGAN exhibited greater variability in Fig. A.4(A) and Fig. A.5, with anomalous generated results showing very large pore sizes, as reflected in the maximum limit. Compared to SAGAN's and SAGAN-short's generated results, SliceGAN tends to produce smaller pores overall but occasionally generates anomalously large pores.

A.3. The descriptor results of case 3 - Cathode materials for solid oxide fuel cell (SOFC)

The TPCF results are shown in Fig. A.6. First, both methods capture the correct geometry pattern on the mean value curve, with SAGAN demonstrating unparalleled accuracy in distribution. In the single-phase results, SliceGAN exhibits a significant overgeneration problem in the first and second phases, with its maximum limit curve far exceeding the results generated by SAGAN and the reference. Additionally, its minimum limit curve shows lower bounds at short distances compared to the reference. Without the constraint of descriptor loss, SliceGAN, relying solely on CV loss, severely overgenerates the first and second phases. However, in the third phase, neither method can express the individual's extremely high proportion of samples in the reference. Both SAGAN and SliceGAN have maximum limit curves lower than the reference's maximum limit curve. Although such high TPCF samples are rare in the reference, both methods show weaker generation capabilities for these less common sample types. For the TPCF between phases, due to the overgeneration of the first and second phases by SliceGAN, there are instances where the minimum limit is zero across all inter-phase results. Overgeneration of one phase reduces the contact surface in the generated results, leading to a lack of interconnectivity in the planar space. While SAGAN's maximum limit is close to the reference, its minimum limit is higher, overemphasizing the connectivity between phases.

For the PSD and LPF results of the multiphase material, shown in Fig. A.7, we only calculate the considered phase rather than the void phase as in two-phase materials. The PSD section of reference reveals that all three phases have distributions of both large and small sizes. SliceGAN overgenerates both large and small size distributions for the first and second phases, reflected in the excessively high maximum limit and low minimum limits. Both methods exhibit some overgeneration of large-sized phases for the third phase, but this is not significantly higher than the standard. In the LPF section, SliceGAN's generated distribution is more dispersed for the first and second phases, showing greater differences in phase connectivity compared to SAGAN. However, SliceGAN shows slightly better connectivity for the third phase than SAGAN.

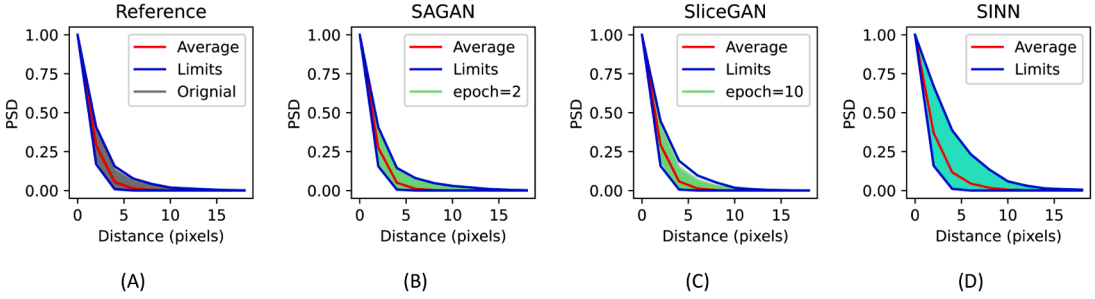

Fig. A.1. Fontainebleau sandstone slice Pore Size Distribution (PSD) results. (A) Reference. (B) SAGAN, epoch 2. (C) SliceGAN, epoch 10. (D) SINN.

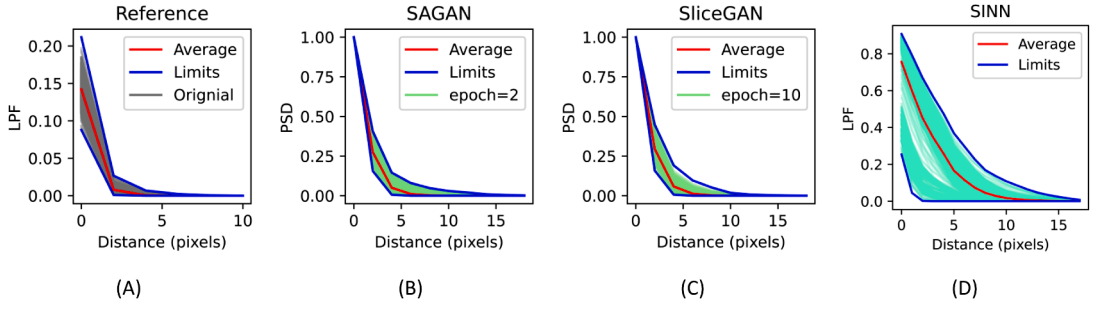

Fig. A.2. Fontainebleau sandstone slice Lineal Path Function (LPF) results. (A) Reference. (B) SAGAN, epoch 2. (C) SliceGAN, epoch 10. (D) SINN.

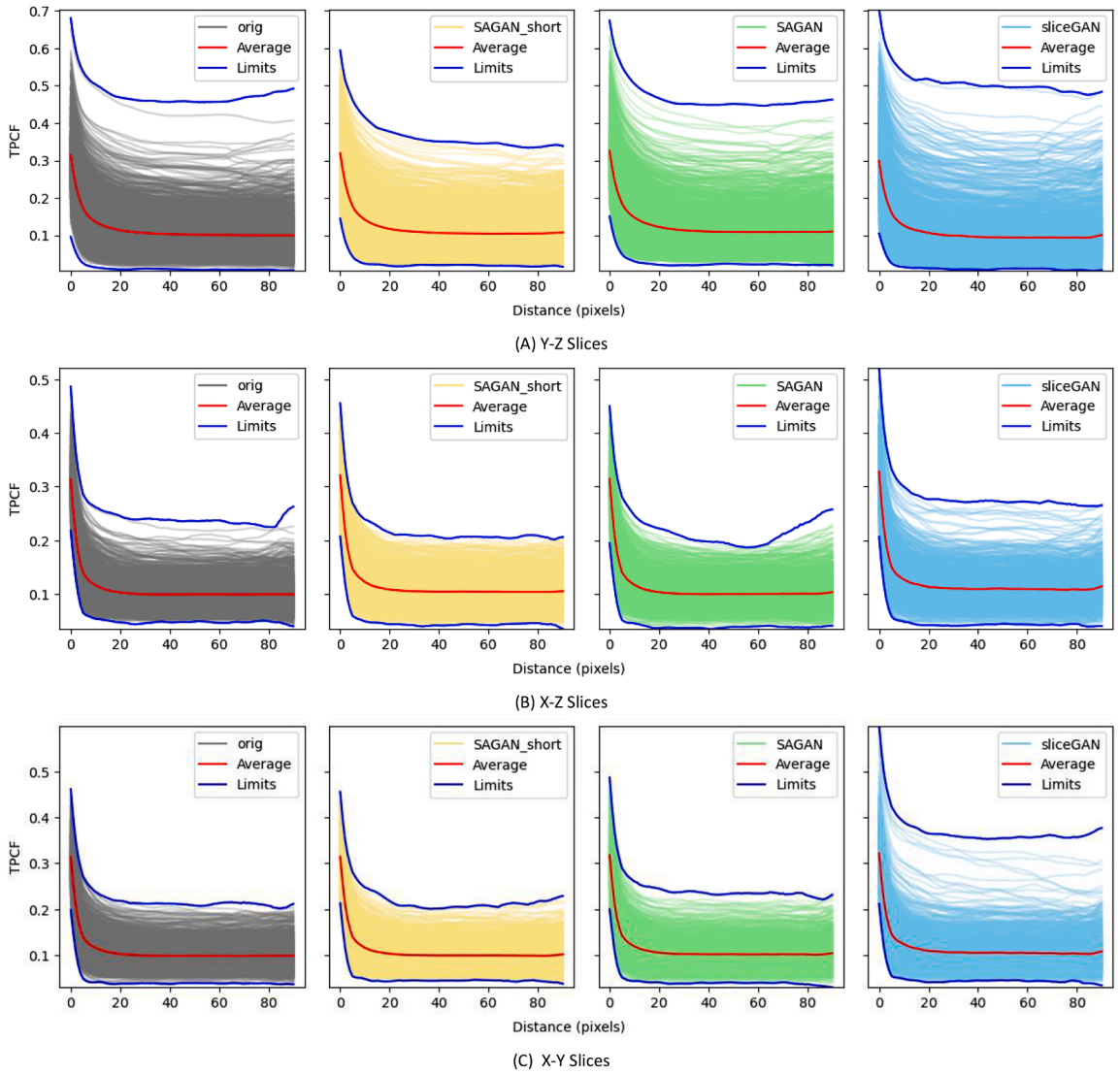

Fig. A.3. Polytetrafluoroethylene slice Two-Point Correlation Function (TPCF) results for different directions. (A) The TPCF results comparisons on Y-Z slices. (B)The TPCF results comparisons on X-Z slices. (C) The TPCF results comparisons on X-Y slices.

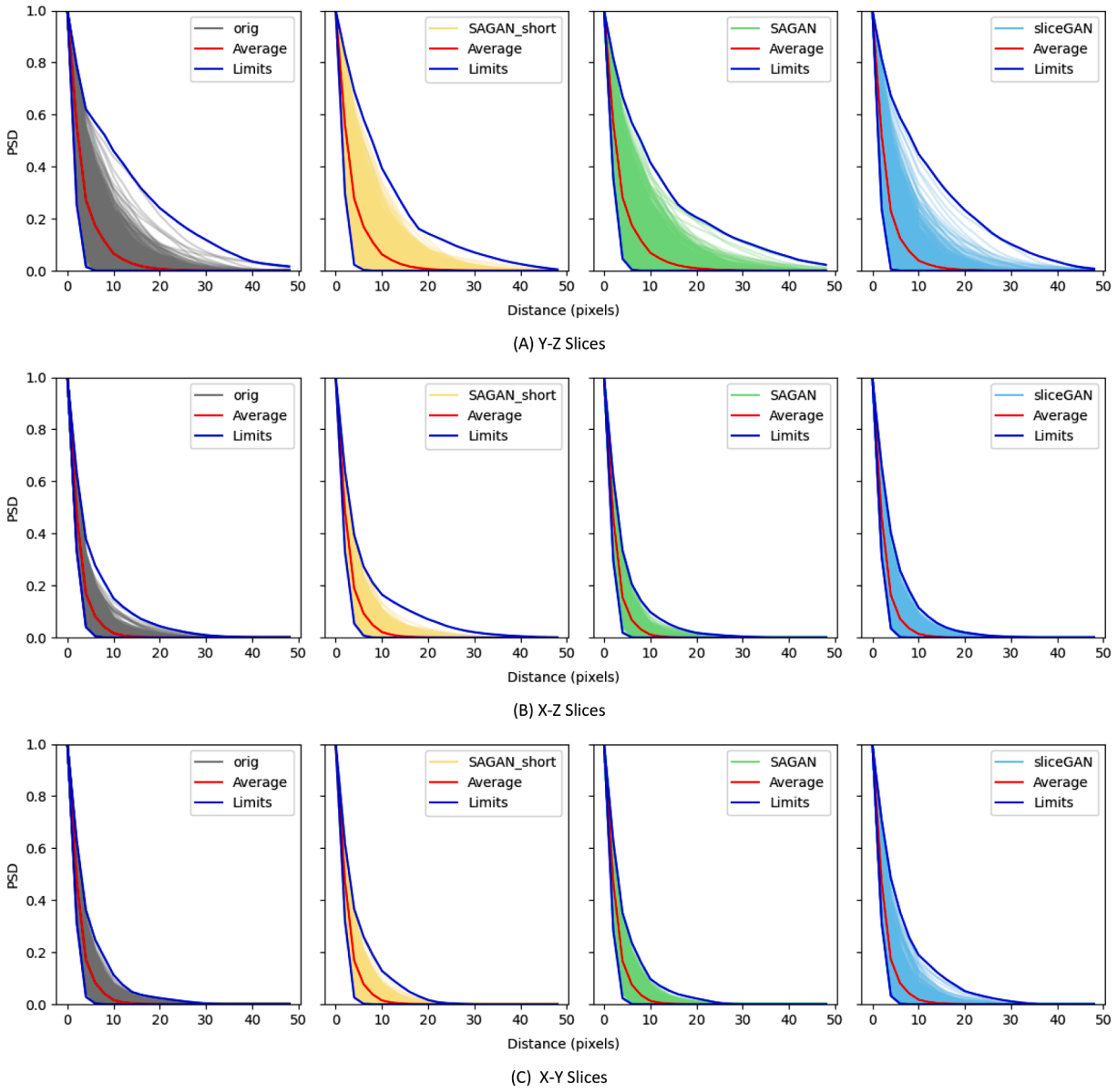

Fig. A.4. Polytetrafluoroethylene slice Pore Size Distribution (PSD) results for different directions. (A) The PSD results comparisons on Y-Z slices. (B) The PSD results comparisons on X-Z slices. (C) The PSD results comparisons on X-Y slices.

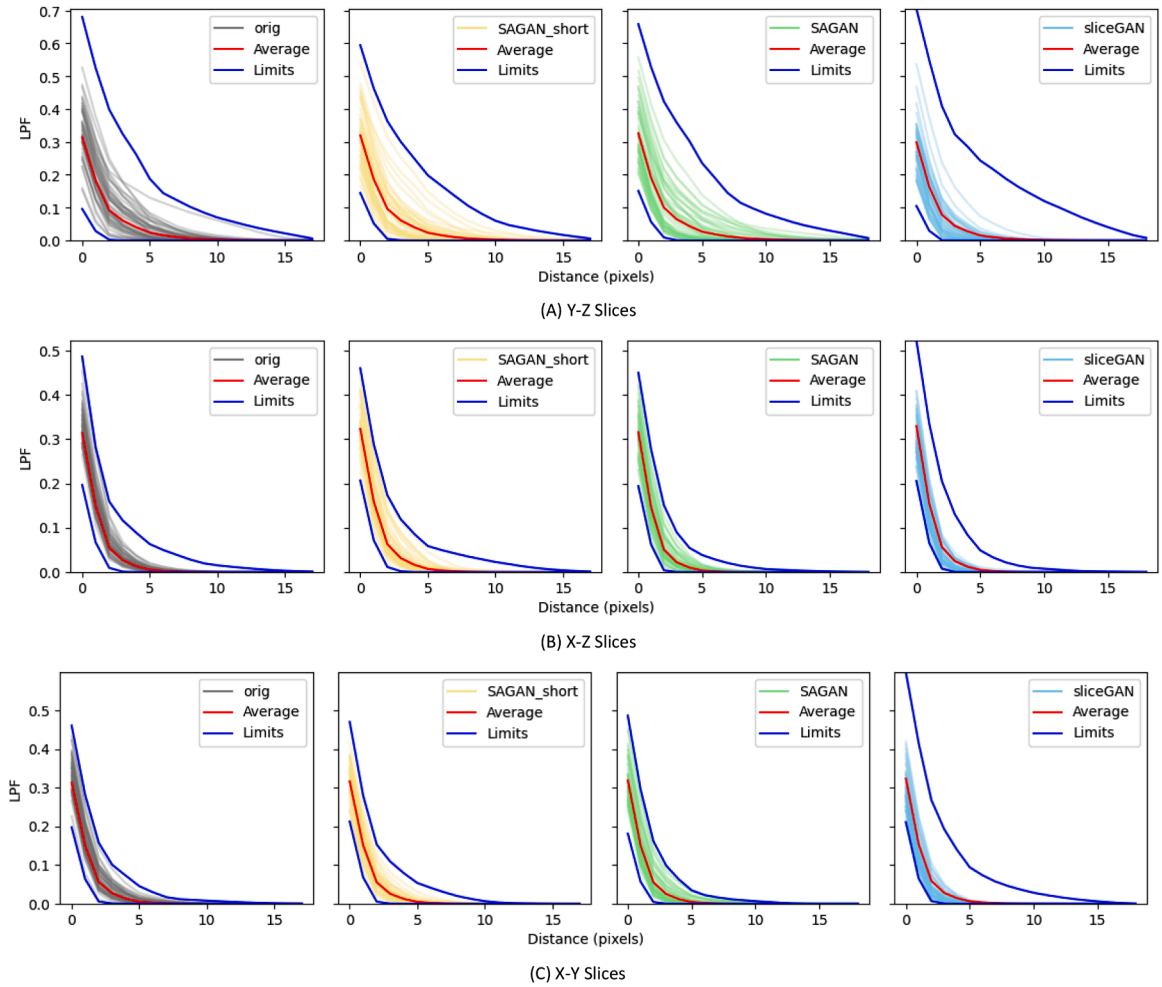

Fig. A.5. Polytetrafluoroethylene slice Lineal Path Function (LPF) results for different directions. (A) The LPF results comparisons on Y-Z slices. (B) The LPF results comparisons on X-Z slices. (C) The LPF results comparisons on X-Y slices.

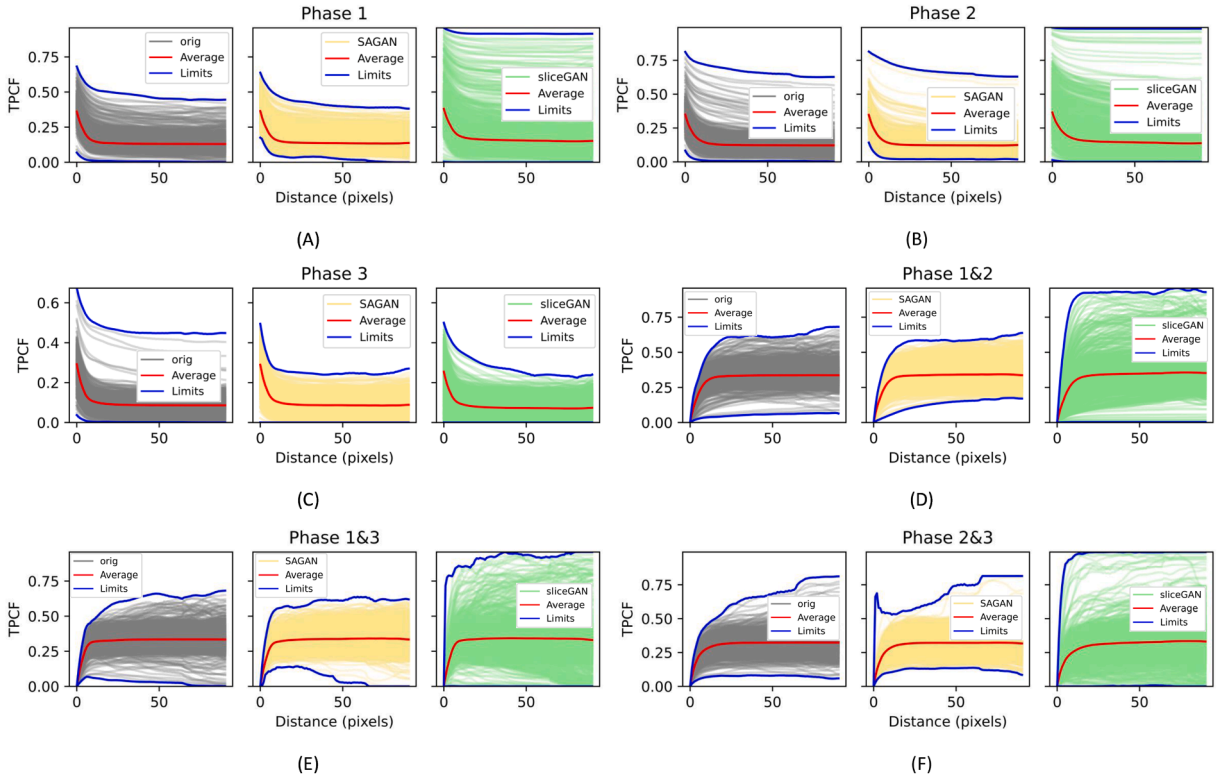

Fig. A.6. Slice the TPCF of Cathode materials for SOFC. SAGAN's accuracy is significantly better than SliceGAN's. (A)Phase 1. (B) Phase 2. (C) Phase 3. (D) Phase 1 and phase 2. (E) Phase 1 and phase 3. (F) Phase 2 and phase 3.

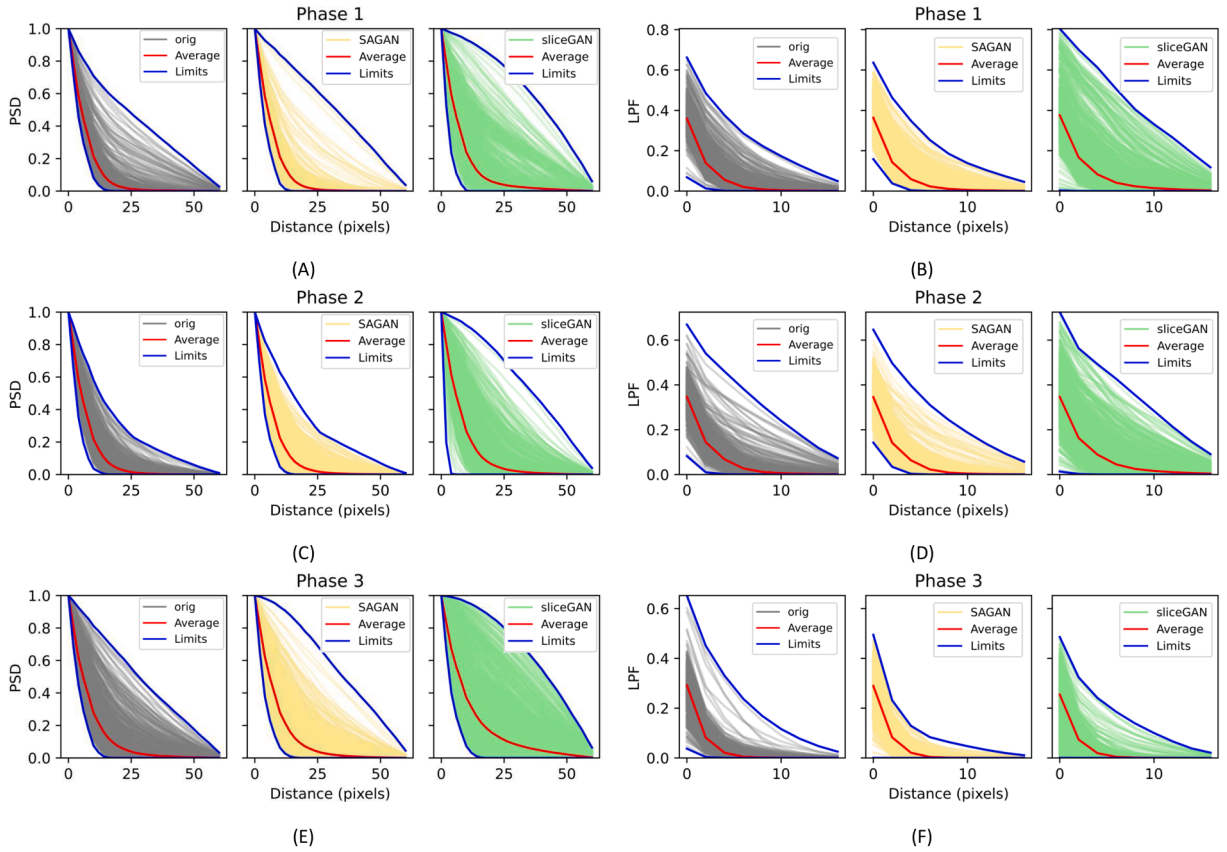

Fig. A.7. Slice Pore Size Distribution (PSD) and slice Lineal-Path Function (LPF) of Cathode materials for SOFC. (A) PSD of phase 1. (B) LPF of phase 1. (C) PSD of phase 2. (D) LPF of phase 2. (E) PSD of phase 3. (F) LPF of phase 3.

References

[1] A. Ahmed, Compressive strength and microstructure of soft clay soil stabilized with recycled bassanite, Appl. Clay Sci. 104 (2015) 27–35.
[2] D. Bouvard, J.-M. Chaix, R. Dendievel, A. Fazekas, J.M. Létang, G. Peix, D. Quenard, Characterization and simulation of microstructure and properties of EPS lightweight concrete, Cem. Concr. Res. 37 (12) (2007) 1666–1673.
[3] V. Kouznetsova, W. Brekelmans, F.F. Baaijens, An approach to micro-macro modeling of heterogeneous materials, Comput. Mech. 27 (2001) 37–48. https://doi.org/10.1007/S004660000212
[4] M. Wang, N. Pan, Predictions of effective physical properties of complex multiphase materials, Mater. Sci. Eng. R Rep. 63 (1) (2008) 1–30.
[5] S. Cui, J. Fu, S. Cen, H.R. Thomas, C. Li, The correlation between statistical descriptors of heterogeneous materials, Comput. Methods Appl. Mech. Eng. 384 (2021) 113948.
[6] M.V. Karsanina, K.M. Gerke, E.B. Skvortsova, D. Mallants, Universal spatial correlation functions for describing and reconstructing soil microstructure, PLoS ONE 10 (5) (2015) e0126515.
[7] H.J. Lim, H. Choi, G.J. Yun, Multiscale failure and damage analysis of sheet molding compound (SMC) composites using micro-CT image-based reconstruction model, Compos. Part B: Eng. 231 (2022) 109593.
[8] H. Al-Marzouqi, Digital rock physics: using CT scans to compute rock properties, IEEE Signal Process. Mag. 35 (2) (2018) 121–131.
[9] K.-H. Lee, H.W. Lee, G.J. Yun, A defect detection framework using three-dimensional convolutional neural network (3D-CNN) with in-situ monitoring data in laser powder bed fusion process, Optics Laser Technol. 165 (2023) 109571.
[10] Y. Wang, J.D. Miller, Current developments and applications of micro-CT for the 3D analysis of multiphase mineral systems in geometallurgy, Earth Sci. Rev. 211 (2020) 103406.
[11] X. Yao, L. Yu, Y. Ke, L. Jin, W. Wang, State-of-the-art research on loess microstructure based on X-ray computer tomography, Appl. Sci. 14 (15) (2024) 6402.
[12] P.I. Guntoro, Y. Ghorbani, P.-H. Koch, J. Rosenkranz, X-ray microcomputed tomography (μCT) for mineral characterization: a review of data analysis methods, Minerals 9 (3) (2019) 183.
[13] A. Mohammed, A. Abdullah, Scanning electron microscopy (SEM): a review, in: Proceedings of the 2018 International Conference on Hydraulics and Pneumatics-HERVEX, Băile Govora, Romania, 2018, 2018, pp. 7–9.
[14] M. Cantoni, L. Holzer, Advances in 3D focused ion beam tomography, MRS Bull. 39 (4) (2014) 354–360.
[15] L. Holzer, P. Marmet, M. Fingerle, A. Wiegmann, M. Neumann, V. Schmidt, Image based methodologies, workflows, and calculation approaches for tortuosity, in: Tortuosity and Microstructure Effects in Porous Media: Classical Theories, Empirical Data and Modern Methods, Springer, 2023, pp. 91–159.
[16] J.A. Quiblier, A new three-dimensional modeling technique for studying porous media, J. Colloid Interface Sci. 98 (1) (1984) 84–102.
[17] J.W. Feng, C.F. Li, S. Cen, D. Owen, Statistical reconstruction of two-phase random media, Comput. Struct. 137 (2014) 78–92.
[18] J.W. Feng, S. Cen, C.F. Li, D. Owen, Statistical reconstruction and Karhunen–Loève expansion for multiphase random media, Int. J. Numer. Methods Eng. 105 (1) (2016) 3–32.
[19] P. Tahmasebi, M. Sahimi, Reconstruction of three-dimensional porous media using a single thin section, Phys. Rev. E 85 (6) (2012) 066709.
[20] Z.M. Wang, A. Kwan, H.C. Chan, Mesoscopic study of concrete I: generation of random aggregate structure and finite element mesh, Comput. Struct. 70 (5) (1999) 533–544.

[21] H. Xu, D.A. Dikin, C. Burkhart, W. Chen, Descriptor-based methodology for statistical characterization and 3D reconstruction of microstructural materials, Comput. Mater. Sci. 85 (2014) 206–216.
[22] C. Yeong, S. Torquato, Reconstructing random media, Phys. Rev. E 57 (1) (1998) 495.
[23] C. Manwart, S. Torquato, R. Hilfer, Stochastic reconstruction of sandstones, Phys. Rev. E 62 (1) (2000) 893.
[24] Y. Wang, J. Lv, L. Zhu, Y. Ma, CALYPSO: a method for crystal structure prediction, Comput. Phys. Commun. 183 (10) (2012) 2063–2070.
[25] D. Beasley, D.R. Bull, R.R. Martin, An overview of genetic algorithms: Part 1, fundamentals, Univ. Comput. 15 (2) (1993) 56–69.
[26] E. Laloy, R. Hérault, J. Lee, D. Jacques, N. Linde, Inversion using a new low-dimensional representation of complex binary geological media based on a deep neural network, Adv. Water Resour. 110 (2017) 387–405.
[27] Y. Du, H. Tu, T. Zhang, Pore space reconstruction of shale using improved variational autoencoders, Geofluids 2021 (2021) 1–11.
[28] R. Bostanabad, Reconstruction of 3D microstructures from 2D images via transfer learning, Comput.-Aided Des. 128 (2020) 102906.
[29] X. Li, Y. Zhang, H. Zhao, C. Burkhart, L.C. Brinson, W. Chen, A transfer learning approach for microstructure reconstruction and structure-property predictions, Sci. Rep. 8 (1) (2018) 13461.
[30] F. Zhang, X. He, Q. Teng, X. Wu, X. Dong, 3D-PMRNN: Reconstructing three-dimensional porous media from the two-dimensional image with recurrent neural network, J. Petrol. Sci. Eng. 208 (2022) 109652.
[31] F. Zhang, Q. Teng, X. He, X. Wu, X. Dong, Improved recurrent generative model for reconstructing large-size porous media from two-dimensional images, Phys. Rev. E 106 (2) (2022) 025310.
[32] J. Fu, D. Xiao, D. Li, H.R. Thomas, C. Li, Stochastic reconstruction of 3D microstructures from 2D cross-sectional images using machine learning-based characterization, Comput. Methods Appl. Mech. Eng. 390 (2022) 114532.
[33] J. Fu, M. Wang, D. Xiao, S. Zhong, X. Ge, M. Wu, B. Evans, Hierarchical reconstruction of 3D well-connected porous media from 2D exemplars using statistics-informed neural network, Comput. Methods Appl. Mech. Eng. 410 (2023) 116049.
[34] J. Fu, W. Tan, Stochastic reconstruction of multiphase composite microstructures using statistics-encoded neural network for poro/micro-mechanical modelling, Comput. Methods Appl. Mech. Eng. 441 (2025) 117986.
[35] J. Phan, M. Sarmad, L. Ruspini, G. Kiss, F. Lindseth, Generating 3D images of material microstructures from a single 2D image: a denoising diffusion approach, Sci. Rep. 14 (1) (2024) 6498.
[36] N.N. Vlassis, W. Sun, K.A. Alshibli, R.A. Regueiro, Synthesizing realistic sand assemblies with denoising diffusion in latent space, Int. J. Numer. Anal. Methods Geomech. 48 (16) (2024) 3933–3956.
[37] P. Fernandez-Zelaia, J. Cheng, J. Mayeur, A.K. Ziabari, M.M. Kirka, Digital polycrystalline microstructure generation using diffusion probabilistic models, Materialia 33 (2024) 101976.
[38] M.O. Buzzy, A.E. Robertson, S.R. Kalidindi, Statistically conditioned polycrystal generation using denoising diffusion models, Acta Mater. 267 (2024) 119746.
[39] C. Düreth, P. Seibert, D. Rücker, S. Handford, M. Kästner, M. Gude, Conditional diffusion-based microstructure reconstruction, Mater. Today Commun. 35 (2023) 105608.
[40] K.-H. Lee, H.J. Lim, G.J. Yun, A data-driven framework for designing microstructure of multifunctional composites with deep-learned diffusion-based generative models, Eng. Appl. Artif. Intell. 129 (2024) 107590.
[41] X. Lyu, X. Ren, Microstructure reconstruction of 2D/3D random materials via diffusion-based deep generative models, Sci. Rep. 14 (1) (2024) 5041.
[42] K.M. Guan, T.I. Anderson, P. Creux, A.R. Kovscek, Reconstructing porous media using generative flow networks, Comput. Geosci. 156 (2021) 104905.
[43] X. Luo, Z. Wang, Q. Wang, J. Lv, L. Wang, Y. Wang, Y. Ma, Crystalflow: a flow-based generative model for crystalline materials, arXiv preprint arXiv:2412.11693 (2024).
[44] H. Mirzaee, S. Kamrava, Inverse design of microstructures using conditional continuous normalizing flows, Acta Mater. 285 (2025) 120704.
[45] L. Mosser, O. Dubrule, M.J. Blunt, Reconstruction of three-dimensional porous media using generative adversarial neural networks, Phys. Rev. E 96 (4) (2017) 043309.
[46] E. Laloy, R. Hérault, D. Jacques, N. Linde, Training-image based geostatistical inversion using a spatial generative adversarial neural network, Water Resour. Res. 54 (1) (2018) 381–406.
[47] A. Valsecchi, S. Damas, C. Tubilleja, J. Arechalde, Stochastic reconstruction of 3D porous media from 2D images using generative adversarial networks, Neurocomputing 399 (2020) 227–236.
[48] E. Kononov, M. Tashkinov, V.V. Silberschmidt, Reconstruction of 3d random media from 2d images: generative adversarial learning approach, Comput.-Aided Des. 158 (2023) 103498.
[49] J. Feng, Q. Teng, B. Li, X. He, H. Chen, Y. Li, An end-to-end three-dimensional reconstruction framework of porous media from a single two-dimensional image based on deep learning, Comput. Methods Appl. Mech. Eng. 368 (2020) 113043.
[50] D. Fokina, E. Muravleva, G. Ovchinnikov, I. Oseledets, Microstructure synthesis using style-based generative adversarial networks, Phys. Rev. E 101 (4) (2020) 043308.
[51] W. Zha, X. Li, Y. Xing, L. He, D. Li, Reconstruction of shale image based on Wasserstein generative adversarial networks with gradient penalty, Adv. Geo-Energy Res. 4 (1) (2020).
[52] S. Kench, S.J. Cooper, Generating three-dimensional structures from a two-dimensional slice with generative adversarial network-based dimensionality expansion, Nat. Mach. Intell. 3 (4) (2021) 299–305.
[53] M. Mirza, S. Osindero, Conditional generative adversarial nets, (2014) arXiv preprint arXiv:1411.1784.
[54] X. Li, Z. Yang, L.C. Brinson, A. Choudhary, A. Agrawal, W. Chen, A deep adversarial learning methodology for designing microstructural material systems, in: International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, 51760, American Society of Mechanical Engineers, 2018, p. V02BT03A008.
[55] R. Yin, Q. Teng, X. Wu, F. Zhang, S. Xiong, Three-dimensional reconstruction of granular porous media based on deep generative models, Phys. Rev. E 108 (5) (2023) 055303.
[56] C. Du, G. Zou, J. Huo, B. Feng, L. Liu, Generative AI-enabled microstructure design of porous thermal interface materials with desired effective thermal conductivity, J. Mater. Sci. 58 (41) (2023) 16160–16171.
[57] R. Shams, M. Masihi, R.B. Boozarjomehry, M.J. Blunt, A hybrid of statistical and conditional generative adversarial neural network approaches for reconstruction of 3D porous media (ST-CGAN), Adv. Water Resour. 158 (2021) 104064.
[58] J. Feng, X. He, Q. Teng, C. Ren, H. Chen, Y. Li, Reconstruction of porous media from extremely limited information using conditional generative adversarial networks, Phys. Rev. E 100 (3) (2019) 033308.
[59] D. Cao, Z. Hou, Q. Liu, F. Fu, Reconstruction of three-dimension digital rock guided by prior information with a combination of InfoGAN and style-based GAN, J. Petrol. Sci. Eng. 208 (2022) 109590.
[60] P. Chi, J. Sun, X. Luo, R. Cui, H. Dong, Reconstruction of 3D digital rocks with controllable porosity using CVAE-GAN, Geoenergy Sci. Eng. 230 (2023) 212264.
[61] R. Zhou, C. Wu, 3D reconstruction of digital rock guided by petrophysical parameters with deep learning, Geoenergy Sci. Eng. 231 (2023) 212320.
[62] K.-H. Lee, G.J. Yun, Denoising diffusion-based synthetic generation of three-dimensional (3D) anisotropic microstructures from two-dimensional (2D) micrographs, Comput. Methods Appl. Mech. Eng. 423 (2024) 116876.
[63] J. Rubinstein, S. Torquato, Flow in random porous media: mathematical formulation, variational principles, and rigorous bounds, J. Fluid Mech. 206 (1989) 25–46.
[64] I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin, A.C. Courville, Improved training of wasserstein gans, Adv. Neural Inf. Process. Syst. 30 (2017) 5769–5779.
[65] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 770–778.
[66] J. Hu, L. Shen, G. Sun, Squeeze-and-excitation networks, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, pp. 7132–7141.