Computer-Aided Design 138 (2021) 103041

Contents lists available at ScienceDirect

Computer-Aided Design

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

ScaffoldGAN: Synthesis of Scaffold Materials based on Generative Adversarial Networks☆

Hui Zhang [a,*], Lei Yang [a], Changjian Li [b], Bojian Wu [c], Wenping Wang [a,1]

[a] Department of Computer Science, The University of Hong Kong, Hong Kong, China
[b] Department of Computer Science, University College London, London, UK
[c] Chinese Academy of Sciences, Shenzhen, China

ARTICLE INFO

Article history:
Received 3 April 2020
Received in revised form 23 February 2021
Accepted 7 April 2021

Keywords:
3D shape synthesis
Generative adversarial networks
Deep learning
Scaffold material
Complex structure

ABSTRACT

Digitally synthesizing scaffold-like materials with complex structures, e.g., bones or metal foam, is a fundamental yet challenging task in tissue engineering and other biomedical applications, because it is difficult to generate synthesized results with equal visual complexity, strong spatial coherence, and similar statistical metrics. To handle these challenges, we present ScaffoldGAN, an efficient end-to-end framework based on generative adversarial networks (GANs) for synthesizing three-dimensional (3D) materials with complex internal structures resembling the given exemplar. Specifically, we propose a novel structural loss to enforce strong spatial coherence in the synthesized results by leveraging the deep features learned by our networks. To demonstrate the effectiveness of our model and the proposed structural loss term, we collected example data containing various structural complexities, covering two categories of materials, i.e., bones and metal foams. Extensive comparative experiments on these collected data showed that our method outperforms state-of-the-art methods, producing synthesized results with better visual quality and desirable statistical metrics. The ablation study proves the structural loss is the main contributor to the performance gain, validating our design choice.

1. Introduction

Digital synthesis of scaffold-like materials with complex structures is a fundamental yet challenging task and is significant for the design and exploration of the cellular materials for biomedical applications such as tissue engineering. Taking the bone scaffold synthesis as an example, a fundamental requirement is that the synthesized scaffold should resemble the original bone as much as possible in order to provide a similar biological environment and mechanical support for bone regeneration [1]. Although recent computational tools [2–4] have facilitated this task, digital synthesis of desired scaffolds is still unsolved and faces two main challenges: enforcing spatial coherence observed in the material exemplar and maintaining statistical metrics to be similar to those of the exemplar.

Some attempts have been made to handle the problem of bone scaffold synthesis. Conventional approaches [1,3,5] for designing bone scaffolds mimic macro-scale properties of materials with interconnected and periodical pores uniformly distributed in a prescribed volume. Procedural methods [6,7] based on random noise sampling are developed to generate stochastic microstructures. Parametric texture methods [8] approach this problem by finding a global statistical descriptor that characterizes the given texture exemplars. However, all these previous approaches could not fully capture the intricate spatial coherence and fine-grained details of example materials to be replicated.

To address the first challenge, one may consider employing example-based texture synthesis approach [9–13] to generate large-scale, scaffold-like structures by taking a small piece of it as an exemplar. Zhang et al. [14] extended previous works to directly use 3D exemplars instead of multiple 2D slices, which achieved better synthesized results than its 2D counterpart. While the visual similarity between the reference and synthesized data may be acceptable, we observed in our experiments that these methods tend to generate results with obvious artifacts, such as repeated patterns and disconnected branches. This is due to the fact that these methods match a *local* patch of the synthesized result and that of the exemplar to enforce such visual similarity, losing a global control over the synthesis process. To overcome the locality nature of the patch-match idea one may enlarge the patch size for matching, but this cubic expansion could soon become computationally intractable.

While producing visually similar results is important, another critical challenge to be tackled is to ensure that the statistical metrics of the generated results are close to those of the exemplar. In the field of material sciences, widely-used statistical

☆ This paper has been recommended for acceptance by Hayong Shin.
* Corresponding author.
E-mail address: huizh@connect.hku.hk (H. Zhang).
[1] IEEE Fellow.

https://doi.org/10.1016/j.cad.2021.103041
0010-4485/

measurements are porosity [15] and two-point correlation function [16], each indicating the spatial distribution and coherence of materials, respectively. Many attempts have been made to extract parametric models (or textons) to encode visual-related statistics relying on engineered features for texture synthesis (see for example [9,11]). But features encoding morphological statistics are hard to be manually crafted. In line with this idea but using deep features, the framework of generative adversarial networks (GANs) is thus a reasonable baseline for our task, because it is designed to model the data distribution during training and output a synthesized result similar to the samples from the training dataset.

Inspired by the success of GANs and relevant deep learning techniques in image synthesis [17–19] and 3D modeling [20,21], we propose ScaffoldGAN, an end-to-end generative adversarial networks based solution, to address the above major challenges for synthesis of scaffold materials with complex structures. Our model is trained on volumetric samples extracted from a large material exemplar acquired via CT scanning, and finally outputs synthesized results with strong similarity to the exemplar in both visual and statistical senses (Fig. 1). Specifically, we introduce a novel structural loss term to mimics the patch matching scheme. However, comparing to the aforementioned texture synthesis methods that perform patch matching in the 2D/3D image domain, our design of the structural loss leverages the convolutional features during the patch matching stage in order to overcome the locality nature of previous methods. This is made possible because convolutional features at different convolution layers embed information of different scales as the receptive fields of the convolution enlarge when the layers get deeper. Therefore, enforcing the similarity between convolutional features allows reproducing strong *spatial coherence* at different scales observed in exemplars. The adversary between the generation and discrimination networks also enforces our synthesized results to closely resemble the samples from the training dataset, and thus implicitly ensures desirable *statistics* for our targeted applications.

We extensively evaluated our results through qualitative and quantitative analyses. A bone dataset and a metal foam dataset (morphologically similar to trabecular bones) acquired via micro-CT were used to demonstrate and validate our method in scaffold materials synthesis. Experiments on these datasets show that our method outperforms conventional example-based methods by far and obtains better statistical metrics comparing to our baseline model 3D-GAN [20]. We also tested our method on the ICL dataset [22] (with four stone exemplars), showing that our method also produces state-of-the-art results and indicating its applicability. The ablation study is presented at last to validate our design choices.

Our contributions are summarized as follows:

(1) ScaffoldGAN, a novel end-to-end generative adversarial network is proposed for the synthesis of 3D scaffold shape, which can efficiently produce high-quality synthesized results with intricate inner structures, while maintaining strong spatial coherence and similar statistical metrics to exemplars.

(2) A new structural loss term is designed to enforce strong spatial coherence of synthesized results by considering deep features similarity in the network.

(3) Training datasets (i.e., bones and metal foams) are collected to facilitate future studies in both computer vision and biomedical engineering.

2. Related work

2.1. Scaffolds structure design

Bone scaffold design is one of the primary application for porous material synthesis. Scaffolds should be made of complex internal structures to resemble the original bones to provide a similar biological environment and mechanical support for tissue repair [5]. Recent advances in digital modeling and tomographic reconstruction provide engineers a set of tools based on computer-aided design [2], image-based design [3] and implicit surfaces design [4] for scaffold design. Scaffold designs with randomly shaped pores [23] or different unit cells [24] in a periodic architecture are reported. However, previous approaches only allow to generate scaffolds with repeated structures under specific rules which are not enough for the requirements of biology similarity in tissue repairing.

2.2. Example-based texture synthesis

Texture synthesis methods [9–11,25,26] have been widely used in generating large images from small given examples. Given a 2D image as exemplar, example-based texture synthesis method can generate large-size images with similar patterns (e.g., [27] and [28]). Based on this example-based philosophy, studies (e.g.Kopf et al. [12], Chen et al. [13], Liu et al. [29]) employed several 2D orthogonal images to synthesize 3D solid structures, aiming to produce results that share visual appearance with the exemplars; however only limited information is contained in 2D slices, which is unable to produce convincing results. Instead, Zhang et al. [14] directly used 3D volumetric data from micro-CT scans to synthesize 3D porous materials, leading to satisfactory results. Yet, repetitive parts and a low degree of spatial coherence can be observed. This is attributed to the locality nature of example-based texture synthesis methods, which makes it difficult to learn global distribution of the exemplar through optimization of local neighborhood matching. While this limitation may be handled by using a relatively larger patch size in 2D, enlarging the size of a 3D cubical region would rapidly increase the computational cost and memory. Thus, traditional example-based methods can be prohibitive in producing convincing 3D results.

Our introduced novel structural loss term is to enforce strong spatial coherence of synthesized results by considering deep features similarity in the network, and target to synthesize shapes in 3D space, which is different from traditional texture synthesis methods that enforce the neighborhood similarity in image space. Once our network is well trained, the generator can produce synthesized results without extra similarity comparison, which again differs from general texture synthesis methods.

2.3. Image synthesis based on deep learning

Recently, deep learning methods [17,18,30–35] have enabled general users to obtain visually compelling image synthesis results. In particular, perceptual loss was proposed in [30] to enhance the performance of image generation tasks by comparing the difference between feature representations. In order to capture the intrinsic structure and style of the exemplar, Gatys et al. [17] proposed to use the Gram matrix as a feature descriptor for texture synthesis, nicely producing visually appealing results. Li and Wand [36] proposed to compute patch-based texture synthesis with deep neural features, which could generate plausible style transfer results but not considering global control. To circumvent the limitations of high computational burden suffered by previous works, feed-forward networks were applied in [31,37]. Sendik and Cohen-Or [38] and Bergmann et al. [19] proposed advanced methods for handling image textures with strong periodicity. A GAN-based method [39] is proposed to perform translation between images, e.g. generating an image given a sketch or vice versa. In this work, a PatchGAN model is devised. It gives real-or-fake scores to all patches in the last hidden layer

and then averages the scores to obtain the final score of the synthesized result. Different from PatchGAN, we follow the conventional GANs framework and produce a real-or-fake score for each generated (3D) image. It does not measure the difference between feature spaces of two images, and thus is different to our method. We compare our results with theirs in the Result section.

Our proposed structural loss is most related to the perceptual loss that is widely used in many different image synthesis tasks. However, we compute our structural loss in different layers only within cubical regions, instead of comparing the loss between feature maps of reference exemplars and synthesized results, in order to measure only regional similarity between the synthesized results and the reference. We also employ cubical regions with different aspect ratios to capture anisotropic perceptual features.

2.4. 3D Content generation using deep learning

Generating and reconstructing 3D shapes and objects has also attracted increasing interest in recent years [20,40–42]. Wu et al. [20] proposed 3D Shapenets, a large scale database for 3D content modeling tasks. Generative adversarial networks [43,44] for 3D content generation and editing [40,45] have also received increasing attention. As GAN frameworks often produce synthetic outputs with limited sizes, Jetchev et al. [18] proposed the spatial GANs to improve the scalability, which is beneficial to 3D content modeling. Most aforementioned studies on 3D content generation aim to synthesize man-made objects, such as chairs and airplanes. However few approaches focus on generating 3D objects with intricate internal structures. A recent work [46] proposes to generate 3D porous materials via a direct usage of GAN, which is barely adequate (as shown in our experiment) to derive materials with desired properties. In summary, current studies have not carefully considered the task of generating 3D models with complex internal structures. Therefore, in this paper we attempt to propose a ScaffoldGAN for this task, and discuss the effectiveness of the novel terms added to the baseline GAN model.

3. Approach

Taking a voxel-based exemplar scanned from real-world scaffold material (e.g., bone) as input, Our approach could output a synthesized 3D scaffold structure that resembles the given exemplar in terms of visual appearance, spatial coherence, and statistical metrics, after training on the samples extracted from input exemplar.

3.1. Model

We adopt the generative adversarial networks as our baseline model, which consists of two networks, i.e., generator G and discriminator D that are trained alternately by solving a minimax problem defined as below:

$$\min_G \max_D \mathcal{L}_{adv}(D,G) = E_{\mathbf{x}\sim p_{data(\mathbf{x})}}[log(D(\mathbf{x}))] + E_{\mathbf{z}\sim p_{\mathbf{z}}(\mathbf{z})}[log(1-D(G(\mathbf{z})))] \quad (1)$$

where $\mathbf{x}$ denotes training samples and $\mathbf{z}$ denotes noise tensors randomly sampled from a given distribution $p_{\mathbf{z}}$. During the adversarial training, performance of discriminator D is maximized so that it can correctly distinguish real samples $\mathbf{x}$ from synthesized results $G(\mathbf{z})$; on the other hand, generator G is trained to minimize $log(1-D(G(\mathbf{z})))$ to produce indistinguishable synthesized results.

The adversarial training facilitates learning the data distribution $p_{data(\mathbf{x})}$ of the training samples, which are extracted from a given input exemplar of a larger size (details are shown in the data preparation section). However, the adversarial loss alone is unable to take into account regional similarity between the synthesized results and training samples in both appearance and spatial coherence.

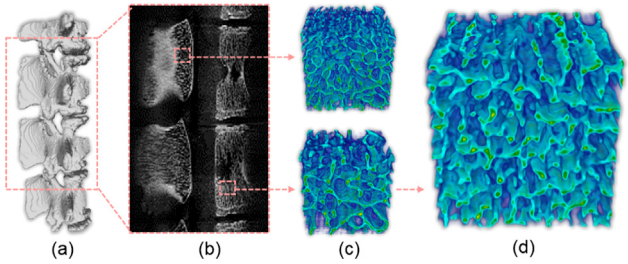

Fig. 1. Bone exemplars (c) extracted from a CT scan (b) of vertebral bone (a). Different morphology can be observed in the samples from different regions. The synthesized result (d) of the ScaffoldGAN strongly resembles the exemplar.

To this end, we incorporate additional structural loss ($\mathcal{L}_s$) and gram loss ($\mathcal{L}_g$) into adversarial training to formulate a novel loss function as follows:

$$\mathcal{L}_{total} = \mathcal{L}_{adv} + \lambda_g \mathcal{L}_g + \lambda_s \mathcal{L}_s, \quad (2)$$

where λ_s and λ_g are balancing weights. This formulation explicitly enforces regional similarity between the synthesized results $G(\mathbf{z})$ and the training samples $\mathbf{x}$, thereby maintaining the spatial coherence as well as visual appearance observed in the exemplar.

3.2. Structural loss for spatial coherence

The structural loss is defined as the cumulative summation of the matching errors between sampled cubical regions from the convolutional feature maps of the synthesized results and their optimal matches from the exemplar. In this way, the spatial coherence similarity at different scales is explicitly measured by the structural loss, and thus minimizing it shall lead to better synthesized results satisfying the aforementioned requirement.

We denote a cubical region centered at point p as $\mathcal{C}_p \in \mathbb{R}^{w\times w\times w}$. Given two cubical regions from the same convolutional features at layer l, the error between them is calculated as following:

$$E(p,q) = \|F^l(\mathcal{C}_p) - F^l(\mathcal{C}_q)\|^2 \quad (3)$$

where $F^l(\mathcal{C})$ returns the corresponding features defined on the cubical region $\mathcal{C}$.

As our goal is to ensure the spatial coherence similarity between the exemplar and the synthesized output, we search for a given cubical region $\mathcal{C}_q$ (from the synthesized result) its best match $\mathcal{C}_{\hat{p}}$ by minimizing the matching error (Eq. (3)) for all candidate cubical regions from the corresponding layer of the exemplar. This leads to the formulation of the structural loss:

$$L_s = \sum_l \sum_q \|F^l(\mathcal{C}_{\hat{p}}) - F^l(\mathcal{C}_q)\|^2 \quad (4)$$

where $\hat{p} = \text{argmin}_{p\in\mathcal{P}} E(p,q)$.

In order to efficiently search the optimal matching cubical region from the training samples for each individual cubical region in the synthesized result, we construct a look-up table $\{F^l(\mathcal{C}_p);\ p \in \mathcal{P}\}$ which contains a large number of features defined on the cubical regions $\mathcal{C}_p$ from the training samples. These cubical regions are obtained as follows. From the feature maps of each convolutional layer l we sample a set of anchor points $\mathcal{P} = \{p\}$ with a stride s along all axes and form the cubical regions $\mathcal{C}_p$

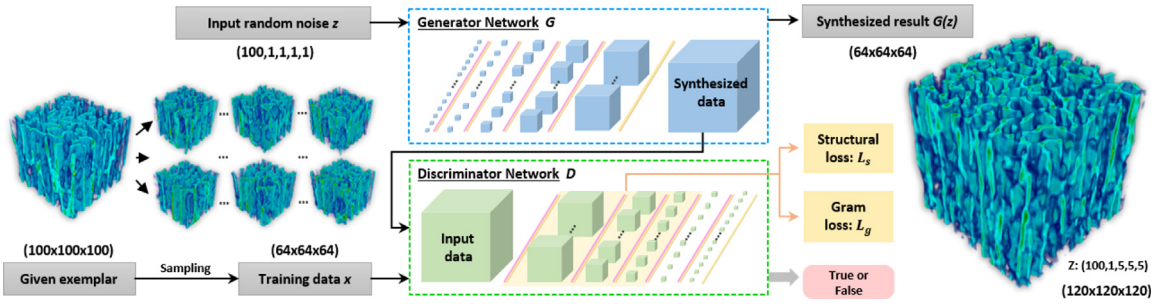

Fig. 2. The architecture of our proposed ScaffoldGAN. Generator G takes a random noise tensor $\mathbf{z}(100, 1, 1, 1, 1)$ as input and outputs a synthesized result $G(\mathbf{z})(64^3)$ voxels during training. Discriminator D is trained to distinguish real training samples $\mathbf{x}$ (extracted from given exemplar(64^3) voxels) from fake synthesized results $G(\mathbf{z})$. After the network is well trained, given a random noise tensor $\mathbf{z}(100, 1, 5, 5, 5)$, we show the synthesized result in the right (120^3) voxels.

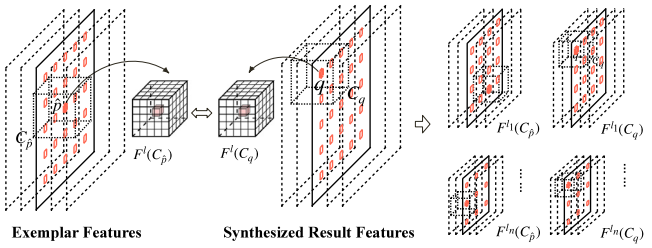

Fig. 3. Illustrate of the cubical region chosen on feature layers.

of width w centered at p. Fig. 3 illustrates the chosen cubical region on feature layers. This is repeated for feature layers in L which construct the look-up table. The anchor points in the look-up table are sampled prior to the training, and the look-up table itself is refreshed once the discriminator is updated during the discriminator training stage for each iteration. Thus, the structural loss is obtained by feeding the synthesized result $G(\mathbf{z})$ to discriminator D and evaluating Eq. (4).

3.3. Gram loss for visual appearance

As shown by previous studies in image synthesis or texture synthesis [17–19], the gram loss, $\mathcal{L}_g$, is helpful for reproducing visually similar appearance of exemplar, and thus considered into our loss function. However, computing the full gram matrices for a volumetric image is prohibitive; therefore, multiple slices with equal spacing sampled from all three dimensions are used to compute the gram matrices. The weight for the gram loss term is derived through careful ablation experiments.

The extended 3D gram loss are described as: $\mathcal{L}_g = \sum_{l\in L} \left\| \mathbf{g}^l_{m,n}(G(\mathbf{z})) - \mathbf{g}^l_{m,n}(\mathbf{x}) \right\|^2$, where $\mathbf{z}$ is the random noise and $\mathbf{x}$ is the training sample. $\mathbf{g}^l_{m,n}(G(\mathbf{z}))$ and $\mathbf{g}^l_{m,n}(\mathbf{x})$ are the gram matrices at layer l of the synthesized result $G(\mathbf{z})$ and the training sample $\mathbf{x}$, respectively. The gram matrix $\mathbf{g}^l_{m,n}$ is the inner product of feature maps $f^l_{m,k}$ and $f^l_{n,k}$. The gram loss can encode the characteristic of 3D complex structures to produce visually plausible synthesized results.

3.4. Network architecture

The architecture of our ScaffoldGAN is illustrated in Fig. 2, which contains two fully convolutional neural networks, i.e., generator G and discriminator D. The discriminator consists of five 3D convolutional layers with 64, 128, 256, 512, and 512 feature maps, respectively. A batch normalization and a leaky ReLU activation function are consecutively applied to each of the first four convolutional layers and a Sigmoid layer is applied to the final layer. The generator consists of five 3D volumetric deconvolutional layers where the number of feature maps at each layer is 512, 512, 256, 128 and 64, respectively. Fully deconvolutional layers are chosen here since they allow generating outputs of arbitrary size from $\mathbf{z}$. Each of the first four deconvolutional layers is followed by a batch normalization and a ReLU activation function, while the final is followed by a Tanh activation. In both convolutional and deconvolutional layers, the size of the kernel filters and the stride is set to 4^3 and 2^3, respectively.

4. Data acquisition and implementation

4.1. Data preparation

Three types of datasets are used for evaluation, a bone dataset and a metal foam dataset that were scanned with micro-CT by ourselves, with an extra public stone dataset [22].

Exemplars in bone and metal foam datasets are acquired from real-world materials via high-resolution micro-CT imaging system (SkyScan 1076). Six bone exemplars are extracted from a trabecular bone scan with 7146 micro-CT slices whose size is 3936×3936 at the resolution of 17 μm. The metal foam dataset contains three categories of exemplars (i.e., Ni, Cu and Al), which also have intricate internal structures similar to trabecular bones. Each exemplar in this dataset consists of 1563 micro-CT slices with 1952×1824 at the resolution of 8 μm.

In order to feed the volumetric data to our framework for training, we down-sample all pieces of materials to a size of 100^3 as given 3D exemplars. It will be shown in the next section, the down-sampled exemplars still well preserve substantial intricate geometric details. Training samples are then extracted from each exemplar (100^3) with a stride of 2 voxels along all axes to form a training dataset. Every training sample has a size of 64^3 voxels, and thus around $6k$ training samples are obtained, ensuring sufficient samples for training.

4.2. Implementation details

Our method is implemented on a desktop featuring an Intel(R) Core(TM) i7-6700 CPU with 3.40 GHz, 32 GB RAM, and an Nvidia GeForce GTX1080 graphics card. Each exemplar can be viewed as a distinctive object due to the differences in their geometry and appearance. Thus, we train a generative model using our ScaffoldGAN for each exemplar. In each iteration, a batch of 100 training samples (of size 64^3 voxels) are randomly selected from each training dataset (formed by $6k$ training samples extracted from an exemplar). The synthesized results are generated by providing with generator G a random tensor $\mathbf{z}$ sampled from the uniform distribution at each iteration. No paired data is needed for training our ScaffoldGAN, which reduces substantial labor. Our method can generate large synthesized results in size (e.g. 200^3

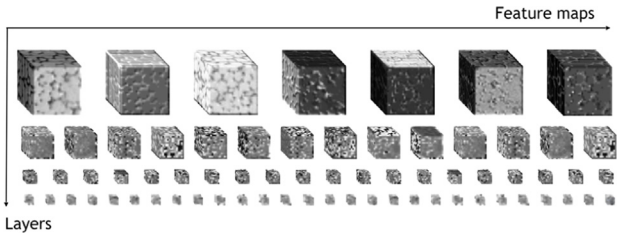

Fig. 4. Visualization of 3D feature maps in different layers of discriminator D (Top to bottom: $layer_0$ to $layer_3$).

voxels), which is larger than the training samples of size 64^3 voxels.

To find out the contribution of each layer to distinguishing structural features in the exemplar, we visualize the feature maps of discriminator D in different layers in Fig. 4. The different feature maps reveal that they encode detailed patterns with various styles. We apply the structural loss to the feature maps at first three convolution layers (l_0 to l_2). Through our observation, the feature maps from l_0 to l_2 can effectively encode spatial features from the training samples. As the size of feature maps drastically decreases at l_3 and l_4, it is difficult to recognize visually significant features. The size of cubical regions for structural loss computation is set to 8^3 for l_0, 4^3 for l_1, and 2^3 for l_2. The sampling strides between neighboring cubical regions are 8, 4 and 2 (along each coordinate axis) for l_0, l_1 and l_2, respectively. These parameters can lead to better spatial coherence, which have been tested through extensive experiments. The weights for balancing the effects of different terms in the loss function are set to be $\lambda_s = 1$, $\lambda_g = 0.1$. More details could be found in the ablation study (Fig. 13) on effects of different loss terms and our design choice of weights. ADAM is chosen as our optimizer with the settings of [44].

To have a better understanding of how our ScaffoldGAN is trained and its training stability, we study the training behavior of our network by visualizing the training curves of generator and discriminator losses with synthesized results at different iterations in Fig. 5. The first circle demonstrates a synthetic result at initial several thousand iterations, which is difficult to recognize any desired structures because the generator has not been fully trained. After fifteen thousand iterations, the details in the synthetic result are refined and structures are much clearer as visualized in the 2*nd* and 3*rd* circles in Fig. 5. When it comes to forty thousand iterations as visualized in the 4*th* circle, the synthesized structures become even clearer that could resemble the exemplar and finally remain almost unchanged as iteration proceeds.

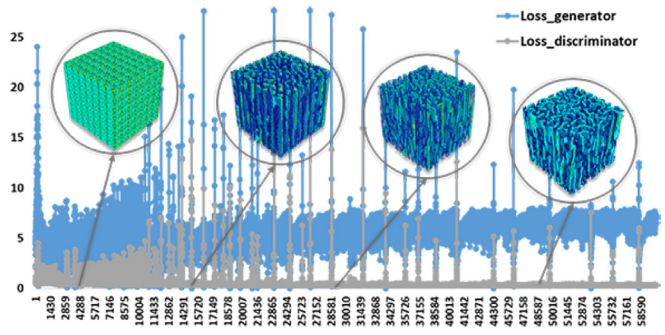

Fig. 5. Training curve and synthetic results visualization from generator G at different iterations.

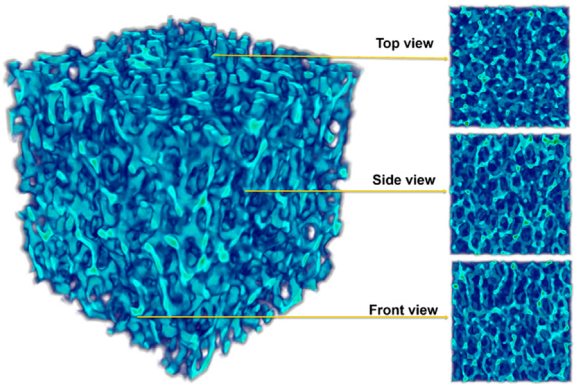

Fig. 6. A 3D synthesized result from foam Ni and its cross-sectional views from different directions. The input is shown in Fig. 7.

5. Experimental results and discussion

We train our network via the combined objective function, the synthesized results are generated by passing the generator G a random noise $\mathbf{z}$ sampled from the uniform distribution.

We first show the synthesized result of *metal foam Ni* is shown in Fig. 6 with top, side and front view, where the given exemplar foam Ni is shown in Fig. 7. Next, synthesized results with a very large size (420^3 voxels) are shown in Figs. 9 and 10. Finally, we also rendered, in Fig. 11, the cross-sectional views of the above 3D synthesized results with 50 consecutive slices chosen randomly.

Synthesized results show that our method could generate compelling results with complex structures similar to the exemplars but in a large size. Overall, the spatial coherence is well preserved as seen from these visualized results, and few blurry or repetitive parts can be observed. We will make our datasets (including both the bone dataset and the metal foam dataset) publicly available for future studies.

To validate the diversity of our synthesized results produced by our networks, we generate the training dataset which contains a large number of samples with rich variations covering the data space. We randomly sample a large number of noise tensors $\mathbf{z}$ for each type of data and pass them to generator G. From the synthesized results, we observed that our trained networks are able to generate results with diverse patterns, experimentally indicating that there is little chance that mode collapse would occurs for both exemplars. For each exemplar (the bone or the foam Ni) three synthesized outputs are randomly selected and illustrated in Fig. 8, showing that the synthesized results have quite different microstructures.

6. Evaluation and ablation study

We validate our method via qualitative and quantitative comparisons with the state-of-the-art methods: (1) GAN-based method: Wu et al. [40] (BaselineGAN with L_{adv} only), PatchGAN [39]; (2) Conventional example-based methods: Kopf et al. [12], Chen and Wang [13] and Zhang et al. [14]. We also conduct the ablation study to demonstrate the effectiveness of the proposed structural loss L_s.

6.1. Qualitative evaluation

Fig. 7 shows the comparison of results on visual appearance using our method and the state-of-the-arts. Two exemplars with highly distinctive materials, $Bone_0$ and *metal foam Ni*, are chosen to demonstrate the performance of our method. The cross-sectional views are rendered with 10 consecutive slices using ParaView [47].

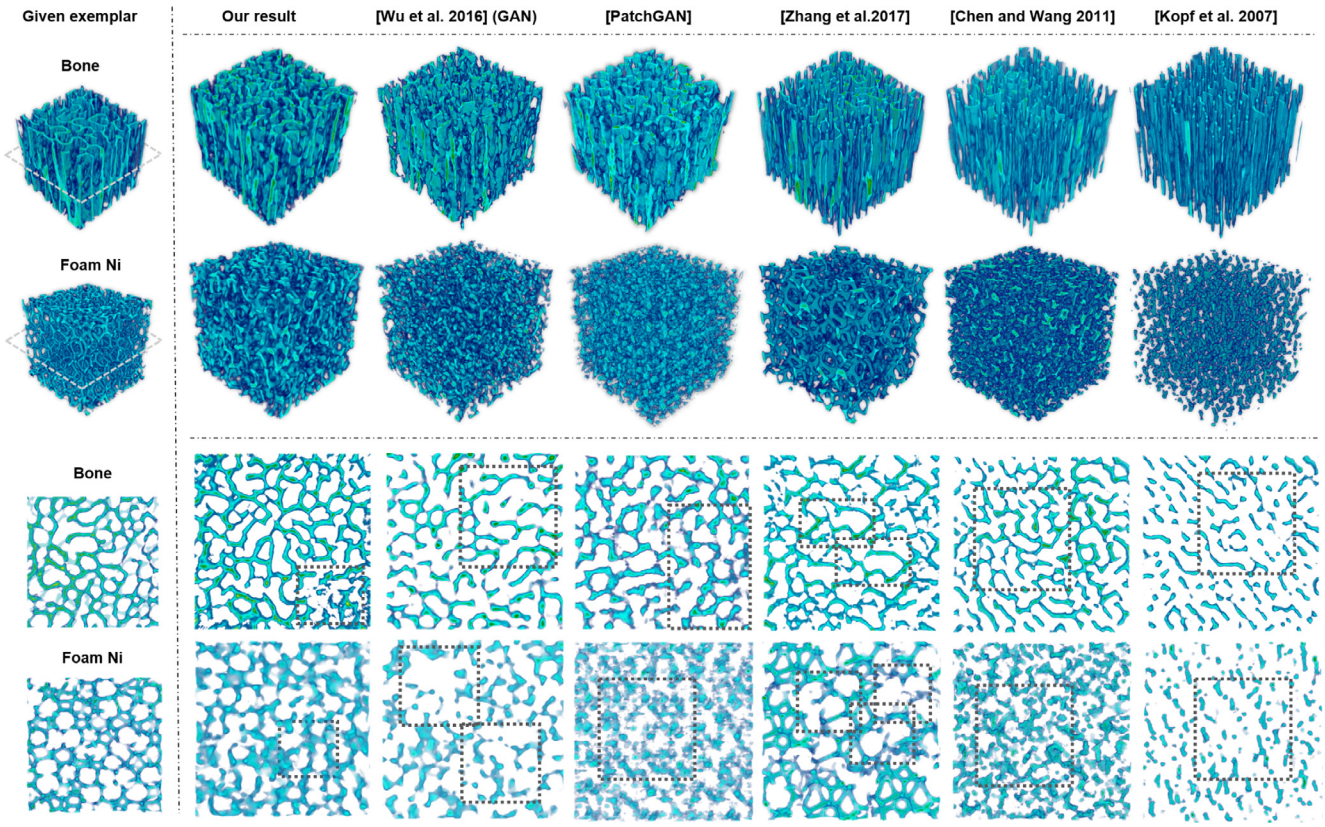

Fig. 7. Visualization and qualitative comparison of our synthesized results with those produced by state-of-the-art methods. 3D renders of synthesized results (120^3 voxels) of ours and the state-of-the-arts methods based on $Bone_0$ and *metal foam Ni* are shown at top two rows. 2D renders of 10 consecutive slices randomly selected from corresponding 3D synthesized results are shown in the bottom two rows. Rectangles highlight failure parts in the results, e.g., voids in Wu et al. repeated parts in Zhang et al. and undesired intertwined structures in ours.

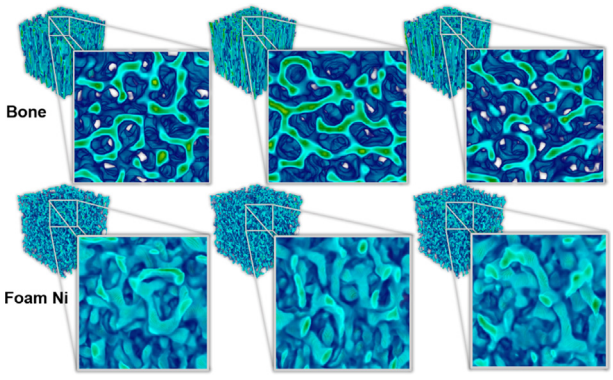

Fig. 8. Diversity analysis of synthesized results generated from three different noise tensors **z** given the bone (shown in top row) and the foam Ni (bottom row) as exemplars.

For the GAN-based method, the generic framework of Wu et al. [40] (BaselineGAN with L_{adv} only) for 3D content modeling is considered as our baseline. As shown in Fig. 7, the results by Wu et al. suffer from the difficulty of reproducing the local structure similar to the given exemplars. This can be easily seen from the top of the 3D result of $Bone_0$ and is very obvious in the 3D result of *metal foam Ni*. From the cross-sectional views, more disconnected branches reveal in the bone result by Wu et al. than in ours. Large voids (highlighted by dashed rectangles) can be observed in the results of *metal foam Ni*. In order to further prove that our method outperforms PatchGAN [39], we generalize it to 3D case and compare their results with ours. Similar problems also exist on the synthesized result of PatchGAN, especially on the synthesized result of *metal foam Ni* where there are many disconnected part and floating structures. Both indicate that the baseline model alone (with only L_{adv}) and patchGAN cannot well preserve the spatial coherence and visual appearance observed in the exemplar.

For conventional example-based methods, disconnected structures are easily observed in 3D results by methods based on 2D information [12] and [13] and highlighted by the dashed rectangles in the cross-sectional views. Using volumetric exemplars as input, results by Zhang et al. [14] could preserve fine-grained, cellular features in a degree. However, repeated artifacts (for $Bone_0$) and low spatial coherence (for *metal foam Ni*) are found in the regions highlighted by the rectangles in Fig. 7.

Overall, our results are better than the others through visual examination. For $Bone_0$, our result has a higher degree of spatial coherence and strong connectivity similar to the exemplar as can be seen from both the 3D model and the cross-sectional view. For *metal foam Ni*, cellular structures are perceivable in our result (as well as in the result by [14]), while other methods fail to reproduce this geometric feature. We also highlighted some of minor issues observed in our results with the dashed rectangles. In summary, our proposed method can effectively enhance the visual quality of the synthesized results from spatial coherence and appearance, and thus outperforms the state-of-the-art methods.

6.2. Quantitative evaluation

To validate our method on spatial coherence and morphological properties, we conduct experiments for each exemplar in the bone dataset and a *metal foam Ni* exemplar, and report

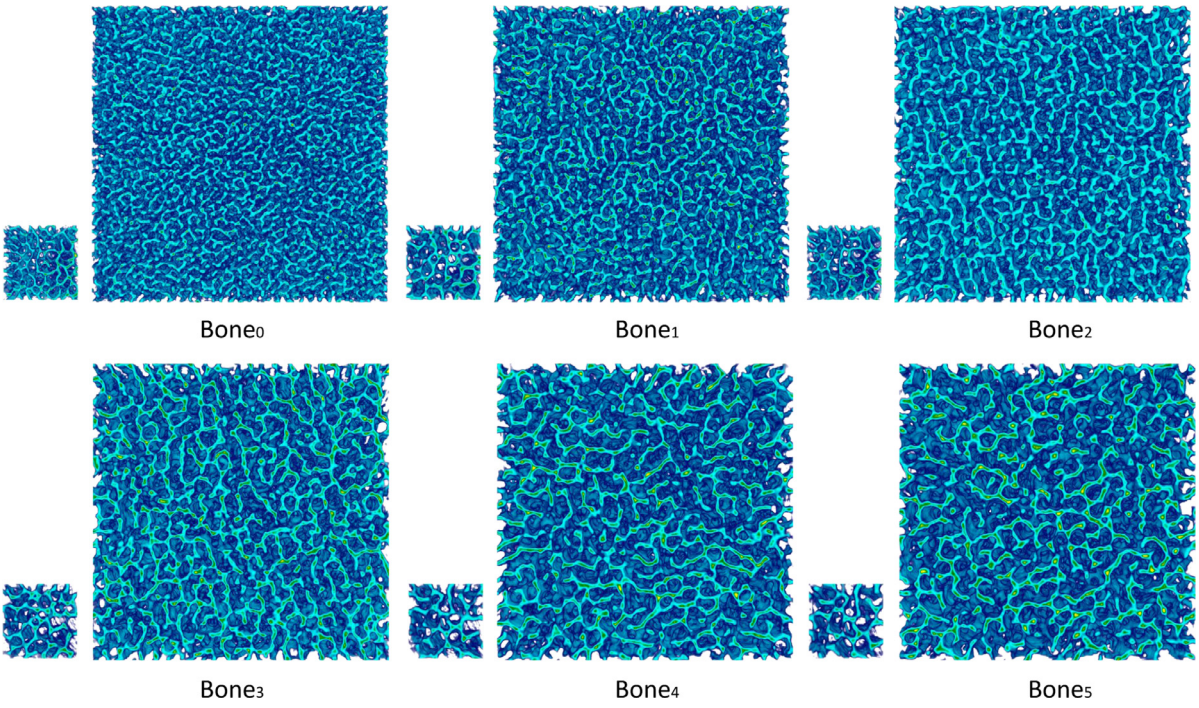

Fig. 9. Large-size synthesized results of different exemplar input from different regions of bone: Each bone exemplar (shown in left, 100^3 voxels) and its 3D synthesized result (shown in right, 420^3 voxels).

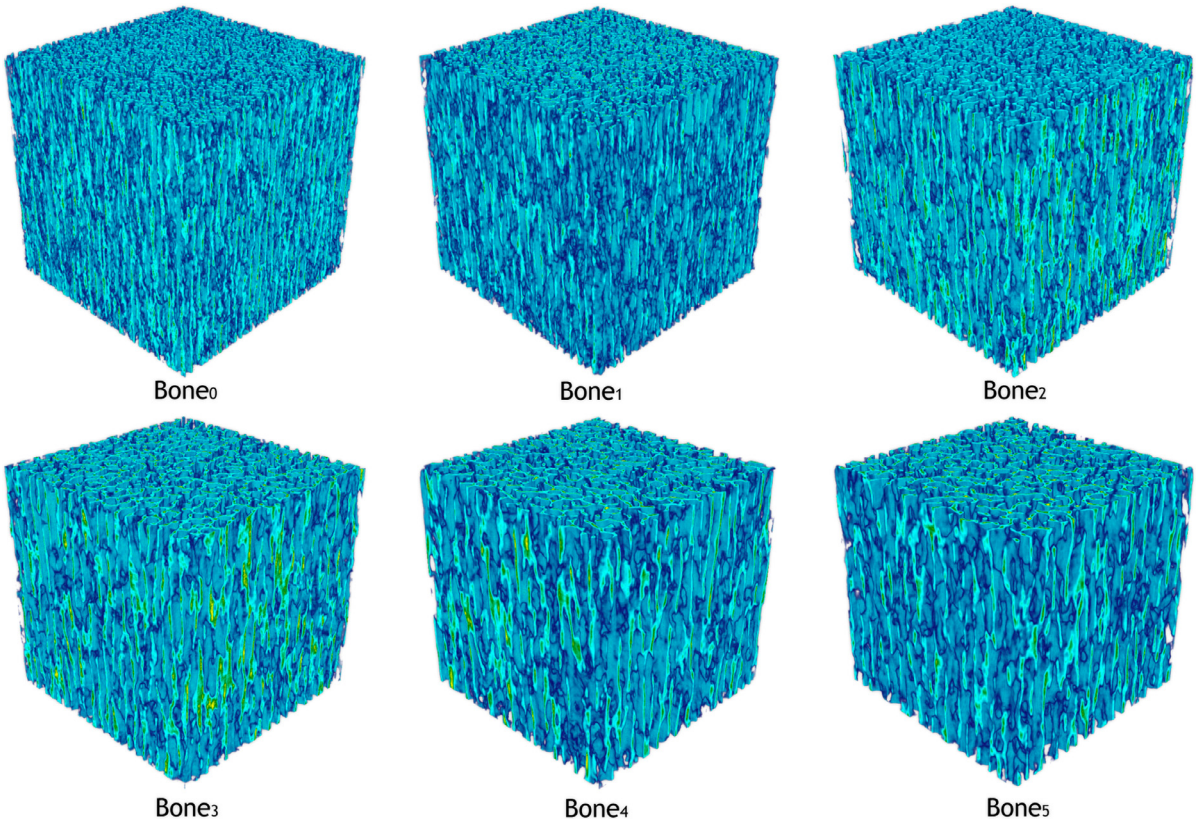

Fig. 10. Visualization of 3D view of the synthesized results (420^3 voxels), bone exemplar are $Bone_0$, $Bone_1$, $Bone_2$, $Bone_3$, $Bone_4$ and $Bone_5$.

Table 1
Quantitative comparison with state-of-the-art methods on $Bone_0$ (columns 2–5) and *metal foam Ni* (columns 6–9). Listed are relative errors of porosity (ε_{err}%), l_2 difference of two-point correlation functions ($D_2(\phi)$) between the synthesized results and the exemplars and connectivity measurements, i.e., the number of connected components (N_{cc}), the number of floating parts (N_f) and the Wasserstein distance (D_w) for connectivity measure. Average timing (s) for generation is listed in last column. The time for training each GAN-based model is around 4 hrs. Rows 2 shows the corresponding ground truth values of the given exemplars.

Methods	ε_{err}%	$D_2(\phi)$	N_{cc}/N_f	D_w	ε_{err}%	$D_2(\phi)$	N_{cc}/N_f	D_w	Time (s)
Metrics (Ground Truth)	73.66	*N.A.*	6/4	*N.A.*	88.24	*N.A.*	4/2	*N.A.*	*N.A.*
Kopf et al. [12]	20.27	43.39	268/159	5.18	10.07	44.88	878/121	2.35	1922
Chen et al. [13]	13.78	32.39	116/106	5.08	3.77	36.91	158/74	2.30	2164
Zhang et al. [14]	2.86	8.97	20/11	4.02	2.11	13.67	22/13	1.91	1441
Wu et al. [40] (baseline GAN L_{adv})	1.51	5.95	40/17	4.60	2.02	18.30	28/8	2.00	0.094
PatchGAN [39]	1.49	5.83	30/23	4.38	1.89	16.31	34/14	2.16	0.095
baseline GAN L_{adv} + L_g	1.16	5.79	26/19	4.26	1.70	15.39	17/9	1.76	0.095
baseline GAN L_{adv} + L_s	0.71	3.27	8/5	1.79	0.85	6.11	8/4	0.74	0.096
Ours ($L_{adv} + L_g + L_s$)	**0.59**	**3.05**	**6/3**	**1.53**	**0.73**	**5.23**	**6/3**	**0.51**	0.096

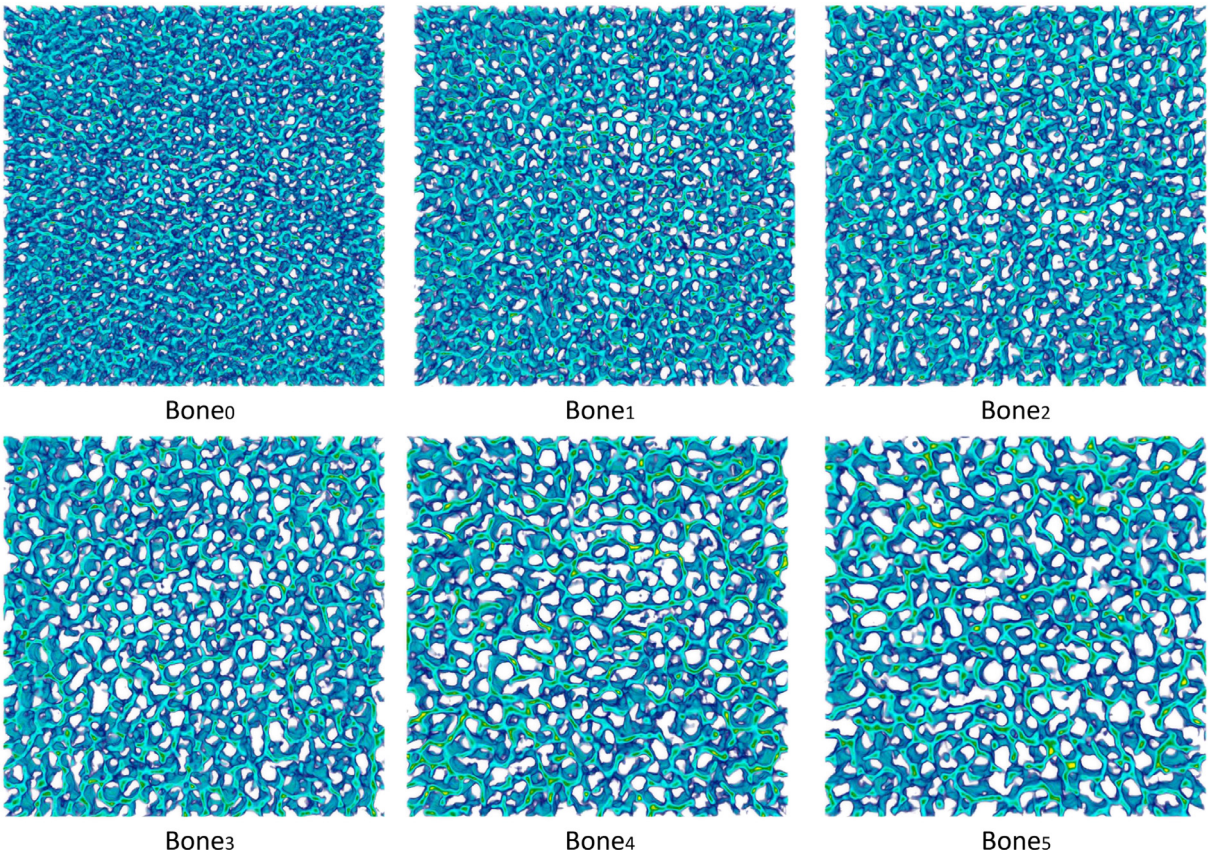

Fig. 11. Visualization of 3D cross-sectional views of the above 3D synthesized results with 50 consecutive slices chosen from synthesized results randomly (420 × 420 × 50 voxels).

the averaged value of each metric from 10 randomly synthesized results for corresponding exemplar.

Porosity ε [15] and two-point correlation function ϕ [16] are adopted in our evaluation, which are two frequently used metrics for characterizing materials. We also measure the connectivity of the synthesized results based on three metrics, i.e., the number of connected components (N_{cc}), the number of floating parts (N_f), and a measure (D_w) based on the Wasserstein distance with respect to the size of connected components. We will explain in detail the definition of each metric later.

Table 1 shows the comparison among different methods on the $Bone_0$ exemplar and the *metal foam Ni* exemplar. To validate the effectiveness of our method on various exemplars, we also list in Table 2 the metric values of another five exemplars extracted from different regions of a vertebral bone from the bone dataset (see Fig. 1(a)).

Table 2
Ground truth metric values (Rows 2–3) and those of synthesized results (averaged over 10 results) of each bone exemplar ($Bone_1$ to $Bone_5$) are presented. Results show that our method performs stably on various exemplars.

Category	$Bone_1$	$Bone_2$	$Bone_3$	$Bone_4$	$Bone_5$	Average
ε	72.50	70.42	69.98	72.63	73.06	*N.A.*
N_{cc}/N_f	6/3	6/4	5/3	4/2	8/6	*N.A.*
ε_{err}%	0.54	0.61	0.67	0.66	0.55	0.60
$D_2(\phi)$	4.26	3.50	2.01	3.79	4.61	3.54
N_{cc}/N_f	9/3	6/4	7/4	8/5	11/4	*N.A.*
D_w	1.87	1.70	2.73	2.12	1.17	1.85

(a) Evaluation on Porosity.

The porosity ε is a measure of the void volume fraction in a specific volume, which is defined as: $\varepsilon = |\mathcal{V}|/|V|$, where $|\mathcal{V}|$ and

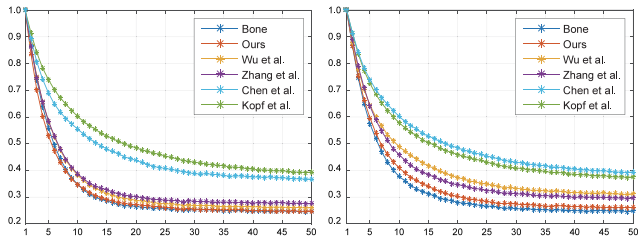

Fig. 12. Two-point correlation functions ϕ ($Bone_0$ and *metal foam Ni*) for our method and the state-of-the-arts methods. Our results well approximate the exemplar.

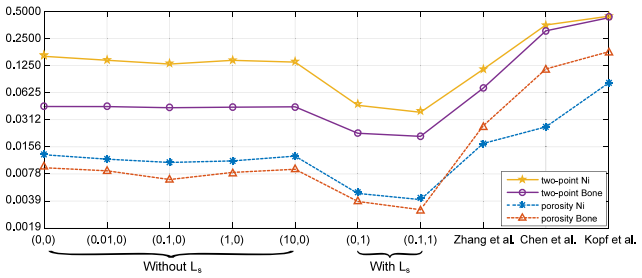

Fig. 13. Ablation study on different loss terms. Values in pair indicate (λ_g, λ_s). Decreases in relative errors of the porosity and two-point correlation function differences can be seen when the structural loss is introduced at (0.1, 1). l_2 difference of two-point correlation functions and the porosity are denoted as *Two-point* and *Porosity*, respectively. (Vertical axis log_2).

$|V|$ are the numbers of void voxels and total voxels in the 3D model, respectively. The relative error between the exemplar and the synthesized is measured as $\varepsilon_{err} = |\varepsilon_E - \varepsilon_S|/\varepsilon_E$. As seen from Table 1, the porosity of our synthesized results is much closer to that of each given exemplar. For $Bone_0$ and *metal foam Ni*, the relative error of porosity to the exemplar (err_ε) of our synthesized results are 0.59% and 0.73%, respectively. Both values are half smaller than those (1.16% and 1.70%) produced by baselineGAN with L_g ($L_{adv}+L_g$). Besides, both values of our approach are largely smaller than PatchGAN [39] (1.49% and 1.89%), and other conventional example-based methods [12–14]. This illustrates that our approach ($L_{adv} + L_g + L_s$) is capable of preserving the porosity of synthesized results similar to that of the given exemplar. Table 2 shows that the porosity of synthesized results is stable.

(b) Evaluation on Two-point correlation function.

Two-point correlation function is a statistical measure for characterizing distribution of different materials in a volume [48], indicating the pore size and spatial coherence. Given a two-phase function $\phi(\mathbf{v})$ defined on domain V which equals 1 when $\mathbf{v}$ lies in the solid phase or 0 in the void, a general definition of the two-point correlation function is as follow: $P(\mathbf{r}) = \langle\phi(\mathbf{v}_1), \phi(\mathbf{v}_2)\rangle$, where $\mathbf{r} = \mathbf{v}_1 - \mathbf{v}_2$ and $\langle\cdot,\cdot\rangle$ denotes the expectation over domain V. In our evaluation, we define the binarized volumetric results as the phase function ϕ and compute $P(|\mathbf{r}|)$ with the distance $|\mathbf{r}|$ ranging from 1 to 50 voxels (neglecting the direction of $\mathbf{r}$ for simplicity). The corresponding two-point correlation functions of the synthesized results are plotted in Fig. 12.

The l_2 differences, denoted $D_2(\phi)$, of two-point correlation functions between synthesized results by different methods and the exemplar are listed in Table 1 for comparison. Our results deviate the least from the exemplars for $Bone_0$ and *metal foam Ni* (3.05 and 5.23, respectively). Differences of the rest of results are more than double (for $Bone_0$) or triple (for *metal foam Ni*) of ours. As it can be seen from Table 1, PatchGAN [39] produces results only comparable with those by the adversarial and gram loss (L_{adv} + L_g), while our methods adding the structural loss (L_{adv} + L_s) lead to large improvement. In particular, lower values of $D_2(\phi)$ and D_w in ours reflect high similarity of spatial coherence between the original and synthesize materials. More statistics on other bone exemplars can be found in Table 2. This clearly evidences the effectiveness of the proposed structural loss in enhancing spatial coherence of the synthesized results.

(c) Evaluation on Connectivity.

We also measure the connectivity of the synthesized results to demonstrate the effects of the combination of the structural loss and the adversarial loss on connectivity. To this end, we first count the number of connected components (N_{cc}) and of floating parts (N_f). A connected component is considered as a floating part if its size is smaller than 10 voxels. Further, we define another connectivity measure based on the Wasserstein distance between two sets of connected components from two respective volumetric models, which is defined as:

$$D_w = \mathrm{WD}(H_S, H_E), \tag{5}$$

where H is the set of connected components in a volumetric model, and the size of each connected component is normalized by the volume of the model. Since the Wasserstein distance measures the difference between two distributions, D_w in fact reflects similarity between the two models in terms of how voxels are distributed and grouped into connected components.

As seen from Table 1, our results have the minimal number of connected components N_{cc} (6) and floating part N_f (3) for both $Bone_0$ and *metal foam Ni* compared to other methods; N_{cc} and N_f of the baseline GAN model (40/17 for $Bone_0$ and 28/8 for *Ni*, respectively) are several times of ours; Zhang et al. and the baseline GAN model plus the gram loss produce slightly better results than the baseline GAN model, but still not comparable to ours. In addition, values of Wasserstein distance based metric D_w for ours (1.53 and 0.51 for the two exemplars, respectively) show that our results are much closer to the exemplars in terms of spatial coherence and connectivity. The averaged value of D_w of synthesized results based on the other five bone exemplars is 1.85 (see Table 2), showing that the connectivity of synthesized results is stable and quite similar to the exemplars.

6.3. Ablation study

To validate our design choice, we conduct ablation experiments (on two exemplars, $Bone_0$ and *metal foam Ni*) to compare each loss term in the full loss (L_{adv} + $\lambda_g L_g$ + $\lambda_s L_s$). First, we train our network with only L_{adv} (with λ_g and λ_s set to 0), which is equivalent to Wu et al. [40] (baselineGAN). Then, we consider the effect of L_g by adding it to L_{adv} (L_{adv} + $\lambda_g L_g$). We gradually increase the weight λ_g = 0.01, 0.1, 1 and 10 and train the network. We find that the best performance in such setting is when the weight equals 0.1. Next, we train the network by only using the structure loss L_s and set the λ_g to zero (L_{adv} + $\lambda_s L_s$), where the related error has dropped obviously. Finally, we train the network by incorporating our proposed structural loss L_s and the previous setting (λ_g = 0.1) together ($L_{adv}+0.1L_g+\lambda_s L_s$), which produces the best results. Performances in terms of the porosity and two-point correlation difference (in l_2 sense) are plotted and compared in Fig. 13.

From the figure, we can see that the synthesized result use baselineGAN with only L_{adv} has similar performance with Zhang et al. [14]. By adding weighted L_g, the relative errors of porosity to the input decrease slightly. However, little improvement in terms of the two-point correlation function difference is observed, indicating the gram loss L_g alone cannot enhance the spatial coherence, which validates our design choice. Better performance is obtained when the structural loss L_s is added. Especially for 2-point correlation (widely used to characterize material distribution), the relative error derived with L_s (Ni) reaches 5.23%,

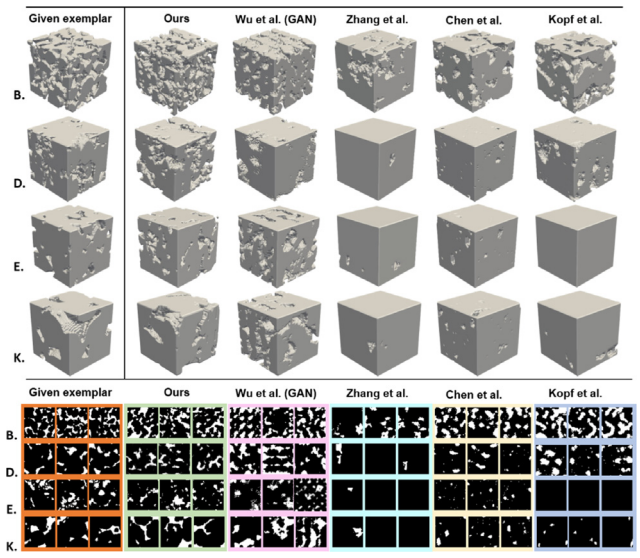

Fig. 14. Qualitative comparison with state-of-the-art methods on the ICL dataset. Synthesized results and three corresponding slices are shown. Our method also produces better results than others in terms of both visual appearance and spatial coherence.

only one third of the most competing result derived without L_s (15.39%, λ_g = 0.1). In particular, two-point correlation function differences with respect to both exemplars are remarkably reduced. Connectivity metrics also show the necessary strengths of L_s. In summary, while the GANs framework (with L_{adv}) generally shows better performances over the conventional methods, our proposed structural loss brings remarkable decrease in all metrics considered here for both exemplars.

6.4. Experiments on more datasets

To analyze the generalizablity of our method, we also present the results on synthesis of four stone exemplars (i.e., *Bentheimer*, *Doddington*, *Estaillades* and *Ketton*) from the dataset [22] and two *metal foams Cu* and *Al*, which also have highly complex internal structures similar to bones. For each exemplar, we randomly generate 10 different results and calculate the mean value for each statistical metric. The comparison on metrics between our methods and the state-of-the-arts is shown in Table 3, and visual comparison containing synthesized results and multiple 2D slices are shown in Fig. 14. These experiments illustrate that our proposed method can also produce convincing results in terms of both statistical metrics and visual appearance, which is especially obvious for *Bentheimer*.

6.5. 3D Fabricated results

We fabricate the digitally synthesized materials obtained by our ScaffoldGAN using a photosensitive resin 3D printer. We post-process the synthesized bone scaffolds by extracting the iso-surface of solid structures for manufacturing. We show the scaffold material from real world and show the structurally complex skeleton of the scaffold materials together with the synthesized result from our network.

As can be seen from Fig. 15, the 3D printed prototypes retain rich fine-grained details and exhibit high visual similarity with the corresponding exemplars, implying that synthesized results by our method can be reliably fabricated while maintaining the design intent. Physically manufacturing bone scaffold may be difficult due to complex geometries, varying mechanical properties

Table 3
Quantitative comparison with state-of-the-art methods on the ICL dataset and *metal foam Cu* and *Al*. Mean values of different metrics are shown for each exemplar.

Category	Benth.	Doddi.	Estai.	Ketto.	Cu	Al
Exemplar ε %	19.34	15.18	13.56	12.10	86.80	84.99
εOurs w L_s	**1.79**	**3.24**	**0.83**	**0.27**	**3.96**	**2.03**
ε w/o L_s	10.13	9.39	11.01	9.85	5.12	4.98
ε Wu et al. [40]	11.72	11.23	12.50	7.89	5.89	6.49
ε Zhang et al. [14]	1.85	11.41	13.31	11.39	10.19	12.87
ε Chen et al. [13]	28.69	4.31	2.57	4.89	4.87	5.46
ε Kopf et al. [12]	6.29	16.16	9.51	9.42	5.94	5.24
ϕOurs w L_s	**0.36**	**1.29**	**1.69**	**1.53**	**0.36**	**0.40**
ϕ w/o L_s	1.24	2.01	1.96	2.73	1.49	2.15
ϕ Wu et al. [40]	0.40	3.10	6.89	2.95	2.30	2.81
ϕ Zhang et al. [14]	3.51	4.86	9.29	5.34	4.06	4.39
ϕ Chen et al. [13]	0.91	7.14	1.99	4.48	0.90	1.20
ϕ Kopf et al. [12]	1.33	4.13	6.87	2.16	1.10	1.63

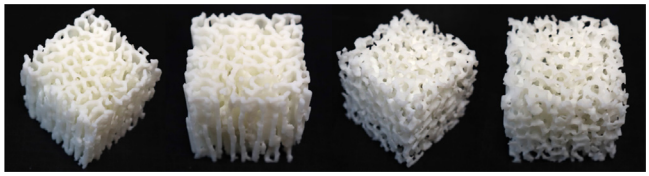

Fig. 15. Two views of the 3D printed synthesized result based on $Bone_0$ (left) and two views based on *the metal foam Ni* (right) are shown. Rich fine-grained details can be maintained after converting our digital results to physical prototypes.

and manufacturability of the models. Hence, success fabrication of the synthesized results indicates that the results are ready for physical prototyping. Thus, it shows our approach could achieve complex bone scaffold synthesis given a natural exemplar and produce manufacturable results that ensure the pore size and inter-connectivity of the scaffolds to be similar to the exemplar.

7. Limitation and future work

Training a general network using all categories of microstructures is very challenging and also the limitation. Currently, training with mixture of data does not generate good results. This is a very interesting question and also one of our future work. Since our approach is data-driven based, the difficulty of obtaining dataset is another limitation, which is difficult for us to test the model on more diverse data. Careful modulation is required to separate data from different categories in the latent embedding space. One future work is to show the functional appropriateness such as endurance to impacts/loads and biological compatibility. Another one is to use sketch or other method to control the synthesized result with different direction and scale, also with different porosity. This makes the final result is regional controllable and could generate the desired 3D shape with complex internal structures.

8. Conclusion

In this paper, we have presented ScaffoldGAN, an efficient end-to-end framework based on the generative adversarial networks and a novel structural loss, to synthesize complex scaffold materials via a single exemplar. With the well-trained generator, we are able to synthesize 3D models of materials with intricate internal structures within seconds. We also collect two datasets containing six bone exemplars from different regions of a real vertebral bone and three metal foams, all of which are acquired via high-resolution microCT scanning. Extensive experimental

results show that our method is able to reproduce 3D scaffold materials with visual appearance, spatial coherence, and statistical metrics closely resembling the given exemplars. Our method also outperforms both conventional example-based texture synthesis methods and the baseline GAN model plus the gram loss. This, together with the ablation study, demonstrates such appealing properties are attributed to the combination of adversarial nets and the deep structural loss. State-of-the-art performance on the ICL dataset (containing four stone exemplars) also indicates that our method can be generalizable to other similar tasks of synthesizing heterogeneous materials. We will make our collected datasets publicly available to facilitate future studies in the relevant communities. How to exploit deep representations to design effective supervision for recovering fine-grained structures in high fidelity observed in the materials is our future work.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We would like to thank the anonymous reviewers for their comments. This work is supported by National Natural Science Foundation of China (NSFC 61572021) and Hong Kong GRF (HKU 717813E). Thanks Microsoft Research Asia for their support.

References

[1] Leong K, Chua C, Sudarmadji N, Yeong W. Engineering functionally graded tissue engineering scaffolds. J Mech Behav Biomed Mater 2008;1(2):140–52.
[2] Smith M, Flanagan C, Kemppainen J, Sack J, Chung H, Das S, et al. Computed tomography-based tissue-engineered scaffolds in craniomaxillofacial surgery. Int J Med Robot Comput Assist Surgery 2007;3(3):207–16.
[3] Bucklen B, Wettergreen W, Yuksel E, Liebschner M. Bone-derived CAD library for assembly of scaffolds in computer-aided tissue engineering. Virtual Phys Prototyping 2008;3(1):13–23.
[4] Derby B. Printing and prototyping of tissues and scaffolds. Science 2012;338(6109):921–6.
[5] Khan Y, Yaszemski MJ, Mikos AG, Laurencin CT. Tissue engineering of bone: material and matrix considerations. JBJS 2008;90:36–42.
[6] Lu L, Sharf A, Zhao H, Wei Y, Fan Q, Chen X, Savoye Y, Tu C, Cohen-Or D, Chen B. Build-to-last: strength to weight 3D printed objects. ACM Trans Graph 2014;33(4):97.
[7] Martínez J, Dumas J, Lefebvre S. Procedural voronoi foams for additive manufacturing. ACM Trans Graph 2016;35(4):44.
[8] Portilla J, Simoncelli EP. A parametric texture model based on joint statistics of complex wavelet coefficients. Int J Comput Vis 2000.
[9] Efros AA, Leung TK. Texture synthesis by non-parametric sampling. In: Proceedings of the seventh IEEE international conference on computer vision. 2, IEEE; 1999, p. 1033–8.
[10] Liang L, Liu C, Xu Y-Q, Guo B, Shum H-Y. Real-time texture synthesis by patch-based sampling. ACM Trans Graph 2001;20(3):127–50.
[11] Wei L-Y, Han J, Zhou K, Bao H, Guo B, Shum H-Y. Inverse texture synthesis. AcM Trans Graph 2008;27(3):52.
[12] Kopf J, Fu C-W, Cohen-Or D, Deussen O, Lischinski D, Wong T-T. Solid texture synthesis from 2d exemplars. ACM Trans Graph 2007;26(3):2.
[13] Chen J, Wang B. High quality solid texture synthesis using position and index histogram matching. Vis Comput 2010;26(4):253–62.
[14] Zhang H, Chen W, Wang B, Wang W. By example synthesis of three-dimensional porous materials. Comput Aided Geom Design 2017;52:285–96.
[15] Karageorgiou V, Kaplan D. Porosity of 3D biomaterial scaffolds and osteogenesis. Biomaterials 2005;26(27):5474–91.
[16] Berryman JG, Blair SC. Use of digital image analysis to estimate fluid permeability of porous materials: Application of two-point correlation functions. J Appl Phys 1986;60(6):1930–8.
[17] Gatys L, Ecker AS, Bethge M. Texture synthesis using convolutional neural networks. In: Advances in neural information processing systems. 2015, p. 262–70.
[18] Jetchev N, Bergmann U, Vollgraf R. Texture synthesis with spatial generative adversarial networks. 2016, arXiv preprint arXiv:1611.08207.
[19] Bergmann U, Jetchev N, Vollgraf R. Learning texture manifolds with the periodic spatial GAN. In: Proceedings of the 34-th international conference on machine learning. 2017.
[20] Wu Z, Song S, Khosla A, Yu F, Zhang L, Tang X, Xiao J. 3d shapenets: A deep representation for volumetric shapes. In: Proceedings of the IEEE conference on computer vision and pattern recognition. 2015, p. 1912–20.
[21] Li C, Pan H, Liu Y, Tong X, Sheffer A, Wang W. Robust flow-guided neural prediction for sketch-based freeform surface modeling. ACM Trans Graph 2018;37(6):238:1–238:12.
[22] Muljadi BP, Blunt MJ, Raeini AQ, Bijeljic B. The impact of porous media heterogeneity on non-Darcy flow behaviour from pore-scale simulation. Adv Water Resour 2016;95:329–40.
[23] Giannitelli S, Accoto D, Trombetta M, Rainer A. Current trends in the design of scaffolds for computer-aided tissue engineering. Acta Biomater 2014;10(2):580–94.
[24] Wettergreen MA, Bucklen BS, Starly B, Yuksel E, Sun W, Liebschner MAK. Creation of a unit block library of architectures for use in assembled scaffold engineering. Comput Aided Des 2005;37(11):1141–9.
[25] Efros AA, Freeman WT. Image quilting for texture synthesis and transfer. In: Proceedings of the 28th annual conference on computer graphics and interactive techniques. ACM; 2001, p. 341–6.
[26] Han C, Risser E, Ramamoorthi R, Grinspun E. Multiscale texture synthesis. ACM Trans Graph 2008;27(3):51.
[27] Kwatra V, Essa I, Bobick A, Kwatra N. Texture optimization for example-based synthesis. ACM Trans Graph 2005;24(3):795–802.
[28] Wei L-Y, Lefebvre S, Kwatra V, Turk G. State of the art in example-based texture synthesis. In: Eurographics 2009, state of the art report, EG-STAR. Eurographics Association; 2009, p. 93–117.
[29] Liu X, Shapiro V. Random heterogeneous materials via texture synthesis. Comput Mater Sci 2015;99:177–89.
[30] Johnson J, Alahi A, Fei-Fei L. Perceptual losses for real-time style transfer and super-resolution. In: European conference on computer vision. Springer; 2016, p. 694–711.
[31] Li C, Wand M. Precomputed real-time texture synthesis with markovian generative adversarial networks. In: European conference on computer vision. Springer; 2016, p. 702–16.
[32] Ulyanov D, Vedaldi A, Lempitsky V. Improved texture networks: Maximizing quality and diversity in feed-forward stylization and texture synthesis. In: Proceedings of the IEEE conference on computer vision and pattern recognition. 2017, p. 6924–32.
[33] Elad M, Milanfar P. Style transfer via texture synthesis. IEEE Trans Image Process 2017;26(5):2338–51.
[34] Li Y, Fang C, Yang J, Wang Z, Lu X, Yang M-H. Diversified texture synthesis with feed-forward networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition. 2017, p. 3920–8.
[35] Huang H, Kalogerakis E, Yumer E, Mech R. Shape synthesis from sketches via procedural models and convolutional networks. IEEE Trans Vis Comput Graph 2017;23(8).
[36] Li C, Wand M. Combining markov random fields and convolutional neural networks for image synthesis. In: Proceedings of the IEEE conference on computer vision and pattern recognition. 2016, p. 2479–86.
[37] Ulyanov D, Lebedev V, Vedaldi A, Lempitsky VS. Texture networks: Feed-forward synthesis of textures and stylized images.. In: ICML. 2016, p. 1349–57.
[38] Sendik O, Cohenor D. Deep correlations for texture synthesis. ACM Trans Graph 2017;36(4):1.
[39] Isola P, Zhu J-Y, Zhou T, Efros AA. Image-to-image translation with conditional adversarial networks. In: Proceedings of IEEE CVPR. 2017, p. 1125–34.
[40] Wu J, Zhang C, Xue T, Freeman B, Tenenbaum J. Learning a probabilistic latent space of object shapes via 3d generative-adversarial modeling. In: Advances in neural information processing systems. 2016, p. 82–90.
[41] Yan X, Yang J, Yumer E, Guo Y, Lee H. Perspective transformer nets: Learning single-view 3d object reconstruction without 3d supervision. In: Advances in neural information processing systems. 2016, p. 1696–704.
[42] Choy CB, Xu D, Gwak J, Chen K, Savarese S. 3d-r2n2: A unified approach for single and multi-view 3d object reconstruction. In: European conference on computer vision. Springer; 2016, p. 628–44.
[43] Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, et al. Generative adversarial nets. In: Advances in neural information processing systems. 2014, p. 2672–80.

[44] Radford A, Metz L, Chintala S. Unsupervised representation learning with deep convolutional generative adversarial networks. 2015, preprint arXiv:1511.06434.

[45] Wang W, Huang Q, You S, Yang C, Neumann U. Shape inpainting using 3d generative adversarial network and recurrent convolutional networks. In: Proceedings of the IEEE international conference on Computer Vision. 2017, p. 2298–306.

[46] Mosser L, Dubrule O, Blunt MJ. Reconstruction of three-dimensional porous media using generative adversarial neural networks. Phys Rev E 2017;96(4):043309.

[47] Ahrens J, Geveci B, Law C. Paraview: An end-user tool for large data visualization. In: The visualization handbook, vol. 717. Elsevier München; 2005.

[48] Pant LM. Stochastic characterization and reconstruction of porous media [PhD thesis], University of Alberta; 2016.