**Proceedings of the ASME 2017 International Design Engineering Technical Conferences &
Computers and Information in Engineering Conference
IDETC/CIE 2017
August 6-9, 2017, Cleveland, United States**

DETC2017-68286

SCALABLE MICROSTRUCTURE RECONSTRUCTION WITH MULTI-SCALE PATTERN PRESERVATION

Ruijin Cang
Mechanical Engineering
Arizona State University
Tempe, Arizona, 85287
Email: cruijin@asu.edu

Aditya Vipradas
Mechanical Engineering
Arizona State University
Tempe, Arizona, 85287
Email: aviprada@asu.edu

Max Yi Ren*
Mechanical Engineering
Arizona State University
Tempe, Arizona, 85287
Email: yiren@asu.edu

ABSTRACT

A key challenge in computational material design is to optimize for particular material properties by searching in an often high-dimensional design space of microstructures. A tractable approach to this optimization task is to identify an encoder that maps from microstructures, which are 2D or 3D images, to a lower-dimensional feature space, and a decoder that generates new microstructures based on samples from the feature space. This two-way mapping has been achieved through feature learning, as common features often exist in microstructures from the same material system. Yet existing approaches limit the size of the generated images to that of the training samples, making it less applicable to designing microstructures at an arbitrary scale. This paper proposes a hybrid model that learns both common features and the spatial distributions of them. We show through various material systems that unlike existing reconstruction methods, our method can generate new microstructure samples of arbitrary sizes that are both visually and statistically close to the training samples while preserving local microstructure patterns.

1 Introduction

One common task in computational material design is to search for an optimal material processing configuration that leads to structures with desired properties. Practically, one could first identify structures that satisfy the target properties before searching for the processing configuration that meet the structural tar-

gets, as this decomposition could mitigate the computational burden of the joint inverse design of both structures and the process. This paper is motivated by one of the key challenges in the first step of the decomposition approach: It is often intractable to optimize for particular material properties with respect to a high-dimensional design space of microstructures, which are represented as 2D or 3D high-resolution images. One potential solution is to identify an encoder that maps microstructures to a lower-dimensional feature space, and a decoder that creates a microstructure based on given features. This two-way mapping allows the optimization to be performed in the feature space, solutions from where can be evaluated through the decoded microstructures.

Traditional approaches following this idea often use hand-picked features, e.g., volume fraction and filler geometric descriptors for composite materials [1], in which cases the encoder and decoder are manually programmed. This approach is not feasible when the microstructure under design has multiscale patterns that are hard to quantify by a human designer, nor would it be favorable when the correlation is not verified between hand-picked features (e.g., volume fraction) and the property of interest (e.g., fracture strength). Feature learning techniques have been applied to enable automated extraction of multiscale patterns from image samples, and to reconstruct new images using these patterns [2–5]. However, existing approaches require the generated images to have the same size as the training samples, and thus are not scalable for designing microstructures of an arbitrary size. Alternatively, Markovian Random Field (MRF) models have been applied to

---

*Address all correspondence to this author.

directly learn the pixel value distribution of images [6]. MRF is complementary to feature learning, in that it allows scalable reconstruction yet may produce undesired microstructures when the Markovian assumption fails to hold (e.g., when complicated local patterns exist, see Fig. 1 in [2] for example).

It is thus natural to propose a hybrid model that combines the advantages of feature learning and MRF. Specifically, we use a convolutional deep belief network (CDBN) [7] to extract local microstructure patterns, and then use a MRF model to learn the distributions of the *activation* of these patterns. We show that the proposed model can generate random microstructure samples of arbitrary sizes, that are statistically and visually similar to the training samples of complex material systems. Fig 1 presents the random generation of Ti-6Al-4V, Pb-Sn alloy, Sandstone, and Spherical Colloids, along with their training samples.

The paper is structured as follows: We review existing feature representations and reconstruction methods for material systems in Sec. 2. Sec. 3 provides specifications of the proposed model. Sec. 4 compares the proposed hybrid model with existing random field and deep belief network models on reconstruction performance using four material systems. We then discuss the advantages and limitations of the proposed algorithm in Sec. 5, followed by conclusions in Sec. 6.

2 Related work

We start by reviewing notable work related to this paper. These include manually defined feature representations, random field methods, and convolutional networks for automated feature learning. We also briefly review material synthesis methods. While synthesis methods produce random microstructures based on training data, they do not provide an encoder-decoder pair that maps the microstructure from and to a lower-dimensional feature space.

2.1 Feature representations for material systems

Existing studies have demonstrated the utility of low-dimensional features in enabling tractable computational material design [1]. Hand-picked features include physical (e.g. composition, dispersion [8–10], geometric [11–15]) and statistical ones (e.g., N-point correlation functions [9,11,12] and spectral density function [16]). The decoding of these features, i.e., the reconstruction process, is to find samples in the microstructure space that match the target descriptors. For example, when physical features are used, the reconstruction process could be to first match the target dispersion defined such as the average nearest cluster center distances, then the target geometry (e.g., distribution of equivalent radius, along the long axis and short axis of the particles), and lastly the volume fraction [17]. When statistical features are used, images can be optimized to match the target N-point correlation or cluster functions. Recently, Yu et al. introduced another statistical representation based on the spectral

density function for quasi-random nanostructured material systems [16]. They showed that such systems can be effectively represented by a low-dimensional feature vector in the spectrum space and the features also matches the bottom-up nanomanufacturing conditions.

2.2 Material microstructure as a random field

Another type of representation that we are particularly interested in is random field [18–20]. Bostanabad et al. [6] models each pixel of a microstructure image as a random variable conditioned on its neighbouring pixels. This implies that the microstructure is a MRF following its locality property, as the value of a pixel is unaffected by those lying outside of its neighborhood. The distribution of each random variable is also independent of its location in the image, further assuming that the MRF is stationary. The parameters of this distribution can be derived (encoded) through supervised learning by treating all pixels (from one or multiple images) as the output variables and their neighbourhoods as the inputs. Inversely, for given distribution parameters, i.e., the features, one can decode a distribution of images by a sequential sampling procedure.

2.3 Learning microstructure features through unsupervised feature learning

Feature learning techniques [7, 21–23] offer an automate two-way conversion between inputs and their lower-dimensional feature representations, and are thus useful in cases where hand-picking features is infeasible or undesirable. Here we focus on unsupervised feature learning methods that incorporate an encoder-decoder pair, for which parameters are tuned to achieve the minimal reconstruction error along with other objectives [4,5,7]. Existing methods along this line include Convolutional Deep Belief Network (CDBN [24]), Variational Autoencoder (VAE [22,25]) and Generative Adversarial Network (GAN [23,26]), all relying on convolutional network as a building block.

Convolutional networks have long been used for feature extraction, with broad applications to 2D and 3D object recognition [24, 27–29], detection [30–32, 32], generation [22, 23] and reinforcement learning [33–35]. Stacking multiple convolutional layers enables the network to learn multiscale features from the training data so that the network can explain the outputs. CDBN is one type of network that incorporates stacked convolutional layers, and allows a two-way mapping between inputs and outputs. Our previous work proposed a 5-layer CDBN [36] specifically designed to simultaneously perform dimension reduction of complex microstructures and microstructure reconstruction based the extracted features. See Fig. 2 for a comparison between random generations from CDBN and MRF on Ti-6Al-4V alloy. We note that reconstructions from CDBN, and other feature learning methods, are mostly of a fixed size, which limits their application to designing microstructures of arbitrary sizes.

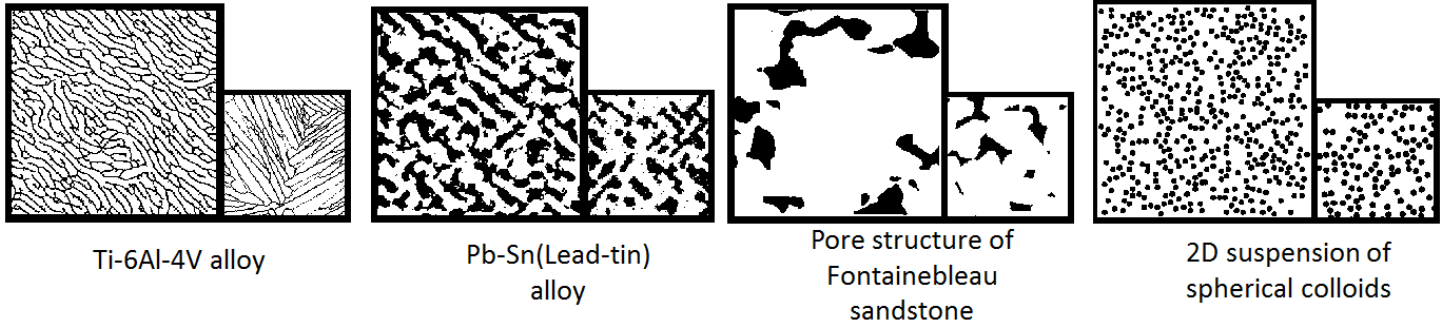VAE, as an extension from autoencoder [21], finds a encoder-

FIGURE 1. Our hybrid method combines advantages from feature learning and Markovian random field to enable scalable random reconstruction with local-pattern preservation. Four material systems are tested: the larger images are random reconstructions by learning from their corresponding samples shown on the right (multiple samples are used for learning, only one is shown).

TABLE 1. Summary of encoding and decoding techniques for microstructure reconstruction

Reconstruction method	Hand-made features	Random fields	Feature learning
Encoding method	Hard-coded feature calculation	Supervised learning	Unsupervised learning
Encoding cost	Low-Mid	High	High
Decoding loss	MSE	Joint distribution	N/A
Decoding method	Loss minimization	Gibbs sampling in image space	Feedforward net w/ independent sampling
Decoding cost	High	Mid	Low
Features	Physical, geometric, or statistical	Distribution parameters	High-level abstraction of images
Interpretability	Good	Poor/unknown	Poor/unknown

decoder pair by minimizing the reconstruction error and a KL-divergence between the data and a target distribution in the feature space. VAE is particularly useful when one needs to find interpretable low-dimensional features embedded in samples, yet often produces blurry outputs through its decoder [23, 37]. GAN, on the other hand, is known for having better quality in generation. A GAN consists of two networks: a decoder (generator) that maps the feature vector to images, and a discriminator that detects whether an input image is authentic (i.e., from the true sample set) or artificial (i.e., from the generator).

2.4 Summary of encoding and decoding techniques in existing microstructure reconstruction methods

In the first two methods, the definition of features entails a loss function, for which the decoding process attempts to minimize. For hand-made features, the loss is often in the form of a mean square error (MSE) between the generated and the target descriptors, and an optimizer is directly applied. The potentially high dimension of the image space imposes a computational burden on the decoding. For random fields, the loss is defined by the joint distribution of all pixel values, and an image with rela-
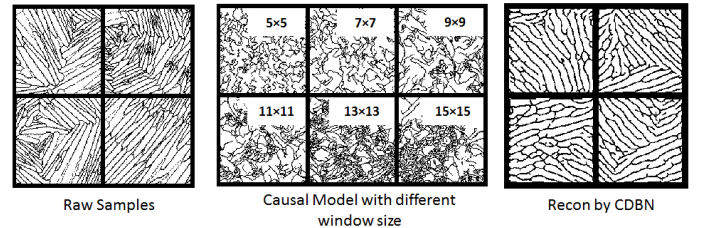

FIGURE 2. Left: Authentic samples of Ti64; Middle: Reconstructions using a MRF model [6] with six window sizes, marked in the images; Right: Reconstructions via CDBN: Better visual quality than MRF but fixed size.

tively low loss is generated through Gibbs sampling. In the case of [6, 38], it is shown that a single iteration of Gibbs sampling will be sufficient at generating visually plausible images for relatively simple microstructures.

The decoders of feature learning methods do not minimize a loss. Instead, the network is trained to output only plausible images with high probability. The variance in the output image is modeled by either randomness in the input (e.g., VAE) or at

3

each network layer (e.g., CDBN). Decoding in this case could be significantly faster than that of random fields, as the calculation can be parallelized thanks to the independence of all random variables involved.

The encoding procedure of these methods are also fundamentally different: For hand-made features, feature-specific encoders are hard coded; for random fields, supervised learning procedures can be adopted; for feature learning, the encoders are jointly learned along with the decoders, in order to minimize task-specific objectives. These differences in encoding and decoding techniques are summarized in Table 1.

2.5 Material synthesis

We note that a similar approach has been used in texture synthesis [39]: instead of using a supervised statistical model, one searches from sample image patches a closest match to the current neighborhood pixels, and assigns the corresponding pixel value to the current location [39]. Lubbers et al [40]. recently demonstrated the use of style transfer [41] on material systems. They also showed that learning a low-dimensional manifold from the texture descriptors could have potential at inversely estimating the processing settings that yield the microstructures. Nonetheless, the method does not include a decoder for generating random microstructures from the manifold.

3 Learning Distributions of Local Microstructure Patterns

Following [36], the proposed CDBN is composed by three stacked CRBM layers, each followed by a probabilistic max-pooling layer that reduces the height and the width of the output by a given factor. The outputs from the last pooling layer is modeled to follow a random field, so that new microstructures can be generated by first sampling an output layer of the CDBN using the random field, and then performing convolutions backward through the CDBN. See Fig. 3.

3.1 Learning local microstructure patterns

We briefly introduce key elements of the CDBN and provide details of the proposed CDBN settings.

Restricted Boltzmann Machine (RBM)   A RBM is a bipartite graphical model with visible ($\mathbf{v}$) and hidden ($\mathbf{h}$) nodes. The model is parameterized by the weight matrix $\mathbf{W}$ for its edges and biases $\mathbf{a}$ and $\mathbf{b}$ for visible and hidden nodes, respectively. The joint probability of $\mathbf{v}$ and $\mathbf{h}$ is modeled as

$$p(\mathbf{v}, \mathbf{h}) = \frac{\exp\left(-E(\mathbf{v}, \mathbf{h}|\boldsymbol{\theta})\right)}{\sum_{\mathbf{v},\mathbf{h}} \exp\left(-E(\mathbf{v}, \mathbf{h}|\boldsymbol{\theta})\right)}, \tag{1}$$

where

$$E(\mathbf{v}, \mathbf{h}|\boldsymbol{\theta}) = -\sum_{i=1}^{n} a_i v_i - \sum_{j=1}^{m} b_j h_j - \sum_{i=1}^{n} \sum_{j=1}^{m} v_i W_{ij} h_j \tag{2}$$

is the Boltzmann energy, and $\boldsymbol{\theta} := \{\mathbf{W}, \mathbf{a}, \mathbf{b}\}$ is the parameter set. Given a set of inputs to the visible nodes (e.g., microstructure images), the network weights and biases can be learned by maximizing the marginal probability

$$p(\mathbf{v}) = \frac{\sum_{\mathbf{h}} \exp\left(-E(\mathbf{v}, \mathbf{h}|\boldsymbol{\theta})\right)}{\sum_{\mathbf{v},\mathbf{h}} \exp\left(-E(\mathbf{v}, \mathbf{h}|\boldsymbol{\theta})\right)}. \tag{3}$$

This optimization problem can be solved by stochastic gradient descent [42] and contrastive divergence [43]. In a *convolutional* RBM, edges share weights in such a way that the encoding ($\mathbf{W}\mathbf{v} + \mathbf{b}$) and the decoding ($\mathbf{W}\mathbf{h} + \mathbf{a}$) are convolution operations.

Probabilistic max-pooling   A probabilistic max-pooling layer [7] is attached to each CRBM to reduce its size. This allows faster training and extraction of features of larger scales for the follow-up CRBMs, and builds into these follow-up filters a degree of invariance to translations and distortions. In forward pooling, hidden channels are divided into smaller blocks, each corresponding to a node in the pooling layer. For max-pooling with factor α, the block is of size $\alpha \times \alpha$. The states of the pooling nodes are drawn from Bernoulli distributions with parameters defined by the states of the corresponding block of hidden nodes. For backward pooling, all nodes in the block are assigned the same state as the corresponding pooling node.

3.2 A Markov Random Field model for pattern activations

We introduce an existing MRF model [39,44], and its application to the output layer of the CDBN, instead of direct application to the microstructure image. While we focus on 2D images, the extension of this model to 3D images can be achieve with minor modification [6].

Model overview   The model assumes that each pixel in a 2D image follows a stationary MRF [38], i.e., a pixel value Y_{ij} at location (i, j) follows a distribution $p(Y_{ij}|X_{ij}; \phi)$ that is conditioned on its neighboring pixels X_{ij} and parameterized by ϕ. When Y_{ij} is binary, the distribution is Bernoulli. The parameters ϕ can be learned through a set of (Y_{ij}, X_{ij}) pairs sampled from given images through supervised learning. Therefore, this model offers both an encoder and a decoder: An image (or a set of images) can be mapped to a single point ϕ in the feature space; inversely, one can generate images by maximizing the log likelihood $\sum_{ij} \log\left(p(Y_{ij}|X_{ij}; \phi)\right)$ with respect to the image, for a given ϕ. A greedy solution to this inverse problem is stated in [38]: Starting from an empty image, we draw Y_{ij} sequentially from the top-left to the bottom-right corner, using only those neighbouring pixels that have been drawn. See Fig. 4 for an illustration.

It is important to note that unlike CDBN, the encoder-decoder pair in this case does not preserve specific images, i.e., the similarity between images is measured only in the feature space. From
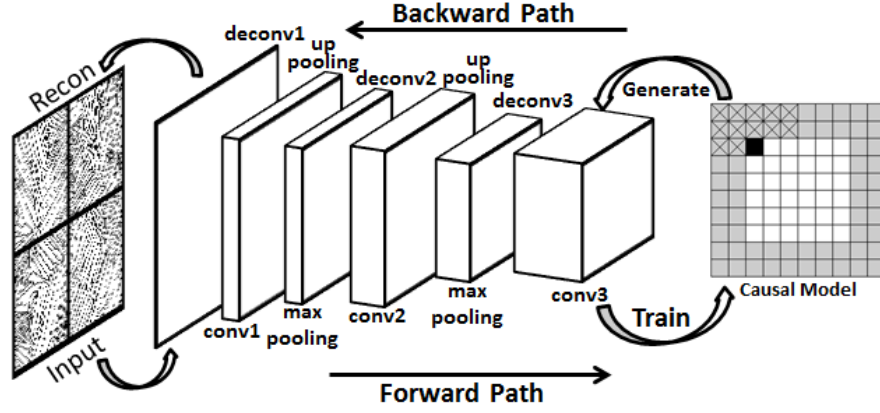
FIGURE 3. The proposed network with three convolutional RBM layers followed by a random field model(causal model). Image generation is done by sampling from the random field and performing backward convolution through the CDBN.
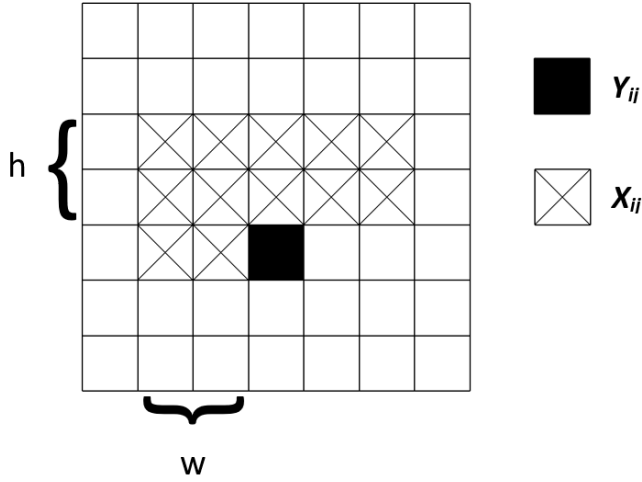

FIGURE 4. The pixel Y_{ij} is conditioned on its $(h(2w+1)+w)$ neighboring pixels X_{ij}
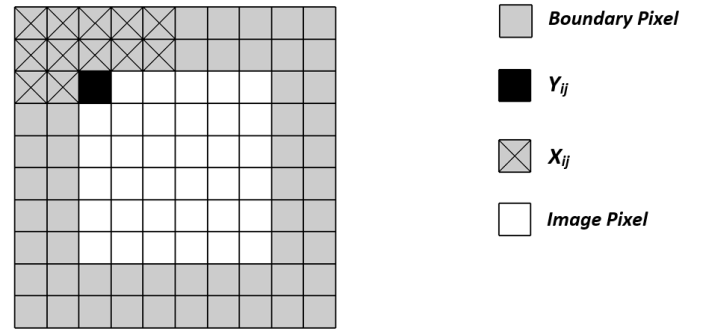
whichever is dominating in the sample images.

FIGURE 5. The padding is shown in grey. Y_{ij} on the boundary depends on X_{ij} in the padding. Padding thickness is $(2h+1)$.

a material design perspective, this model is useful for material systems where the statistics (such as ϕ) rather than the image representations is sufficient and effective at explaining the variances in properties and processing parameters. We will elaborate on this point in Sec. 5.

Setting boundary padding   One may notice that there is a need to provide boundary padding to images for both encoding and decoding. This is because the MRF model requires each pixel to be dependent on its neighbors, including the ones on the boundary. Boundary padding is thus created around the image to provides sufficient preceding pixels to draw Y_{ij}s on the boundary. See Fig. 5 for an illustration. The padding is set to black or white,

Learning the conditional probability model   We learn K conditional probability models, where K is the number of channels of the CDBN output layer of a particular material system. To do so, we first randomly sample $(Y_{ij}^{(k)}, X_{ij}^{(k)})$ pairs from each channel across all sample images, where $X_{ij}^{(k)}$ are preceding pixels in the neighbour of $Y_{ij}^{(k)}$. As will be explained in Sec. 4.1, the best neighbouring window size $(h+w+1)$ varies with material systems. Note that in addition to the neighboring pixels, we also include in $X_{ij}^{(k)}$ an independent variable $\gamma_{ij}^{(k)}$ that represents the percentage of activated nodes from the preceding channels at location (i, j): $\gamma_{ij}^{(k)} = \sum_{k'=1}^{k-1} Y_{ij}^{(k')}/(k-1)$. For the first channel, this variable is omitted. This treatment prevents overly activating one location across channels by accident. The prepared data is

5

then used to train K models using a classification tree algorithm (*fitctree*) in MATLAB. The tree depth is $n-1$ where n is the training sample size. The minimal leaf size is 1, and the minimal parent size is 10.

Reconstruction with a consistent volume fraction   Once trained, the classifiers output the probabilities $p_{ij}^{(k)} := p(Y_{ij}^{(k)} = 1|X_{ij}^{(k)}; \phi)$ for given $X_{ij}^{(k)}$. Each channel is generated using the corresponding classifier through the method described above. To ensure volume fraction conservation in each channel, we use the iterative procedure described in [38]: In each iteration, we calculate the volume fraction gap (δ) between the current reconstruction and the training images. If $\delta > 0.0025$, we adjust the conditional probability as $p_{ij}'^{(k)} = p_{ij}^{(k)} + c\sqrt{p_{ij}^{(k)}(1-p_{ij}^{(k)})}$, where the algorithmic parameter c is updated as $c = c + 0.005\text{sign}(\delta)$, initialized at $c = 0$. This procedure reduces the probability of activating a node when the reconstructed volume fraction is higher than that of the samples, and vice versa.
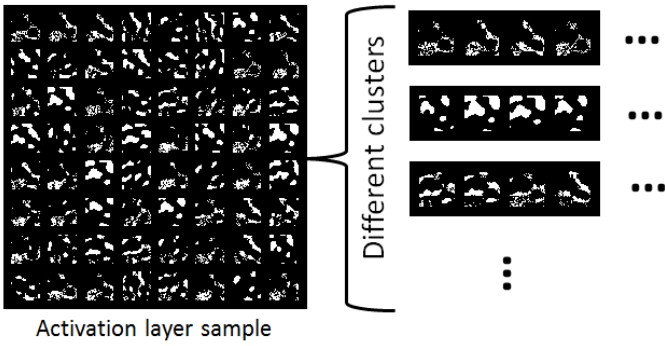

FIGURE 6.   Sample activation layer from sandstone. Clusters are also observed in other material systems.

Addressing correlations among channels   We noticed that the activations across channels in the CDBN output layer are correlated. And the same correlation occurs across images. See Fig. 6 for example. The proposed random field model, however, does not capture this correlation, as the classification trees are trained almost independently except for the consideration of γ. To enable correlations, we propose to identify the clustering of channels based on their activation patterns from the training samples. Specifically, we first calculate the channel-wise mean activation across all samples, although the activation pattern is very different for every image, the same correlation occurs across all the training samples, in this case K-means [45] is applied to cluster these mean activation. The resultant clusters will be applied during the reconstruction, where we use the same patterns

for channels within a cluster. The appropriate number of clusters is determined by calculating the reconstruction error, i.e., image-wise difference, between the original samples and the reconstructed ones with shared patterns. When calculating this difference, the shared pattern for each cluster is chosen randomly from one of its channels.

3.3   Post-processing

Post-processing is needed to improve the quality of reconstructions for Ti64. As discussed in [2], the third hidden layer of the CDBN can be overly saturated, which leads to regions in the input layer with overlapped grain boundary or undesirable voids. The same issue is observed in the hybrid model. Our current solution is to set an upper bound on the activation rate of all channels. Fig. 7 shows the activations of the CDBN output layer above and their corresponding reconstructions below. One can see that directly applying the generated activation will lead to less desirable outcomes. Experiments show that reducing the activation rate to 15% will sufficiently improve the reconstruction quality, because by reducing the activation rate cross all the channels avoids the over-activation issue Fig 10. Finally, skeletonization is used to further match the style of the generated images to the original ones.
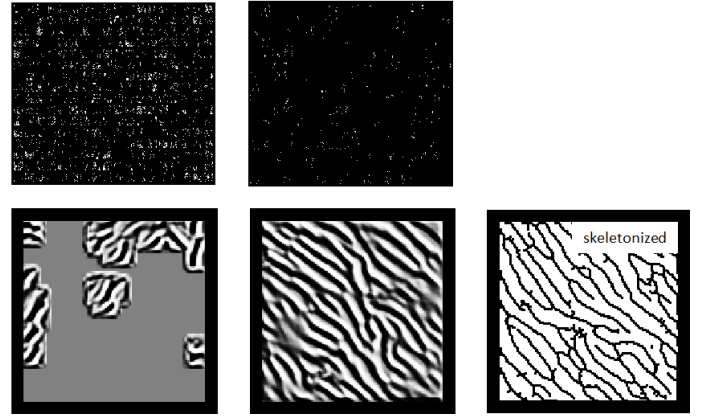

FIGURE 7.   (Top) CDBN output activations generated by MRF for Ti64; (Bottom) The corresponding generated images; (Left) No activation reduction after sampling; (Middle) 85% reduction on the total activation; (Right) Result after skeletonization

4   Verification of the Proposed Method

This section elaborates on the algorithmic settings of the proposed hybrid model for four material systems: Ti-6Al-4V alloy (Ti64), Pb-Sn alloy (Pb-Sn), sandstone (SS), and spherical

6

colloids (SC). Due to differences in their local patterns and distributions of these patterns, the settings vary across these systems. We then compare the generation quality between the hybrid and the MRF models. Our data consists of 100 samples for Ti64, 60 samples for Pb-Sn and SS, and 80 samples for SC. All sample images used for model training are listed in the Appendix. We also show generations of different image sizes using the proposed method.

4.1 Algorithmic settings

Settings of the MRF model   We apply the MRF model directly to microstructure images to produce benchmark images. Due to the sensitivity of the MRF approach to the window size (as illustrated in Fig 8), we determine the optimal size for each material system via cross-validation (CV) by starting with a relatively large window size and reducing the size till the CV error of the model does not further decrease. The CV errors are measured by the differences between the two-point correlation functions of the generated images and those of the original samples. The optimal sizes are thus derived as 7×7, 17×17, 19×19, 15×15, for Ti64, Pb-Sn, SS and SC, respectively. The minimum CV errors are 7.7% for Ti64, 4% for Pb-Sn, 0.95% for SS, and 2.2% for SC. This data indicates that Ti64 follows the Markov assumption of MRF less strictly than the other material systems.

Settings of the hybrid model   The configurations of the CRBM layers are tuned according to each material in order to achieve good generation quality, as different material systems may have patterns of different length scales. All models have three CRBM layers followed by probabilistic max-pooling layers, except for SC due to its simple spherical feature, in which case a single CRBM layer is used. For Ti64, Pb-Sn, and SS, the convolutional filter sizes are 6×6, 9×9, and 9×9 for the first, second, and third CRBM layers, respectively; the number of filters for the first and second layers are 24 and 40 respectively; The number of filters for the third layer varies according to the microstructure complexity. Experiments showed that 288, 64, and 64 filters will be sufficient for Ti64, Pb-Sn and SS, respectively. We notice that using these reduced amount of filters for Pb-Sn and SS leads to 6% of average pixel-wise reconstruction error, compared with those with 288 filters. Max-pooling layers with a factor of $\alpha = 2$ is applied to Ti64, Pb-Sn, and SS, and $\alpha = 3$ for SC.

The MRF model is applied to the last activation layers. Similar to the direct MRF models, the window size for training the models is chosen through CV. For all materials, the optimal window size is found to be 5×5.

Target sparsity for the hidden nodes are set during the training of the CRBM network in order to prevent learning trivial features. This ensures that, during the training, a filter can only activate a limited number of hidden nodes and thus represent a specific image pattern. Table 2 summarizes the model and training parameters.

The number of clusters for the MRF applied to the CDBN output is tuned for each material system. It is identified that 8 clusters work well for Pb-Sn and SS. Fig. 9 compares the reconstructions from 4 to 64 clusters. Pb-Sn is shown as an example. Note that the reconstruction error decreases with more clusters. No cluster is needed for SC since only one channel exists. Ti64 material system is not suitable for clustering as doing so leads to undesirable reconstructions, as is demonstrated in the top row of Fig. 10. By checking the summation of all the activation channels (the bottom row of Fig. 10), we notice that the regions with high activations, due to shared patterns within clusters, lead to voids in the reconstructions.

Settings of the similarity measure   A similarity measure is needed for evaluating the quality of reconstructions and random generations. In this paper, we measure the similarities between sample images and reconstructions both visually and through 2-point correlation functions (S_2), which is a very effective choice of measurement for homogeneous and isotropic materials [46, 47], such as Pb-Sn, SS and SC. While commonly agreed similarity measure for complicated material systems such as Ti64 is currently missing, the same measure is applied for consistency. To calculate 2-point correlation functions for biphase microstructures, we sample all the possible pairs of points (x_1, x_2) for distances from 1 to half of the length of the image. At each distance, the percentage of pairs with matched pixel values is calculated.

4.2 Comparisons on random generations

Fig. 11 compares the performance between MRF, CDBN [2] and the proposed hybrid method, using the four material systems. For each system, we present the original microstructures, and random generations by MRF (with the best window size), hybrid and CDBN, from top to bottom. 2-point correlation functions are calculated for 30 random generations from each method, in addition to the original samples.[1] Mean and standard deviations (as the shades) are reported.

Major findings are as follows: (1) The proposed method produces reconstructions that are visually similar to the samples from CDBN, while MRF has less favorable reconstructions particularly on Ti64 due to the existence of local patterns which breaks the Markovian assumption. MRF also produces more white areas in Pb-Sn in comparison with the hybrid and the CDBN methods. In addition, a closer investigation shows that MRF fails to constrain the filler particles in SC to be spherical. Nonetheless, MRF produces local patterns closer to the samples in SS. We believe that this is due to the drawback of the current feature learning implementation that is less sensitive to capturing sharp edges of local patterns. (2) None of the reconstructions for Pb-Sn, SS, and SC are statistically different from the original samples with

---

[1]The number of random generations is limited by the high cost of sampling the large number (288) of channels for the hybrid method for Ti64.

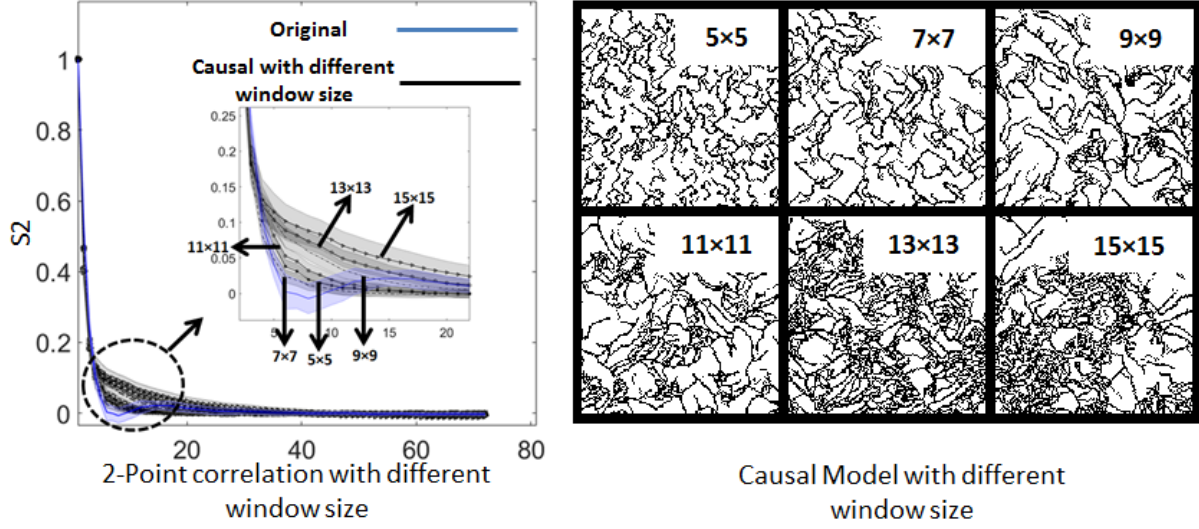**FIGURE 8.** (Left) 2-point correlation functions with window sizes: 5×5, 7×7, 9×9, 11×11, 13×13 and 15×15 (Right) Corresponding generations

Material system	1st layer				2nd layer				3rd layer			
	#filter	filter size	sparsity	pλ	#filter	filter size	Sparsity	pλ	#filter	filter size	sparsity	pλ
Ti-6Al-4V	24	6	0.1	10	40	9	0.1	10	288	9	0.1	10
Pb-Sn alloy2	24	6	0.1	10	40	9	0.1	10	64	9	0.1	10
Sandstone	24	6	0.1	10	40	9	0.1	10	64	9	0.1	10
Spherical Colloid	1	12	0.1	10	NA				NA			

TABLE 2. CDBN settings
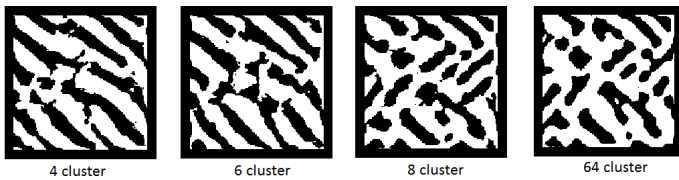
FIGURE 9. Different cluster number on the 64 channels' activation layer based reconstruction for Pb-Sn alloy is shown, as the cluster number increases the final reconstruction gets close.

respect to S_2. For Ti64 however, the findings are intriguing: While MRF produces closer S_2 to the samples, these functions do not capture the critical characteristics of this material system, as they monotonically decreases along the distance. To elaborate, S_2 for Ti64 should exhibit clear oscillations for small distance

values, to reflect significant short range correlations due to the mutual exclusion volume between the grains. The wavelengths associated with the oscillations also indicate the average grain width. The characteristics are preserved in reconstructions from the hybrid and the CDBN methods, while their S_2s have higher amplitudes in the oscillation than those from the samples. This is due to the more uniform grain sizes and shape distributions in the reconstructions and the missing of local details exist in the samples.

4.3 Generation at an arbitrary size

With the help of the MRF model on the CDBN output, we can produce reconstructions with an arbitrary size. Fig. 12 shows samples for larger reconstructions: 288×288 for Ti64, Pb-Sn and SS, 378×378 for SC. Their 2-point correlation functions are also calculated and compared with those of the original samples, and generations of the original sizes (144×144 for Ti64, Pb-Sn and
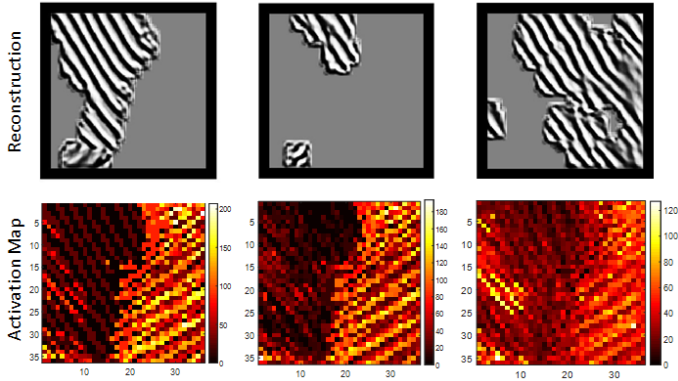
FIGURE 10. Clustering the the activation channels for Ti64 into three different groups, 8, 16, 36, from left to right. None of them produce a good reconstruction as shown in the upper row. The lower row presents the summation of their corresponding activation channels.

SS, and 189×189 for SC). The result shows that characteristics of the 2-point correlation functions are preserved across generation sizes.

5 Discussion

Limitations and future directions are summarized as follows.

Reconstruction quality We note that the proposed model is yet to achieve high accuracy in reconstruction for complex material systems, especially when detailed features (e.g., the connection of grain boundaries in Ti64) are critical for material properties, even when post-processing techniques are manually applied. The application of post-processing also introduces additional human knowledge into the reconstruction procedure, making the method not fully automated for all material systems. The lack of reconstruction accuracy is largely due to the blurring effect of the convolutions. We also note that neither VAE nor GAN in their original formulations [22, 23] produces high quality random generations for complex material systems such as Ti64, as shown in Fig. 13, although the networks can reconstruct the training samples well. Note that these generations are too blurred to skeletonize. VAE suffers from its limitation of only allowing continuous features, i.e., a real-valued weighting on various discrete local patterns does not produce meaningful patterns. The GAN model may suffer from its intrinsic difficulty in balancing the convergence of the generator and the descriminator.

The Markovian assumption It is also important to note that MRF relies on the strong assumption that the image to be generated follows a stationary random field, and each pixel value is only conditioned on its limited neighbors. In this paper, we used CDBN to remove local patterns, as an attempt to acquire

the resulting activation images these MRF properties. This attempt will fail if a material system, after abstracted by CDBN, cannot be considered as a MRF. CDBNs used in this paper also requires careful tuning, based on the designer's understanding of the representative length scales of local patterns. Compared with conventional procedures for hand-picking features, our approach avoids asking for specifying details of human knowledge during the construction of the decoder, yet still requires heuristic inputs from the human designer.

Optimal material design in a statistical sense As we mentioned earlier, the application of MRF allows a distribution of microstructures to be represented by the same feature vector, i.e., the MRF parameters. This model is useful when the variances in the properties of interest can be effectively explained by these parameters. This is, however, not always the case. For example, as is shown in [36], the fracture strength of Ti64 samples from the same processing setting is sensitive to their microstructure details. Thus it is not meaningful to represent a distribution of designs with the same feature when the property is to be optimized with respect to a particular microstructure. Nonetheless, like other material properties, the fracture strength of Ti64 exhibits correlation with laser sintering parameters, as can be seen in Fig. 14. Therefore, using MRF parameters (with the help of CDBN) to represent microstructures is meaningful when the goal is to identify a distribution of microstructures from the same processing setting that has the optimal mean property. Our future work will include microstructure samples from multiple processing settings in training the hybrid model. We will examine if the resulting MRF parameters are effective at explaining the properties and the processing parameters.

6 Conclusions

In this paper we developed a method to enable random generation of 2D material microstructure images of an arbitrary size that are visually and statistically similar to the authentic ones, based on a given set of samples. This is achieved by learning multiscale patterns through a convolutional deep belief network, and modeling the distributions of these patterns through a Markov Random Field. Technical challenges in integration these two models are discussed in detail. Model parameters are tuned by matching 2-point correlation functions to those of the training samples. We show that the proposed hybrid method performs particularly well on Ti-6Al-4V alloy both visually and statistically, in comparison with an existing MRF model.
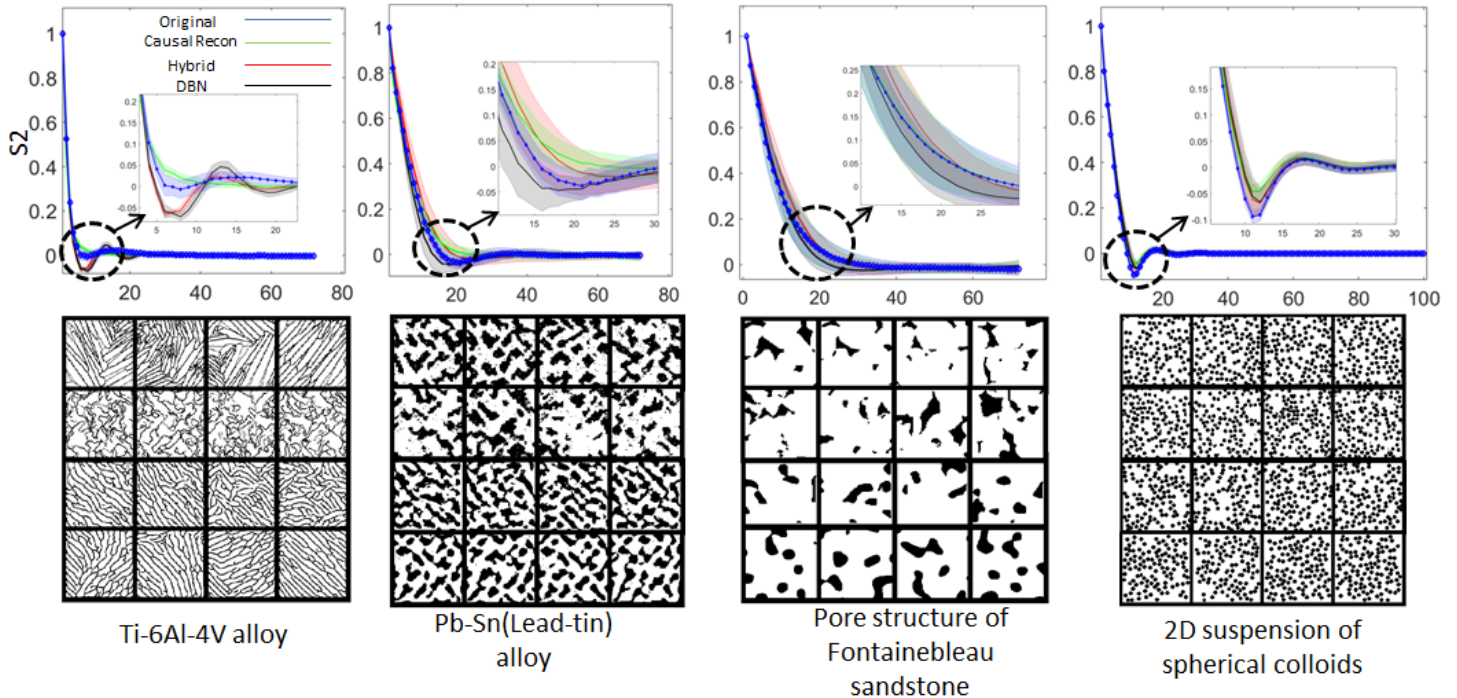
FIGURE 11. (Top) Comparison of the sample means and sample standard deviations of 2-point correlation functions of: the original samples, reconstructions from MRF with the best window sizes, the hybrid method, and CDBN (Bottom) The corresponding reconstructions (from top to bottom: original, MRF, hybrid, CDBN)

REFERENCES

[1] Xu, H., Dikin, D. A., Burkhart, C., and Chen, W., 2014. "Descriptor-based methodology for statistical characterization and 3d reconstruction of microstructural materials". *Computational Materials Science, 85*, pp. 206–216.

[2] Cang, R., and Ren, M. Y., 2016. "Deep network-based feature extraction and reconstruction of complex material microstructures". In ASME 2016 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, American Society of Mechanical Engineers, pp. V02BT03A008–V02BT03A008.

[3] Coates, A., Lee, H., and Ng, A. Y., 2010. "An analysis of single-layer networks in unsupervised feature learning". *Ann Arbor, 1001*(48109), p. 2.

[4] Boureau, Y.-l., Cun, Y. L., et al., 2008. "Sparse feature learning for deep belief networks". In Advances in neural information processing systems, pp. 1185–1192.

[5] Vincent, P., Larochelle, H., Lajoie, I., Bengio, Y., and Manzagol, P.-A., 2010. "Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion". *Journal of Machine Learning Research, 11*(Dec), pp. 3371–3408.

[6] Bostanabad, R., Chen, W., and Apley, D. W., 2016. "Characterization and reconstruction of 3d stochastic microstructures via supervised learning". *Journal of Microscopy, 0*, pp. 1–16.

[7] Lee, H., Grosse, R., Ranganath, R., and Ng, A. Y., 2009. "Convolutional deep belief networks for scalable unsupervised learning of hierarchical representations". In Proceedings of the 26th Annual International Conference on Machine Learning, ACM, pp. 609–616.

[8] Rollett, A. D., Lee, S.-B., Campman, R., and Rohrer, G., 2007. "Three-dimensional characterization of microstructure by electron back-scatter diffraction". *Annu. Rev. Mater. Res., 37*, pp. 627–658.

[9] Borbely, A., Csikor, F., Zabler, S., Cloetens, P., and Biermann, H., 2004. "Three-dimensional characterization of the microstructure of a metal–matrix composite by holotomography". *Materials Science and Engineering: A, 367*(1), pp. 40–50.

[10] Tewari, A., and Gokhale, A., 2004. "Nearest-neighbor distances between particles of finite size in three-dimensional uniform random microstructures". *Materials Science and Engineering: A, 385*(1), pp. 332–341.

[11] Torquato, S., 2013. *Random heterogeneous materials: microstructure and macroscopic properties*, Vol. 16. Springer

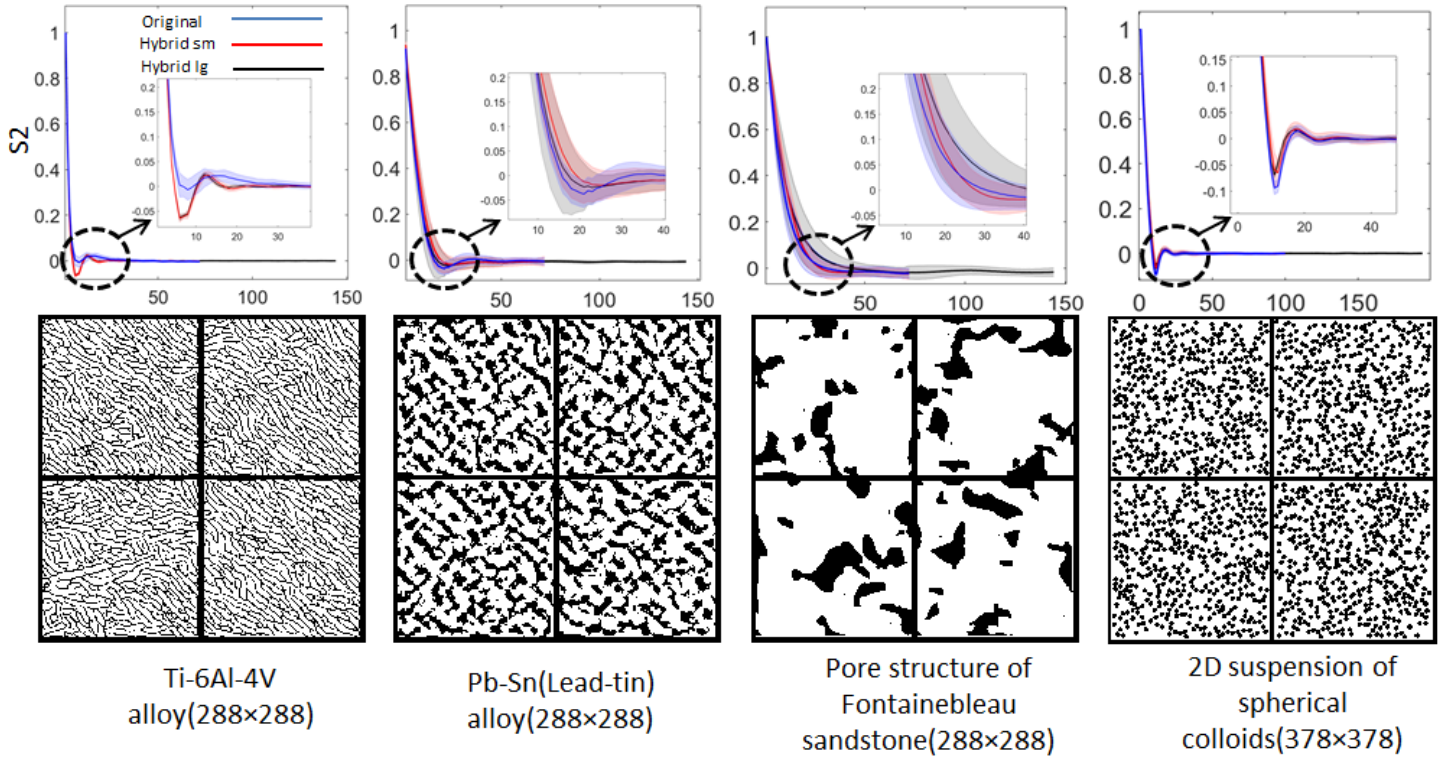**FIGURE 12.** Generations of size 288×288 for Ti64, Pb-Sn, and SS, 378×378 for SC, and their 2-point correlation functions.
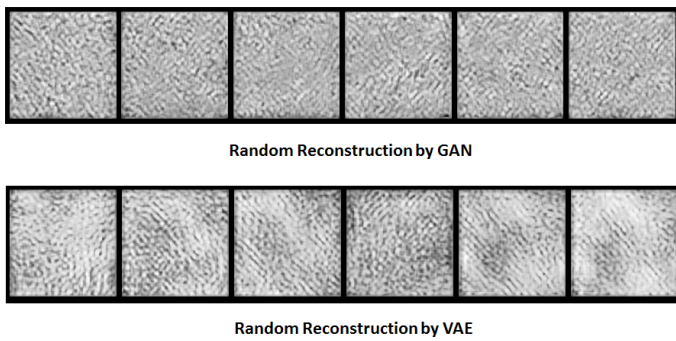

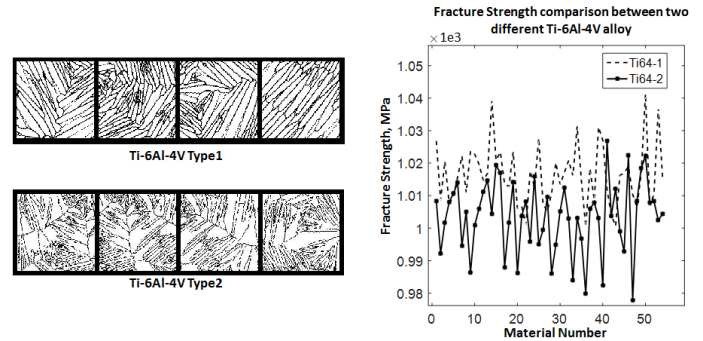**FIGURE 13.** Random generations from VAE and GAN.

FIGURE 14. (Right) Two types for Ti64 samples from different processing parameters (Left) Their corresponding fracture strengths

Science & Business Media.

[12] Sundararaghavan, V., and Zabaras, N., 2005. "Classification and reconstruction of three-dimensional microstructures using support vector machines". *Computational Materials Science, 32*(2), pp. 223–239.

[13] Basanta, D., Miodownik, M. A., Holm, E. A., and Bentley, P. J., 2005. "Using genetic algorithms to evolve three-dimensional microstructures from two-dimensional micrographs". *Metallurgical and Materials Transactions A, 36*(7), pp. 1643–1652.

[14] Klaysom, C., Moon, S.-H., Ladewig, B. P., Lu, G. M., and Wang, L., 2011. "The effects of aspect ratio of inorganic fillers on the structure and property of composite ion-exchange membranes". *Journal of colloid and interface science, 363*(2), pp. 431–439.

11

[15] Gruber, J., Rollett, A., and Rohrer, G., 2010. "Misorientation texture development during grain growth. Part II: Theory". *Acta Materialia, 58*(1), pp. 14–19.

[16] Yu, S., Zhang, Y., Wang, C., Lee, W.-k., Dong, B., Odom, T. W., Sun, C., and Chen, W., 2016. "Characterization and design of functional quasi-random nanostructured materials using spectral density function". In ASME 2016 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, American Society of Mechanical Engineers, pp. V02BT03A012–V02BT03A012.

[17] Xu, H., Li, Y., Brinson, L. C., and Chen, W., 2014. "A Descriptor-based Design Methodology for Developing Heterogeneous Microstructural Materials System". *Journal of Mechanical Design, 136*(c), pp. 1–25.

[18] Wei, L.-Y., and Levoy, M., 2000. "Fast texture synthesis using tree-structured vector quantization". In Proceedings of the 27th annual conference on Computer graphics and interactive techniques, ACM Press/Addison-Wesley Publishing Co., pp. 479–488.

[19] Efros, A. A., and Freeman, W. T., 2001. "Image quilting for texture synthesis and transfer". In Proceedings of the 28th annual conference on Computer graphics and interactive techniques, ACM, pp. 341–346.

[20] Efros, A. A., and Leung, T. K., 1999. "Texture synthesis by non-parametric sampling". In Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on, Vol. 2, IEEE, pp. 1033–1038.

[21] Zemel, R. S., 1994. "Autoencoders, minimum description length and helmholtz free energy". NIPS.

[22] Kingma, D. P., and Welling, M., 2013. "Auto-encoding variational Bayes". *arXiv preprint arXiv:1312.6114*.

[23] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., and Bengio, Y., 2014. "Generative adversarial nets". In Advances in neural information processing systems, pp. 2672–2680.

[24] Hinton, G., Deng, L., Yu, D., Dahl, G. E., Mohamed, A.-r., Jaitly, N., Senior, A., Vanhoucke, V., Nguyen, P., Sainath, T. N., et al., 2012. "Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups". *Signal Processing Magazine, IEEE, 29*(6), pp. 82–97.

[25] Rezende, D. J., Mohamed, S., and Wierstra, D., 2014. "Stochastic backpropagation and approximate inference in deep generative models". *arXiv preprint arXiv:1401.4082*.

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

[27] Krizhevsky, A., Sutskever, I., and Hinton, G. E., 2012. "ImageNet classification with deep convolutional neural networks". In Advances in neural information processing systems, pp. 1097–1105.

[28] Simonyan, K., and Zisserman, A., 2014. "Very deep convolutional networks for large-scale image recognition". *arXiv preprint arXiv:1409.1556*.

[29] He, K., Zhang, X., Ren, S., and Sun, J., 2016. "Deep residual learning for image recognition". In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 770–778.

[30] Krizhevsky, A., Sutskever, I., and Hinton, G. E., 2012. "ImageNet classification with deep convolutional neural networks". In *Advances in Neural Information Processing Systems 25*, F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, eds. Curran Associates, Inc., pp. 1097–1105.

[31] Girshick, R., Donahue, J., Darrell, T., and Malik, J., 2014. "Rich feature hierarchies for accurate object detection and semantic segmentation". In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 580–587.

[32] Ren, S., He, K., Girshick, R., and Sun, J., 2015. "Faster r-cnn: Towards real-time object detection with region proposal networks". In Advances in neural information processing systems, pp. 91–99.

[33] Mnih, V., Kavukcuoglu, K., Silver, D., Graves, A., Antonoglou, I., Wierstra, D., and Riedmiller, M., 2013. "Playing Atari with deep reinforcement learning". *arXiv preprint arXiv:1312.5602*.

[34] Schmidhuber, J., 2015. "Deep learning in neural networks: An overview". *Neural Networks, 61*, pp. 85–117.

[35] Levine, S., Finn, C., Darrell, T., and Abbeel, P., 2015. "End-to-end training of deep visuomotor policies". *arXiv preprint arXiv:1504.00702*.

[36] Cang, R., Xu, Y., Chen, S., Liu, Y., Jiao, Y., and Ren, M. Y., 2016. "Microstructure representation and reconstruction of heterogeneous materials via deep belief network for computational material design". *arXiv preprint arXiv:1612.07401*.

[37] Larsen, A. B. L., Sønderby, S. K., Larochelle, H., and Winther, O., 2015. "Autoencoding beyond pixels using a learned similarity metric". *arXiv preprint arXiv:1512.09300*.

[38] Bostanabad, R., Bui, A. T., Xie, W., Apley, D. W., and Chen, W., 2015. "Stochastic microstructure characterization and reconstruction via supervised learning". *Acta Materialia, 103*, pp. 89–102.

[39] Liu, X., and Vadim, S., 2015. "Random heterogeneous materials via texture synthesis". *Computational Materials Science, 99*, pp. 177–189.

[40] Lubbers, N., Lookman, T., and Barros, K., 2016. "Inferring low-dimensional microstructure representations using convolutional neural networks". *arXiv preprint arXiv:1611.02764*.

[41] Gatys, L. A., Ecker, A. S., and Bethge, M., 2016. "Image style transfer using convolutional neural networks". In Pro-

ceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2414–2423.

[42] Bottou, L., 2010. "Large-scale machine learning with stochastic gradient descent". In *Proceedings of COMPSTAT'2010*. Springer, pp. 177–186.

[43] Hinton, G. E., 2002. "Training products of experts by minimizing contrastive divergence". *Neural computation, 14*(8), pp. 1771–1800.

[44] Sundararaghavan, V., 2014. "Reconstruction of three-dimensional anisotropic microstructures from two-dimensional micrographs imaged on orthogonal planes". *Integrating Materials and Manufacturing Innovation, 3(1)*, p. 19.

[45] MacQueen, J., et al., 1967. "Some methods for classification and analysis of multivariate observations". In Proceedings of the fifth Berkeley symposium on mathematical statistics and probability, Vol. 1, Oakland, CA, USA., pp. 281–297.

[46] Liu, Y., Greene, M. S., Chen, W., Dikin, D. A., and Liu, W. K., 2013. "Computational microstructure characterization and reconstruction for stochastic multiscale material design". *Computer-Aided Design, 45*(1), pp. 65–76.

[47] Jiao, Y., Stillinger, F., and Torquato, S., 2008. "Modeling heterogeneous materials via two-point correlation functions. ii. algorithmic details and applications". *Physical Review E, 77*(3), p. 031135.
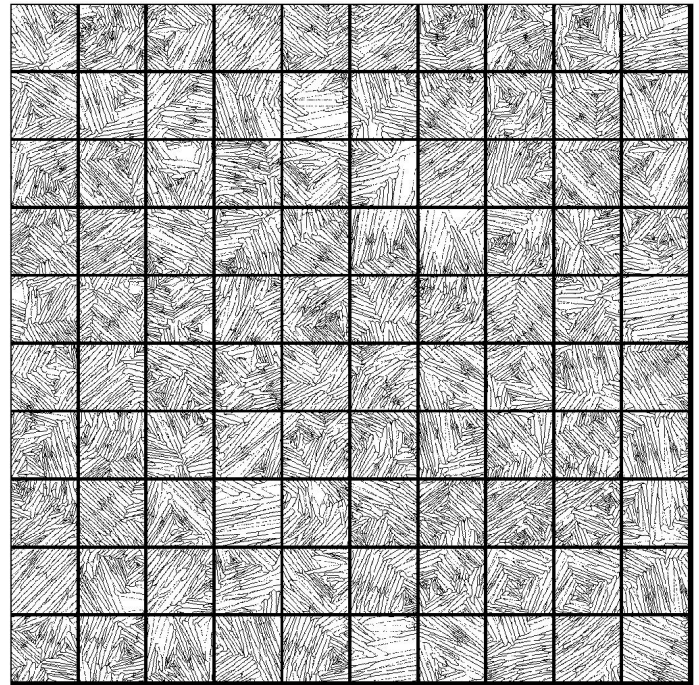
Appendix: Sample Microstructure Images

The following figures are reproduced from [2].

FIGURE 15. Ti-6Al-4V alloy

FIGURE 16. Pb-Sn (lead-tin) alloy

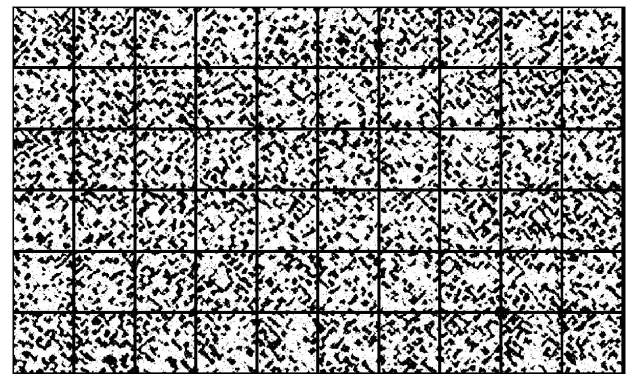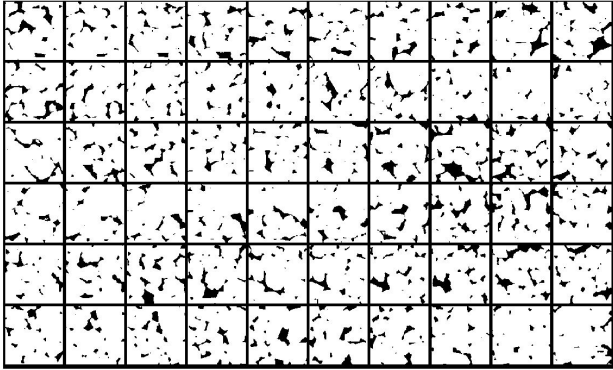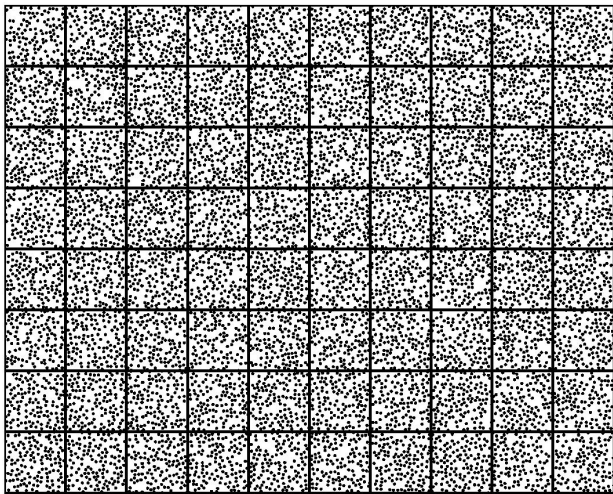**FIGURE 17**. Pore structure of Fontainebleau sandstone

FIGURE 18. 2D suspension of spherical colloids