Ultra-sparse View X-ray Computed Tomography for 4-D Imaging

Yanjie Zheng[†] and Kelsey B. Hatzell[*,‡]

†*Department of Mechanical and Aerospace Engineering Department, Princeton University, Princeton, NJ 08540, USA*

‡*Mechanical and Aerospace Engineering Department, Princeton University, Princeton, NJ 08540, USA*

¶*Andlinger Center for Energy and the Environment, Princeton, NJ 08540, USA*

E-mail: kelsey.hatzell@princeton.edu

Abstract

X-ray Computed Tomography (CT) is a non-invasive, non-destructive approach to imaging materials, material systems and engineered components in two- and three-dimensions. Acquisition of 3D images requires the collection of hundreds or thousands of through-thickness X-ray radiographic images from different angles. Such 3D data acquisition strategies commonly involve sub-optimal temporal sampling for *in situ* and operando studies (4D imaging). Herein, we introduce a sparse-imaging approach, Tomo-NeRF, which is capable of reconstructing high-fidelity 3D images from <10 two-dimensional radiographic images. Experimental 2D and 3D X-ray images were used to test the reconstruction capability in two-view, four-view, and six-view scenarios. Tomo-NeRF is capable of reconstructing 3D images with a structural similarity of 0.9971-0.9975 and voxel-wide accuracy of 81.83–89.59% from 2-D experimentally obtained images. The reconstruction accuracy for the experimentally obtained images

is less than the synthetic structures which demonstrated a similarity of 0.9973-0.9984 and voxel-wise accuracy of 84.31-95.77%.

Introduction

X-ray computed Tomography (XCT) tools are non-invasive imaging tools which can capture sub-surface morphological and structural features in a range of materials and material systems.[1] The fidelity of microstructure representation is essential for material property estimation, simulation, and materials design.[2] Three-dimensional reconstructions of materials are often utilized to accurately model components and critically assess degradation and material transformation pathways.[3] The resiliency of any quantitative assessment is highly dependent on the reconstruction of a 3D image from hundreds of 2D radiographic images acquired from different projection angles[4].[5] The inverse Radon transform provides the mathematical basis for the conventional reconstruction process.[6] Thousands of 2-D images at varying angles are typically acquired in order to reconstruct images into a 3D rendering. For dynamic or *operando* investigations, where you combine an action with imaging (e.g. imaging particles flowing), it is challenging to take thousands of images in the time period of the action. Thus, 3D imaging which requires fast acquisition times is fundamentally limited.[7] Sparse view reconstruction approaches attempt to reduce the number of 2D images (e.g. sampling rate) necessary for 3D reconstruction. Decreasing the sampling rate enables transient *in situ* and operando studies. Reducing the number of images needed for a reconstruction can have the added benefit of less instrument maintenance, longer system lifetime, and simpler system design and hardware.[8] Despite the promising features, sparse view reconstruction via conventional reconstruction methods suffers from the inevitable information loss that hinders correct image interpretation and quantification.

Image reconstruction in computed tomography has undergone a paradigm shift as a result of recent advancements in deep learning for computer vision applications.[9] The two

primary methodologies of deep sparse view reconstruction are image quality enhancement [8] and stochastic reconstruction. [2] The former adapts deep learning networks to reduce noise and strike artifacts from the sparse view reconstruction. The networks in the data domain interpolate missing sinogram data (i.e., Radon transform data) of sparse view reconstruction by learning from complete projection sinograms. [3,10–14] The networks in the image domain restore the image quality by learning to detect and subtract streak-type artifacts from the undersampled reconstruction images. [15,16] Either by learning from data or images, these deep learning models incorporate analytical transform (e.g., Radon transform) as physical-based knowledge for image reconstructions, cutting the number of projection views needed from hundreds to dozens. However, these methods still face the bottleneck of reconstructing images from single-digit views since they need enough views to perform the reconstruction via the Radon transform for further quality enhancement. An alternative approach is stochastic reconstruction. Stochastic reconstruction utilizes deep convolutional generative adversarial networks (GANs) as layered architecture to extract and reproduce the hierarchical features of image datasets. [2] From representative 2D CT scan slices, GANs may stochastically generate a statistically equivalent 3D structure of porous media, [17] three-phase electrodes, [2,18] and anisotropic polymer membranes. [2] The synthetic volumes via stochastic approaches demonstrate statistical similarity with the original datasets, while they cannot reveal the real 3D structure from the input 2D images. The loss of localized geometric information is inevitable. Therefore, they are ineffective for transient *in situ* and operando studies or heterogeneity studies, where realistic variations in localized properties across time or space matter.

Ultra-sparse view 3D tomography reconstruction with deep learning methods is an under explored area due to model limitations and insufficient training data. Herein, we introduce a learnable framework capable of reconstructing high-fidelity 3D microstructures from single-digit radiographic images (Tomo-NeRF). Tomo-NeRF is distinct from the quality enhancement and stochastic reconstruction methods and does not rely on the analytical knowledge of inverse transform for reconstruction. Instead, Tomo-NeRF is a physics based

approach which utilizes the Beer-Lambert law to reconstructe samples. Enlightened by pixelNeRF,[19] a photorealistic novel view synthesis model, Tomo-NeRF aligns pixel and voxel in the 3D coordinate system via relating the x-ray attenuation to the material properties of the sample. The algorithm learns to predict the phase of each voxel with the grayscale value from 2D radiographic images. We train the model with a set of multi-view radiography projections along with the 3D voxel labels so that the model can learn the scene prior to reconstruction from two-view radiographic images. Obtaining thousands of 3D and 2D training datasets from real experiments for training is cost-prohibitive. Herein, we utilize a numerical simulator to create synthetic data. Synthetic reconstructions are combined with physics based equations (e.g. Beers-Lambert) to generate artificial radiographic images for the 2D training datasets. To demonstrate the functionality of Tomo-NeRF and the artificial training datasets, we conduct real full-view tomography imaging experiments on a two-phase granular ceramic proppants with mono-dispersed sizes. These ceramic proppants are made of bauxite and are commonly utilized for high temperature heat transfer from concentrated solar power plants. The bauxite particle heat exchangers are comprised of spheroid materials which can exhibit a range of aspect ratios and varying morphologies. Herein, we combine synthetic data with experimental data to evaluate the ability for Tomo-NeRF to reconstruct reconstruct 3D image from less than 10 views.

Tomo-NeRF

Tomo-NeRF is a learning framework that can reconstruct 3D images from a limited number of fixed radiographic projection views. Neural radiance field (NeRF) was previously developed for novel photorealistic views synthesis with an incomplete set of photos.[20] It leverages a continuous volumetric radiance field of color and density for volumetric scene prediction and uses gradient descent to optimize the scene using the input photos. Our Tomo-NeRF employs the convolutional approach to the conventional NeRF, pixelnerf.[19] Convolutional

layers in convolutional neural networks (CNNs) utilize a local receptive field defined by the size of the convolutional kernel.[21] This design enhances network's flexibility and computational efficiency when extracting image features, compared to fully connected structure. Therefore, CNNs would be more suitable for computed tomography reconstruction in tolerating potential noise within the input radiographic images. To realize the ultra-sparse view reconstruction in X-ray computed tomography, we embed the physical field of X-ray imaging into NeRF, making it correspond to real experiments.

Physical field of X-ray imaging

Radiographic images are 2D images which capture variable attenuation characteristics in a material or component.[22] An X-ray source emits fluxes of X-ray photons that pass through or interact with the sample.[23] Any interaction, via scattering or absorption, removes the photon from the X-ray beam.[24] The attenuation coefficient (μ) is a material property which describes how easily a X-ray can pass through a material. The attenuation coefficient of a material depends on the type of material (atomic number and density) and the photon energy of the radiation.[25] The X-ray attenuation often varies in space for heterogeneous materials comprised of either multiple materials or local variations in density. The detector electronically detects the photons that pass through the sample. Other than the material types and incident photon energy, the decrease in detected X-ray intensity is dependent upon the depth of X-ray beam penetration through the sample (i.e., material heterogeneity).[24] The computer will further process the through-thickness projection data (e.g., adjust brightness and contrast) and visualize it as a grayscale radiographic image. Beer-Lambert law relates the X-ray beam attenuation to the material properties (attenuation coefficient) and the penetration depth (z). With a parallel beam approximation, the general equation to calculate detected X-ray intensity I_d is:

$$I_d(x,y) = \int_0^{\varepsilon_{max}} \eta(\varepsilon) I_0(\varepsilon) \exp(-\int \mu(x,y,z;\varepsilon) dz) d\varepsilon \tag{1}$$

where $\eta(\varepsilon)$ is the quantum efficiency of the detector and $I_0(\varepsilon)$ is the incident X-ray intensity with the unit of photons per unit area per unit electron energy (photons/(m^{2}keV)). The incident X-ray intensity is a function of X-ray energy (ε). Variations in material properties are more detectable than variations in X-ray energies, particularly when utilizing a synchrotron facility that can generate an almost mono-chromatic X-ray source.[24] Therefore, Tomo-NeRF neglects the incident energy variation and simplifies the Beer-Lambert Law function as only an integral function of penetration thickness:

$$I_d(x, y) = \eta \cdot I_0 \cdot \exp(-\int \mu(x, y, z)dz) \tag{2}$$

The Beer-Lambert law function allows for a continuous 3D field representation and aligns material microstructure in spatial location (x, y, z) with the radiographic projection on the 2D plane (x, y).

X-ray computed tomography 3D reconstruction requires through-thickness information from different views (conventional methods would require hundreds of views). Therefore, the tomography imaging experiment involves radiographic scanning around a sample by fixing the X-ray source and detector and rotating the sample at a constant altitude around the central axis.[26] To represent spatial location of X-ray beams from different imaging direction, we encode the projection direction into Cartesian coordinate system. We sample points along the 0 deg projection beam as a reference coordinate (x_0, y_0, z_0), and rotate the X-ray parallel beams in the reverse direction of the sample rotation. For a field of view of $w \times w$, the reference point (x_0, y_0, z_0) on X-ray parallel beams rotates angle θ around the central vertical axis (i.e., y-axis) where $(x_c, y_c) = (w/2, w/2)$ will be (x_1, y_1, z_0). Here,

$$x_1 = (x_0 - x_c)\, cos\,(\theta) - (y_0 - y_c)\, sin\,(\theta) + x_c \tag{3}$$

and

$$y_1 = (x_0 - x_c)\, sin\,(\theta) - (y_0 - y_c)\, cos\,(\theta) + y_c \tag{4}$$

Model pipeline

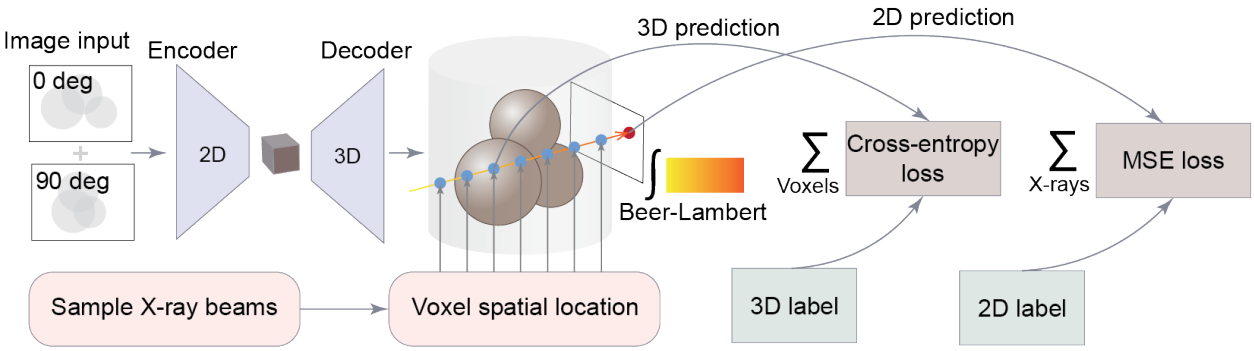

Figure 1: The overview of Tomo-NeRF pipeline (memory size of whole model is ∼80MB)

Tomo-NeRF (Figure 1) utilize a encoder-decoder convolutional neural network to extract features from images and align the features of the input image with 3D geometry in a voxel grid (see **Methods**). Tomo-NeRF uses a modified ResNet34 as a 2D encoder to extract local and global information from radiography images. 3D decoder further process the output feature vector from 2D encoder with the upsampling methods of transposed convolution and bilinear interpolation. It reshapes and restores the feature vector into a 3D feature map that aligns with the 3D ground truth in the voxel grid. The last layer of convolution neural network within 3D decoder is a sigmoid activation function $S(x) = \frac{1}{1+e^{-x}}$ to bound the output between 0 and 1,[27] turning each voxel grid into a classifier. For a two-phase sample (e.g., ceramic proppants), the numerical prediction in each voxel within the 3D feature map represents the probability *prob* of finding the solid phase in that spatial location. If $prob > 0.5$, the grid has a higher chance to be occupied by a solid material (e.g., ceramic), otherwise it is unoccupied (e.g., air).

During the training process (see **Supporting Information**), the encoder-decoder convolutional neural network takes the 0 deg and 90 deg projection images as input and produces a 3D feature grid of probability information that matches the original size of the volumetric grid ($100 \times 100 \times 150$). The model samples voxel grids, along X-ray beams among the six projection directions (i.e., beam rotation angle θ). Each direction ranges from 0-180 deg with a step of 30 deg. It sends the query specification of the spatial location along with the input

image into the Tomo-NeRF network to retrieve the corresponding feature from the predicted 3D feature grid. The model compares the retrieved 3D prediction along the X-ray beam with the 3D ground truth and has a cross-entropy loss (L_{CE}) for each grid to learn from the 3D dataset. Tomo-NeRF programs that each X-ray beam in the network may produce a grayscale pixel on the radiographic image under a given projection direction. Therefore, the model can further make a 2D prediction of pixel value by integrating the 3D predictions on the specified X-ray beams with the Beer-Lambert law (Equation 2). By comparing the 2D prediction (i.e., a radiographic image from 3D predictions) with the radiography ground truth corresponding to its projection direction, the model can learn from the 2D dataset (i.e., six direction radiographic images for each sample) via mean squared error (MSE) loss. By continuously minimizing cross-entropy (Equation S2 in **Supporting Information**) and MSE (Equation S3 in **Supporting Information**) losses during the training process, Tomo-NeRF can learn to increase the probability of predicting the voxel phase from 3D labels and six views. Learning from one projection direction can serve as relevance and share knowledge for other directions. The model can rely more directly on the input image feature if the query projection direction and input orientation are similar; otherwise, the model needs to leverage the learned prior.

There are three reasons for further making projections on the 3D predictions. Firstly, it enables Tomo-NeRF to learn from both 3D labels and 2D labels, allowing for learn the scene prior to reconstruction. In addition, for the testing or experimental validation process, in which only radiographic images are available, the trained network optimizes itself through image features and the grayscale label from the input of projection views. Furthermore, for the actual experiments (e.g., *in situ* imaging) that lack 3D labels to verify the fidelity of the deep reconstruction results, we may compare the 2D predictions on the same projection angle with the input images and assess the novel view quality.

Results

Tomo-NeRF reconstruction capabilities were evaluated on an bauxite granular material. Granular media are commonly used for high temperature heat transfer, catalysis, food, and pharmaceutical applications.[28,29] Understanding the kinetics of dense granular flow via a non-invasive method is essential to flow phenomena studies (e.g., Brazil nut effect,[30] pattern formation,[31] jamming transition,[32,33] and local rearrangement,[34] etc.) and heat transfer analysis.[35,36] However, dynamic granular flow studies with a non-invasive method have been limited to radiographic studies due to the rate dependency in granular flow. Tomo-NeRF enables prediction of complex 3D microstructures from a limited number of radiographic images.

X-ray micro-computed tomography imaging was conducted on a sintered bauxite proppant sample (see **Methods**). We obtain both real tomographic (reconstruction from full-view) and radiographic images as a reference sample (i.e., real sample) for evaluating reconstruction performance. Training Tomo-NeRF requires thousands of 2D and 3D image training datasets for an accurate reconstruction. Acquisition of such a large number of real datasets faces the limitations of time, cost, and synchrotron source availability. Herein, we propose an artificial image generator (see **Methods**) as a means for obtaining an affordable training dataset. The artificial image generator prepares two different types of datasets: (1) synthetic data with regular spheres and (2) synthetic data with irregular spheres (non-smooth). The first type of dataset is randomly packed identical spherical particles with 600 μm diameter in a 3 mm inner diameter tube (i.e., artificial sphere sample). The second dataset is randomly packed mono-dispersed particles in a 3.95 mm inner diameter tube. The irregular spheres more closely resemble the real bauxite materials (e.g. ceramic proppant) used in the experimental tomography imaging experiment.

To validate the reconstruction capabilities of Tomo-NeRF, we conduct four deep learning experiments (Table 1). Each experiment used a specific type of synthetic data set as the training data and a specific input to test the algorithm's ability to convert two-, four- and

Table 1: Summary of four experiments with input data and training input data for model.

Experiment #	**Input Data**	**Training Data for Model**
Experiment 1	Real XCT	Synthetic Data - Irregular Spheres
Experiment 2	Synthetic Data- Irregular Sphere	Synthetic - Data Irregular Sphere
Experiment 3	Synthetic Data- Regular Sphere	Synthetic Data - Regular Sphere
Experiment 4	Synthetic Data - Irregular Sphere	Synthetic Data - Regular Sphere

six-views into a 3D image. Two-view means two radiographic images taken from two different angles, while six-views means six radiographic images taken from six different angles. Conventional tomography combines 1000s of radiographs from 1000s of different angles to created 3D images. All input data sets were original and were not utilized in the training process. Experiment #1 uses real x-ray radiographic images of the bauxite particles as an input for the model trained with synthetic data comprised of irregular spheres (non-smooth). The first experiment specifically aims to evaluate the practicality of the proposed Tomo-NeRF in real XCT applications. Experiment #2 and #3 use the synthetic data as the input data and utilize algorithms trained with their respective types of data-sets (Table 1). These experiments aim to validate the feasibility of the proposed Tomo-NeRF's reconstruction principal and to investigate the the role of geometric complexity (e.g. uniform vs. irregular spheres) on the reconstruction results. Experiment #4 utilizes radiographic images of artificial irregular sphere samples into the model trained by the artificial spherical shape sample dataset. Tomo-NeRF takes the two-view, four-view, or six-view images as the input (Figure 2a) and reconstructs a 3D image as the output (Figure 2b). We visualize the 2D prediction images by further making projections on the 3D output in both the same and different directions of the input images (i.e., validation view and novel view). The novel view represents the model reconstruction capability in a direction that is different from the projecting direction of the input images. Other than visual comparison, we quantitively evaluate the reconstruction accuracy with 3D structural similarity (SSIM), voxel-wise accuracy and the standard deviation of localized porosity (see **Methods**) by using the 3D ground truth sample as the reference.

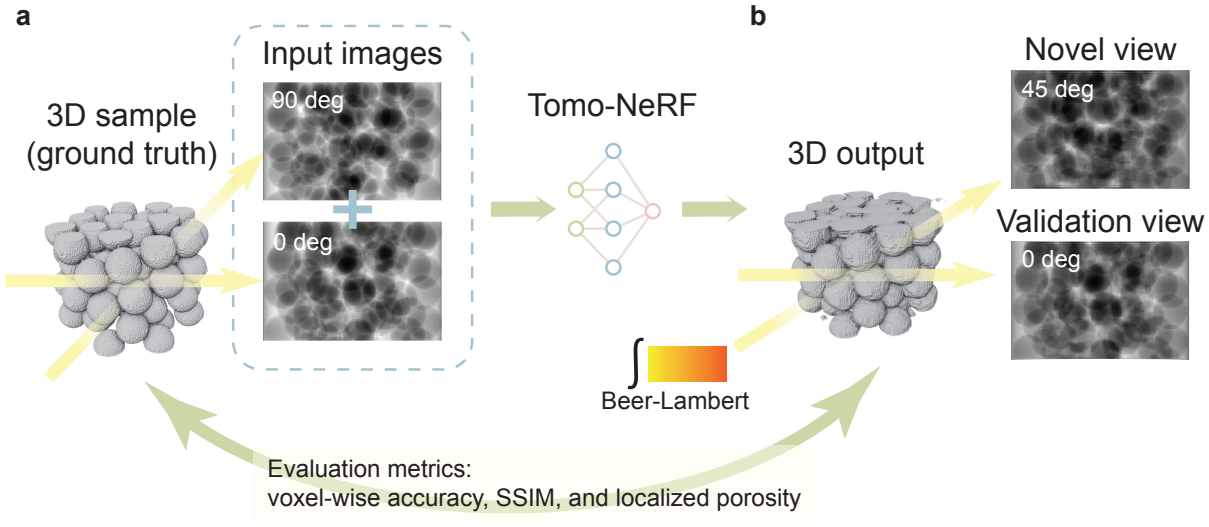

Figure 2: An example of a two-view reconstruction for an artificial sphere sample (in a voxel-wise accuracy of 90.39%) for illustrating the experimental workflow. **a**. The artificial image generator first produces a 3D sample of identical sphere particles with a discrete element method simulator and then makes artificial radiographic images in 0 and 90 deg projecting direction, respectively. **b**. Tomo-NeRF takes the concatenated two images as input and predicts a 3D image. Beer-Lambert Law applies to the 3D output to make 2D radiographic projections in the same projection angle of input images as the validation view (e.g., 0 deg) and the different projection angle as the novel view (e.g., 45 deg).

The reconstruction accuracy increases with an increasing number of input views for each of the four experiments (Figure 3a). The rate of improvement decreases for Experiment #1, #2, and #3 but steadily increases for Experiment #4. Experiment #3 which uses uniform spheres as an input and an algorithm trained on synthetic data with uniform spheres has the highest accuracy across all number of views. Eight synthetic samples in the testing process have an average accuracy of 88.71% in two-view reconstruction, and this increases to 96.63% with six-input views. It is unsuprising that Tomo-NeRF performs better in reconstructing images smooth and predictable surfaces and regular sizes (e.g. uniform spheres). Reconstructing 'real XCT' data (Experiment # 1) results in lower reconstruction accuracy. The real x-ray dataset has voxel-wise accuracy around 81.83%, 87.77%, and 89.59% for two-view, four-view, and six-view reconstructions, respectively. The fourth experiment, which incorporates irregular shaped spheres into a model trained for regular spheres has the lowest

reconstruction accuracy with 2 views. However, the fourth experiment shows the greatest improvement in voxel-wise accuracy with increasing projection views.

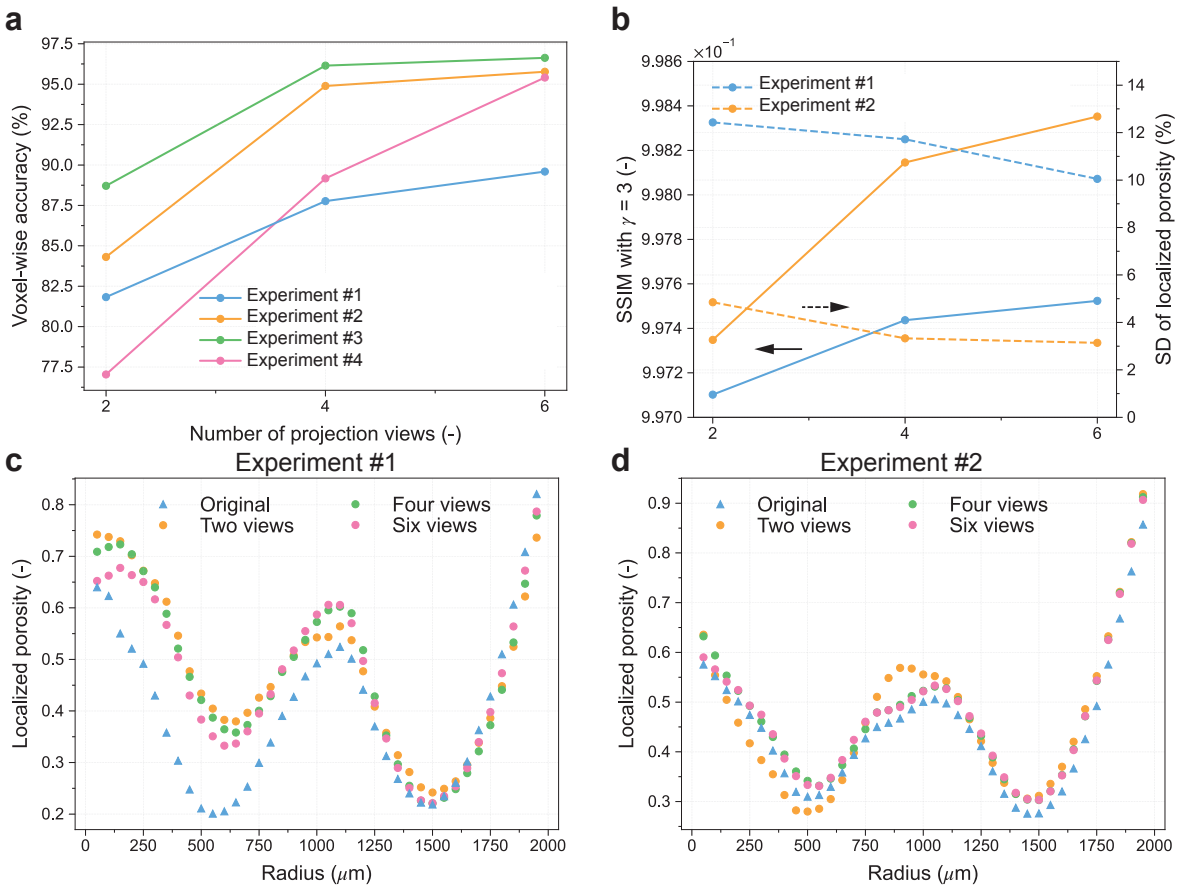

Figure 3: **a**. The reconstruction voxel-wise accuracy increases with more available input projection views, and the new task samples have the highest increasing rate. **b**. Structural similarity (SSIM) with exponent for structural term (γ) of 3 exceed 0.997 and it increases with more input views. The standard deviation (SD) of localized porosity decreases with more views. **c**.The localized porosity radial profile from the center to the near wall region for the real sample (voxel-wise accuracy ranges from 81.83–89.59%) shows the major deviations occur near the central region. **d**.The localized porosity radial profile for the artificial sample (voxel-wise accuracy ranges from 86.40–95.37%) has a smaller deviation than the real sample.

Similar to the voxel-wise accuracy, the trend for structural similarity (SSIM) increases with the number of projection views (Figure 3b) for the real XCT data (e.g. Experiment #1) and synthetic irregular spheres (e.g. Experiment #2). Tomo-NeRF shows a high fidelity in structural restoration. When the input data is real XCT data the two-view reconstruction has a low accuracy but the structural similarity (SSIM) still reaches as high as 0.9971. The

highest value of structural similarity (SSIM) is 0.9984 for the six-view reconstruction of the artificial sample (e.g. synthetic data with irregular spheres). Another way to evaluate the resiliency of the reconstruction is to extract local microstructure properties (e.g. porosity) for the different particle beds. The porosity radial profile through the reconstructed sample was systematically compared for the true XCT data (e.g. Experiment #1) and input data containing irregular spheres (e.g. Experiment #2). Figure 3c and d demonstrates the radial porosity for input data (e.g. original) and 2-, 3-, and 6-view reconstructions. The real XCT data (e.g. Experiment #1) demonstrates wide porosity variations between the different views (Fig. 3c) and capture the sub-surface variability in the learning exercise. The standard deviation of localized porosity distribution is between 3.14-4.85% for the synthetic data with irregular spheres (Fig. 3d, Experiment #2) and 10.05-12.43% for the real XCT sample (Fig. 3c, Experiment #1)

Synthetic data was constructed with irregular shaped spheres of varying sizes (e.g. Experiment #2). The 3D ground truth (Figure 4a) and predicted images (Figure 4c) demonstrate that all the reconstructions (2-, 4- and 6-views) demonstrate similar packing densities as the ground truth. As the number of projection views increases, the details of the particles' shape are more accurate and the surface is smoother. The adhesion between some of the particles is attenuated since the reconstructed particles' shape is more accurately represented. The 2D predictions in the validation views (Figure 4d) shows a high similarity to the input view (Figure 4b) at the same projection angle (0 deg) in all reconstructions except for the high contrast particle overlapping region in the two-view reconstruction. The novel view projections define more clearly particle boundaries as the number of views increases to four (Figure 4d).

Similar to the synthetic sample with irregular particles, the reconstruction of the real XCT data appears more similar to the input data with higher input views (Figure 5c). However, the overall reconstruction quality is not as high as the artificial sample in Figure 4. The XCT data has more artifacts due to x-ray scattering and local variations in the material which leads

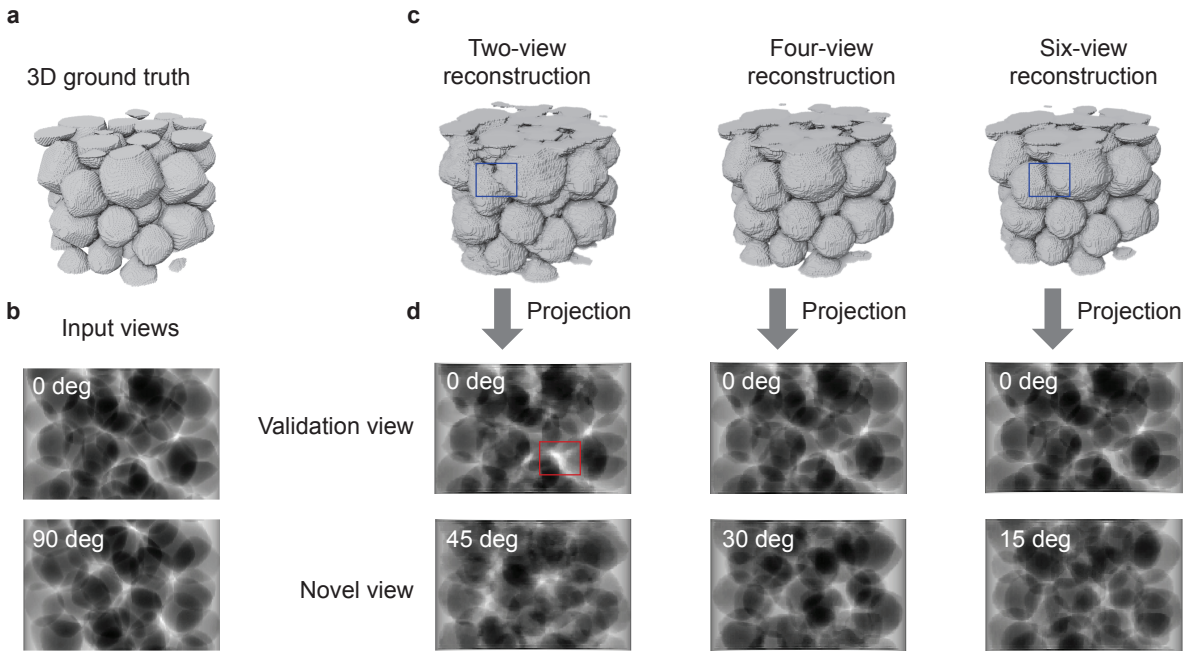

Figure 4: **a**. 3D ground truth of an example of synthetic sample of irregular shape (Experiment #2). **b**. Input view examples in 0 deg and 90 deg. **c**. 3D visualization of two-view, four-view, and six-view predictions. There is less adhesion between some of the particles as more input views (blue box). **d**. The validation views exhibit a strong overall structural similarity to the input view. In the two-view reconstruction, there is a high-contrast region has a marginally poor restoration of details (red box) and it gets improved with more input views. In the novel view visualizations, individual particles outlines is becomes more distinct with more input views.

to deviation between the assumptions of the model and artificial image generator. Tomo-NeRF and artificial image generators assume a monochromatic beam. In reality, synchrotron facilities produce monochromatic beams (I_0) through single-crystal monochromators and X-ray diffraction at a higher intensity than the in-house imaging facility in our experiment.[37] The in-house imaging facility uses the filter to decrease the intensity of the beams with an unwanted wavelength, yet the emitted radiation is still polychromatic, which is different from the ideal assumption (Eq. 2). This deviations results in an attenuation coefficient which is dependent on the local spatial heterogenity in the sample and X-ray energy (i.e., $\mu(x, y, z; \varepsilon)$). In addition, the lower energy of the in-house facility increases the contrast I_d/I_0 (e.g., the overexposure for the void region near the overlapping particles), and results

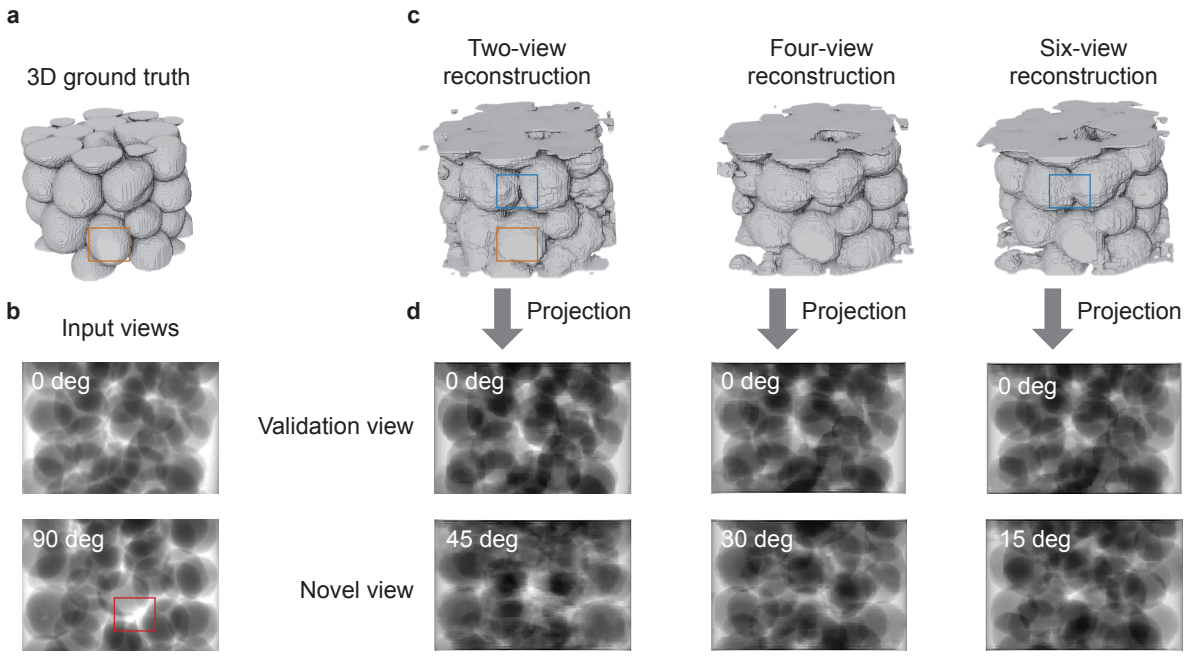

Figure 5: **a**. 3D ground truth of tomographic image reconstructed from the full-view X-ray imaging experiment (Experiment #1) **b**. Input view examples in 0 deg and 90 deg.The lower energy intensity and noise in the real in-house imaging experiment blur the outline of particles' projection on the radiographic images, resulting in information loss (red box). **c**. 3D visualization of two-view, four-view, and six-view predictions. More input views tend to aggravate the particle adhesion (blue box). **d**. With more input views, the validation views shows a higher structural similarity to the input view. In the novel view visualizations, the outline of particles' projection gets clearer with more input views.

in more noise in the real experiment (Figure 5b). By comparing the input images from the real sample (Figure 5b) and that from the artificial samples (Figure 4b), we can see that there is information loss in the particle shape, where the boundaries of the overlapping particles cannot be resolved due to high contrast and noise. It is more difficult for the convolution network to extract features from blurrier input radiographic images, leading to higher deviations of localized porosity in the central region (Figure 3c).

More input views mitigate the particle adhesion phenomenon when synthetic data is in the input (Figure 4c). However, more artifacts due to particle stacking with increasing views is observed with real XCT data (blue box in Figure 5c). The cone beam of the x-ray source and sample shifting during imaging can lead to this observed flaw when the input data is real

radiograph. In Tomo-NeRF and the artificial generator, we assume parallel-beam geometry for the flux of photons with the X-ray point source at an infinite distance from the testing object. In a real experiment, the distance between the X-ray source and the object is limited, forming a cone beam at an angle of 3.28 degrees. The cone beam can distort the projection image away from the angle bisector of the cone. The higher energy for the photon flux near the angle bisector results in a brighter image (i.e., overexposure). In addition to cone-effect, slight misalignment during the experimental setup can result in part of the sample to be cut off in the field of view (orange box in Figure 5a). Particles obstructed or cut-off from the field of view will not emerge in reconstructed experimental data (orange box in Figure 5b). The final challenge with experimental data is that Tomo-NeRF assumes the sample rotates around a central axis. Any shift in the central axis will change the projection angle in the input data. Over long imaging experiments, where 1000s of images are taken, it is common for some shifting to occur. One of the benefits of the sparse-view imaging approach is a decrease in imaging time and thus reduction of these artifacts.

According to the 2D validation views and the novel views from the 3D predictions, the validation views capture the input views structure (Figure 5d). More input views increase the sharpness of novel views, indicating a better reconstruction quality in the given projecting angle. There are visible differences between the 2D predictions and the input views. It is due to the inaccurate assumption of the attenuation coefficient in particles and quartz containers. The incident energy of the polychromatic beam is uncertain and the particles are made of mixtures with indeterminate proportions of materials. The model optimizes by minimizing the difference between the pixel predictions and the 2D labels in a batch. Therefore, as long as the assumption of the attenuation coefficients is within the acceptable range and the 2D predictions can reflect the main structure, the inaccurate assumption of the attenuation coefficient only has a limited effect on the 3D reconstruction results.

In addition to the above-mentioned variations between the experiment and modeling, non-uniform information density within the 3D decoder (Figure 1) can also impact the re-

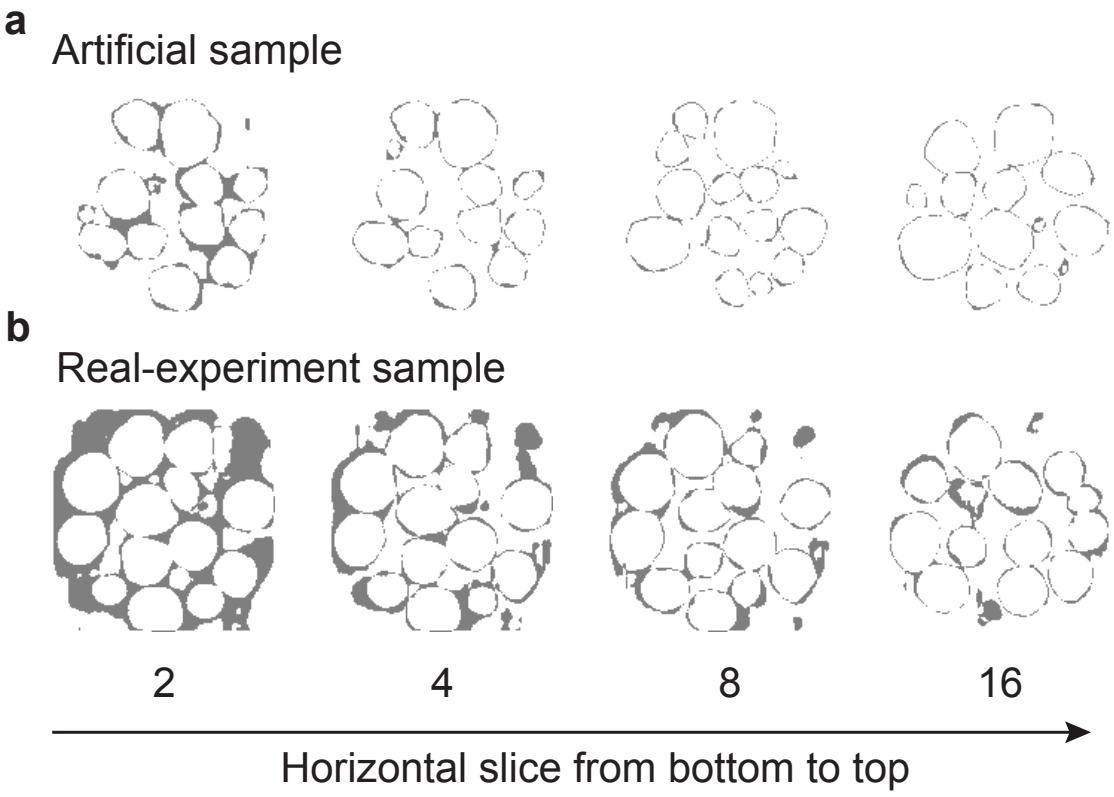

Figure 6: Visualization of prediction errors of synthetic data sample (**a**) (Experiment #2) and real sample (**b**) (Experiment #1) from bottom to top. For the slices closer to the central, the error mainly occurs near the particle surfaces. The information density gradient formed and exacerbated as it propagates in the multilayer transposed convolution neural network. It causes the inferior prediction accuracy near the top/bottom surface of the sample.

construction quality near the outer surfaces of the 3D predictions. We visualize the prediction error of the six-view reconstructions for the synthetic irregular sphere sample (Figure 6a) and the real sample (Figure 6b) on the selected horizontal slices. The model is less likely to accurately discern and reconstruct the particles at the bottom of the sample which makes the particles look fused together. This issue is more exacerbated with real XCT data (Experiment #1, Figure 6b). When the distance from the bottom/top is more than four voxels, the reconstruction quality improves. During the transposed convolution operation, the voxels further away from the edge are able to receive a contribution from the adjacent voxel, i.e., overlapping information.[2] The lack of overlapping information at the edge of the sample forms an information density gradient. The information density gradient becomes

exacerbated as it propagates in the multilayer transposed convolution neural network. Furthermore, the particle projections near the top and bottom edges of the radiographic images are incomplete, thereby more irregular than the particles of the complete shape. The low information density and the more irregular shapes cause the low-quality reconstruction region near the top and bottom of the sample. This can be potentially resolved by subsequently feeding partially overlapping radiographic images into the model, and then stitching together the central part of the reconstruction (with a higher prediction accuracy). The error map for the region away from the bottom/top shows that Tomo-NeRF manages to restore the position and general shape of the particles. However, errors primarily occur at the surface of the particles, indicating that the restoration of shape details is insufficiently accurate, particularly for the real-experiment sample.

Conclusion

Three-dimensional x-ray imaging is an important materials characterization approach which enables non-destructive and sub-surface evaluation of a wide range of materials. There is a growing trend in 4-D imaging approaches which combine 3D imaging with time. The latter is specifically useful for observing dynamic operating conditions also known as *in situ* or *operando*. One of the limitations in 4D imaging is the need for 1000s of images which severely limits the temporal resolution of such a technique. Herein, we introduce a sparse-view reconstruction algorithm known as Tomo-NeRF. Tomo-NeRF is a physics based learning approach which enables the reconstruction of 3D images with limited 2D views (<10 views). The results show the ability of Tomo-NeRF to reliably restore the structure of 3D two-phase granular media from two-view 2D radiographic images in a high structural similarity of 0.9971 for real XCT images and 0.9973 for synthetic images with irregular particle shape. The voxel-wise accuracy for real images (81.83–89.59%) is lower than synthetic images (86.40–95.37%). The deviation is likely because of experimental considerations which

do not obey model assumptions (e.g., low incident energy, differences in attenuation coefficients, polychromatic and cone beams, and sample shifting). Reconstruction of complex structures with high fidelity with single-digit radiographic images can significantly impact the temporal resoulution of 4D-imaging. Future work aims to test the resiliency of these algorithms on dynamic imaging experiments.

References

(1) Taillandier-Thomas, T.; Roux, S.; Hild, F. Soft route to 4D tomography. *Physical review letters* **2016**, *117*, 025501.

(2) Kench, S.; Cooper, S. J. Generating three-dimensional structures from a two-dimensional slice with generative adversarial network-based dimensionality expansion. *Nature Machine Intelligence* **2021**, *3*, 299–305.

(3) Dong, J.; Fu, J.; He, Z. A deep learning reconstruction framework for X-ray computed tomography with incomplete data. *PloS one* **2019**, *14*, e0224426.

(4) Andersson, F. Fast inversion of the Radon transform using log-polar coordinates and partial back-projections. *SIAM Journal on Applied Mathematics* **2005**, *65*, 818–837.

(5) Withers, P. J.; Bouman, C.; Carmignato, S.; Cnudde, V.; Grimaldi, D.; Hagen, C. K.; Maire, E.; Manley, M.; Du Plessis, A.; Stock, S. R. X-ray computed tomography. *Nature Reviews Methods Primers* **2021**, *1*, 1–21.

(6) Quinto, E. T. An introduction to X-ray tomography and Radon transforms. Proceedings of symposia in Applied Mathematics. 2006; p 1.

(7) Dixit, M. B.; Moreno, D.; Xiao, X.; Hatzell, M. C.; Hatzell, K. B. Mapping charge percolation in flowable electrodes used in capacitive deionization. *ACS Materials Letters* **2019**, *1*, 71–76.

(8) Wang, G.; Ye, J. C.; De Man, B. Deep learning for tomographic image reconstruction. *Nature Machine Intelligence* **2020**, *2*, 737–748.

(9) Ahishakiye, E.; Bastiaan Van Gijzen, M.; Tumwiine, J.; Wario, R.; Obungoloch, J. A survey on deep learning in medical image reconstruction. *Intelligent Medicine* **2021**, *1*, 118–127.

(10) Dong, X.; Vekhande, S.; Cao, G. Sinogram interpolation for sparse-view micro-CT with deep learning neural network. Medical Imaging 2019: Physics of Medical Imaging. 2019; pp 692–698.

(11) Lee, H.; Lee, J.; Cho, S. View-interpolation of sparsely sampled sinogram using convolutional neural network. Medical Imaging 2017: Image Processing. 2017; pp 617–624.

(12) Lee, D.; Choi, S.; Kim, H.-J. High quality imaging from sparsely sampled computed tomography data with deep learning and wavelet transform in various domains. *Medical physics* **2019**, *46*, 104–115.

(13) Gjesteby, L.; Yang, Q.; Xi, Y.; Zhou, Y.; Zhang, J.; Wang, G. Deep learning methods to guide CT image reconstruction and reduce metal artifacts. Medical Imaging 2017: Physics of Medical Imaging. 2017; pp 752–758.

(14) Lee, H.; Lee, J.; Kim, H.; Cho, B.; Cho, S. Deep-neural-network-based sinogram synthesis for sparse-view CT image reconstruction. *IEEE Transactions on Radiation and Plasma Medical Sciences* **2018**, *3*, 109–119.

(15) Xie, S.; Zheng, X.; Chen, Y.; Xie, L.; Liu, J.; Zhang, Y.; Yan, J.; Zhu, H.; Hu, Y. Artifact removal using improved GoogLeNet for sparse-view CT reconstruction. *Scientific reports* **2018**, *8*, 1–9.

(16) Davoudi, N.; Deán-Ben, X. L.; Razansky, D. Deep learning optoacoustic tomography with sparse data. *Nature Machine Intelligence* **2019**, *1*, 453–460.

(17) Feng, J.; Teng, Q.; Li, B.; He, X.; Chen, H.; Li, Y. An end-to-end three-dimensional reconstruction framework of porous media from a single two-dimensional image based on deep learning. *Computer Methods in Applied Mechanics and Engineering* **2020**, *368*, 113043.

(18) Gayon-Lombardo, A.; Mosser, L.; Brandon, N. P.; Cooper, S. J. Pores for thought: generative adversarial networks for stochastic reconstruction of 3D multi-phase electrode microstructures with periodic boundaries. *npj Computational Materials* **2020**, *6*, 1–11.

(19) Yu, A.; Ye, V.; Tancik, M.; Kanazawa, A. pixelnerf: Neural radiance fields from one or few images. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2021; pp 4578–4587.

(20) Mildenhall, B.; Srinivasan, P. P.; Tancik, M.; Barron, J. T.; Ramamoorthi, R.; Ng, R. Nerf: Representing scenes as neural radiance fields for view synthesis. *Communications of the ACM* **2021**, *65*, 99–106.

(21) Yu, W.; Yang, K.; Bai, Y.; Yao, H.; Rui, Y. Visualizing and comparing convolutional neural networks. *arXiv preprint arXiv:1412.6631* **2014**,

(22) Lenoir, N.; Bornert, M.; Desrues, J.; Bésuelle, P.; Viggiani, G. Volumetric digital image correlation applied to X-ray microtomography images from triaxial compression tests on argillaceous rock. *Strain* **2007**, *43*, 193–205.

(23) Stevenson, A. W.; Gureyev, T. E.; Paganin, D.; Wilkins, S.; Weitkamp, T.; Snigirev, A.; Rau, C.; Snigireva, I.; Youn, H.; Dolbnya, I., et al. Phase-contrast X-ray imaging with synchrotron radiation for materials science applications. *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms* **2003**, *199*, 427–435.

(24) Russo, P. *Handbook of X-ray imaging: physics and technology*; CRC pres: Boca Raton, 2018.

(25) Akman, F.; Durak, R.; Turhan, M. F.; Kaçal, M. R. Studies on effective atomic numbers, electron densities from mass attenuation coefficients near the K edge in some samarium compounds. *Applied Radiation and Isotopes* **2015**, *101*, 107–113.

(26) Dixit, M. B.; Zaman, W.; Bootwala, Y.; Zheng, Y.; Hatzell, M. C.; Hatzell, K. B. Scalable manufacturing of hybrid solid electrolytes with interface control. *ACS applied materials & interfaces* **2019**, *11*, 45087–45097.

(27) Narayan, S. The generalized sigmoid activation function: Competitive supervised learning. *Information sciences* **1997**, *99*, 69–82.

(28) Rycroft, C. H.; Grest, G. S.; Landry, J. W.; Bazant, M. Z. Analysis of granular flow in a pebble-bed nuclear reactor. *Physical review E* **2006**, *74*, 021306.

(29) Jovanović, A.; Pezo, M.; Pezo, L.; Lević, L. DEM/CFD analysis of granular flow in static mixers. *Powder technology* **2014**, *266*, 240–248.

(30) Möbius, M. E.; Lauderdale, B. E.; Nagel, S. R.; Jaeger, H. M. Size separation of granular particles. *Nature* **2001**, *414*, 270–270.

(31) Sandnes, B.; Flekkøy, E.; Knudsen, H.; Måløy, K.; See, H. Patterns and flow in frictional fluid dynamics. *Nature communications* **2011**, *2*, 1–8.

(32) Cates, M.; Wittmer, J.; Bouchaud, J.-P.; Claudin, P. Jamming, force chains, and fragile matter. *Physical review letters* **1998**, *81*, 1841.

(33) Liu, A. J.; Nagel, S. R. Jamming is not just cool any more. *Nature* **1998**, *396*, 21–22.

(34) Zhai, C.; Albayrak, N.; Engqvist, J.; Hall, S. A.; Wright, J.; Majkut, M.; Herbold, E. B.; Hurley, R. C. Quantifying local rearrangements in three-dimensional granular materials: Rearrangement measures, correlations, and relationship to stresses. *Physical Review E* **2022**, *105*, 014904.

(35) Cheng, G. J.; Yu, A. Particle scale evaluation of the effective thermal conductivity from the structure of a packed bed: radiation heat transfer. *Industrial & engineering chemistry research* **2013**, *52*, 12202–12211.

(36) Zheng, Y.; Hatzell, K. B. Thermal-Economic Optimization of Moving Packed Bed Particle-to-SCO2 Heat Exchanger Using Particle Swarm Optimization. Energy Sustainability. 2021; p V001T02A010.

(37) Caciuffo, R.; Melone, S.; Rustichelli, F.; Boeuf, A. Monochromators for x-ray synchrotron radiation. *Physics Reports* **1987**, *152*, 1–71.