CS230

Enhanced visualization of fractured shale rocks using super resolution neural network

Manju Pharkavi Murugesu
Department of Energy Resources Engineering
Stanford University
mmuruges@stanford.edu

Abstract

Injection into deep geological formations is a promising approach for the utilization, sequestration, and removal from the atmosphere of CO_2 emissions. Laboratory experiments are essential to characterize how CO_2 flows and reacts in various types of geological media. We reproduce such dynamic injection processes while imaging using Computed Tomography (CT) at sufficient temporal resolution to visualize changes in the flow field. The resolution of CT, however, is on the order of 100's of μm and insufficient to characterize fine-scale reaction-induced alterations to micro-fractures. Super resolution deep learning is, therefore, an essential tool to improve spatial resolution of dynamic CT images. We acquired and processed pairs of multi-scale low- and high-resolution CT rock images. We also show the performance of our baseline model on fractured rock images using peak signal to noise ratio and structural similarity index. Coupling dynamic CT imaging with deep learning results in visualization with enhanced spatial resolution of about a factor of 4 thereby enabling improved interpretation.

1 Introduction

Low-carbon energy resources and large-scale greenhouse emissions reductions are important to meet the global demand for energy while combating climate change. Subsurface geology provides large scale solutions to our climate goals through technologies such as CO_2 utilization and sequestration, subsurface hydrogen and compressed air storage, and enhanced geothermal systems. For example, CO_2 can replace water in hydraulic fracturing processes while simultaneously sequestering, creating a reduced-carbon or carbon negative energy process [1, 2].

Injected CO_2 tends to flow preferentially through naturally existing or induced fractures. CO_2 and brine mixtures are quite acidic. The low pH brine chemically interacts with surface minerals along fractures, altering rock-fluid interactions through dissolution and precipitation of certain rock minerals. Such reactions can significantly impact the porosity and permeability along the fractures and near-fracture matrix, affecting both CO_2 flow and storage processes. Visualization techniques, such as Computed Tomography (CT) imaging, capture alterations of fractures during reactive transport experiments at the laboratory scale.

CT is an important technique to image rocks during dynamic transport experiments. The resolution of CT images, however, is insufficient to characterize accurately the changes in fractures and rock matrix porosity on the order of μm or less. Micro-CT (μCT) scanners have the necessary superior resolution to characterize reactive transport processes at the microscale. The scan time of μCT is longer than that of a CT by at least an order of magnitude and therefore limits the temporal resolution

CS230: Deep Learning, Spring 2021, Stanford University, CA. (LateX template borrowed from NIPS 2017.)

significantly. To overcome these challenges, we propose to image dynamic reactive transport using CT scanners and later upsample the CT image to have comparable resolution to μCT.

Dong et al. (2015) shows that deep convolutional neural networks (CNN) are state of the art superior methods for image super resolution tasks [3]. Past literature reports supervised super resolution deep learning models based on micro-CT and synthetically downsampled micro-CT image pairs for geological samples [4, 5, 6]. Low-resolution CT images are more applicable inputs to super resolution deep learning models due to the widespread use of CT in geoscience. The availability of CT and micro-CT imaging capability motivates us to train CNN models with low resolution CT as input images and high-resolution micro-CT as ground truth images. Hence, this work extends super resolution deep learning for multi-scale data obtained from separate imaging platforms.

2 Related Work

Following the trajectory of development in CNN models that improved image classification tasks, the Image Restoration and Enhancement community increasingly adapted similar structures in super resolution-based CNN models. Inspired by the VGG16 and VGG19, Kim et al. (2016) shows better performing CNN model with smaller filter sizes (3 x 3 spatial size of filter) and deeper network (20 layers), which increased both the non-linearity and lowered the number of parameters to tune [7]. Super resolution architectures are also improved by the addition of residual blocks that preserves input features by connecting output of previous layers to the output of new layers [8].

The feed forward deep CNN models learn non-linear mapping from low resolution inputs to higher resolution output. Due to homogeneity of rock images, it is important that the model is sensitive to fine details. Ledig et al. (2016) successfully combined feedforward deep residual network with generative adversarial network (GAN) for image super resolution and we implement our model based on this work as baseline [9]. Due to the lack of comparison between the image material, we also refer to super resolution work related to rock materials, where Wang et al. (2019) shows super resolution reconstruction for CT-obtained rock samples using 3D convolutional neural network, and Da Wang et al. (2019) shows resolution enhancement for digital rock images obtained from micro-CT [4, 5]. One possible challenge in comparing with similar materials is that the resolution of input rock image used by Wang et al (2019), and Da Wang et al. (2019) are superior to the input image in our work and may not be directly comparable.

3 Workflow

This project proposes a workflow shown in Figure 1 to conduct neural network super resolution of multi-scale CT rock images. The workflow outlines the steps to prepare images, train models and later effectively deploy the models for our application. The focus of the project has been dataset curation and model training.

3.1 Dataset and features

Image pre-processing is a crucial step in the project as the images obtained from CT and micro-CT are not directly comparable or translatable unless aligned by following the pipeline outlined in Figure 2. First, the 2D slices of CT and micro-CT go through the process of cropping, normalization and rotation until they are aligned. The pixel distribution of the low- and high-resolution image pairs should be as close as possible to indicate that the only difference between the two images is resolution. Once the 2D images are processed, they can be stitched together using nearest neighbour interpolation along the z-direction to form 3D volumes. For the initial study in this work, we first randomly cropped 256 x 256 pixels of micro-CT images that are equivalent to a 64 x 64 pixels of CT images. Due to memory bottleneck, we randomly cropped 64 x 64 x 64 voxels of micro-CT volumes that are equivalent to a 16 x 16 x 16 voxels of CT volumes. We used 10,000 images for 2D model training and 2000 volumes for 3D model training.

3.2 Model Training

We base the study on two types of network: deep residual convolutional neural network (CNN) and generative adversarial network (GAN). Feed-forward deep CNN models learn non-linear mapping

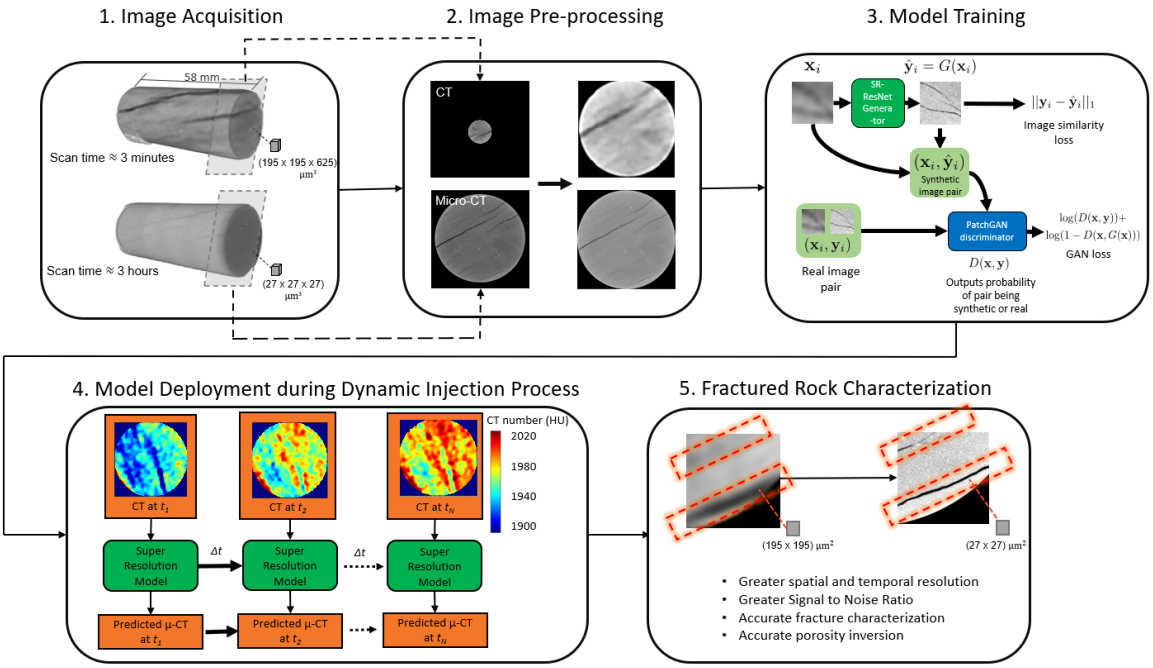

Figure 1: Proposed workflow for application of super resolution deep learning to characterize better fractured geologic porous media using enhanced visualization during dynamic injection processes.

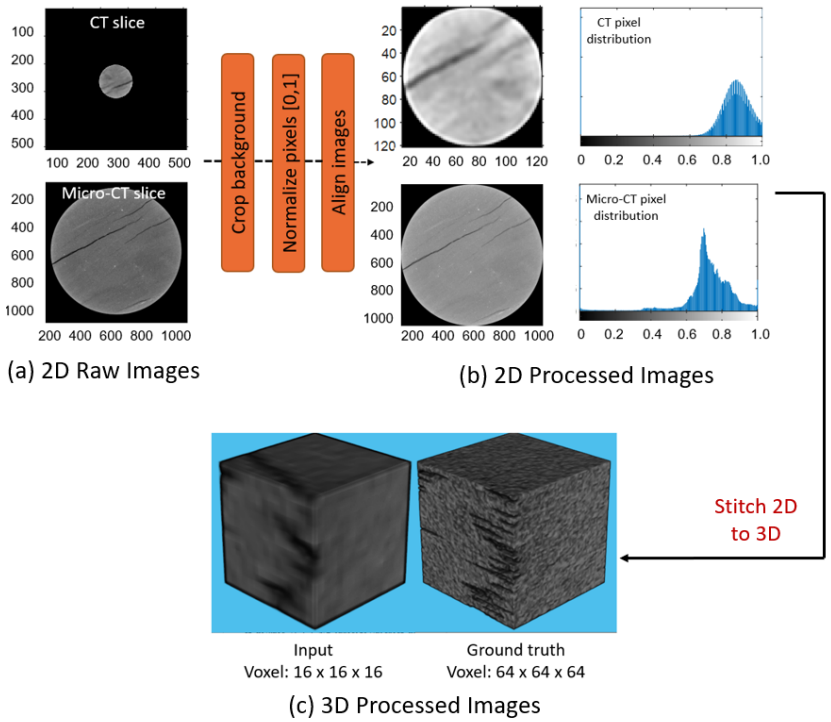

Figure 2: Image pre-processing pipeline from (a) raw 2D images to (b) processed 2D images. The 2D images are interpolated along the z-direction using the nearest neighbour method to create (c) processed 3D volume pairs

from low resolution inputs to higher resolution output. The feed-forward CNN model is based on the ResNet architecture proposed by Johnson et al. (2016) as shown in Figure 3 [10]. Due to homogeneity of rock images, it is important that the model is sensitive to fine details and produce realistic looking grains. GAN addresses this problem by incorporating an additional discriminator network that discriminates between synthetic high resolution image and actual high resolution counterpart. Although the images for this study are all obtained from CT-based instruments, the images are possibly offset due to default settings of separate CT and μCT machines. Therefore, the project adapts the multimodal image enhancement implementation suggested by Anderson et al. (2020) based on the Pix2pix conditional-GAN (cGAN) model from Isola et al. (2017) [11, 12]. Conditional GAN has a generator and a discriminator network. The generator is basically a feedforward CNN model, and the discriminator is a 3-layer CNN that outputs the probability for whether the generated high resolution image looks real or fake. The conditional-GAN is optimized based on two loss terms: GAN and image similarity loss terms, whereas the feedforward CNN model is optimized based on L1 loss only [11]. The 2D models are converted to 3D models by modifying the current layers, such as the convolutional and padding layers, in the neural network.

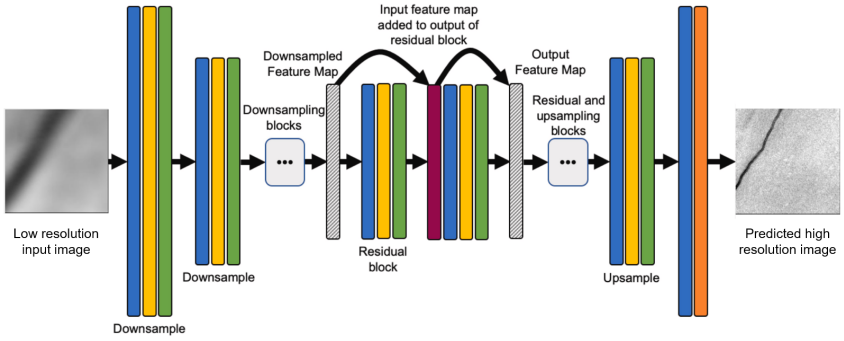

Figure 3: ResNet architecture based on Johnson et al. (2016)

4 Experiments/Results/Discussion

Several experiments are conducted using feedforward CNN and cGAN models to provide proof of concept for this study. Each model follows the original hyperparameters recommended by Anderson et al. (2020). The models were run for 20 epochs. The initial learning rate for 2D and 3D model is 0.0002 and 0.0001 respectively. The models could not converge at greater learning rate. We trained the 2D CNN models by tuning the number of residual blocks (6 or 9 residual blocks), downsample factor (a scale of 1, 2 and 4) and loss functions (L1, wasserstein and vanilla loss). Initial results for validation set are shown in Table 1. PSNR quantifies the proportion between the maximum possible power of a signal and the noise, and SSIM measures the structural difference between two images. No model in Table 1 refers to bicubic interpolation without model training.

Table 1 shows higher PSNR and SSIM for feedforward models compared to conditional-GAN models. For example, the cGAN predicted images from Figure 4 generates photo-realistic images with high perceptual content, but are lower in pixel to pixel accuracy as shown by the poor fracture segmentation [5]. Feedforward model accurately predicts the fractures, but the matrix and spatial resolution are washed out (Figure 4). Both Conditional-GAN and feedforward models show higher PSNR and SSIM for 4x downsampling. The model architectures are suitable for 4x scale difference between input and predicted images. In addition, L1 training loss generates best pixel-to-pixel accuracy for cGAN models.

The 3D models were run using 6 residual blocks and small minibatch sizes of one and two, to be able to efficiently run the models without running out of memory. The PSNR of the 3D models is greater than 2D models. The current challenge in the project is to visualize the 3D images to be able to confirm the superior performance of 3D models compared to the 2D models.

$$PSNR = 10log_{10}1MSE \tag{1}$$

$$SSIM(x,y) = (2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2) \quad (2)$$

Table 1: PSNR and SSIM metrics predicted from various CNN models.

Model	PSNR (dB)	SSIM
No Model	15.719 $\pm$0.207	0.242 $\pm$0.009
Feedforward CNN ResNet-6	19.151 $\pm$0.296	0.248 $\pm$0.009
Feedforward CNN ResNet-9	19.332 $\pm$0.285	0.248 $\pm$0.009
cGAN ResNet-6	17.894 $\pm$0.211	0.180 $\pm$0.011
cGAN ResNet-9	17.998 $\pm$0.227	0.183 $\pm$0.011
cGAN ResNet-9 (2x downsample)	18.784 $\pm$0.238	0.218 $\pm$0.010
cGAN ResNet-9 (4x downsample)	19.040 $\pm$0.253	0.237 $\pm$0.010
cGAN ResNet-9 (8x downsample)	18.105 $\pm$0.280	0.227 $\pm$0.009
Feedforward CNN ResNet-9 (1x downsample)	18.994 $\pm$0.351	0.188 $\pm$0.002
Feedforward CNN ResNet-9 (2x downsample)	19.330 $\pm$0.255	0.224 $\pm$0.007
Feedforward CNN ResNet-9 (4x downsample)	19.587 $\pm$0.266	0.228 $\pm$0.009
cGAN ResNet-6 (L1 Loss)	19.258 $\pm$0.293	0.217 $\pm$0.009
cGAN ResNet-6 (Wasserstein Loss)	18.657 $\pm$0.369	0.210 $\pm$0.007
cGAN ResNet-6 (Vanilla Loss)	15.957 $\pm$0.190	0.148 $\pm$0.010
3D Feedforward CNN ResNet-6 BatchSize-1	24.419 $\pm$0.318	0.351 $\pm$0.017
3D Feedforward CNN ResNet-6 BatchSize-2	24.061 $\pm$0.319	0.294 $\pm$0.010

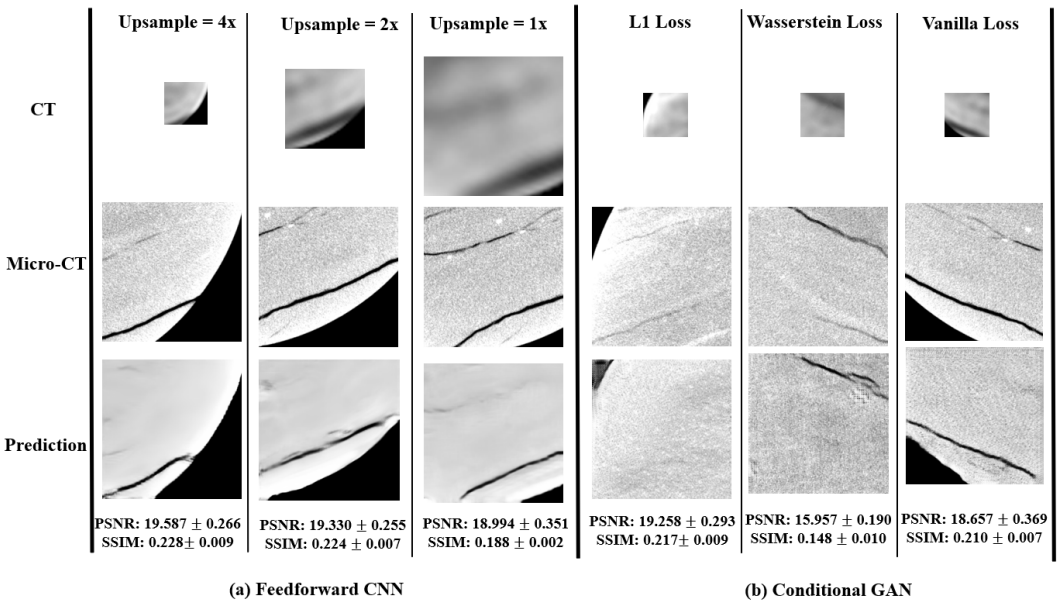

Figure 4: Preliminary results for 2D super resolution based on different hyperparameters: loss functions and upsample factors

5 Future Work

The 2D models are validated both qualitatively and quantitatively to be deployed for imaging applications. The 3D models are built based on the current 2D feedforward CNN and conditional-GAN models. While the quantitative metrics for 3D models are promising, we need to conduct more hyperparameter tuning, model training and visualize the super resolution on 3D volumes in order to validate the models before applying to real systems.

6 Acknowledgement

The author received meaningful input from Professor Anthony Kovscek and Timothy Anderson from the Department of Energy Resources, Stanford University.

References

[1] Richard S Middleton, J William Carey, Robert P Currier, Jeffrey D Hyman, Qinjun Kang, Satish Karra, Joaquín Jiménez-Martínez, Mark L Porter, and Hari S Viswanathan. Shale gas and non-aqueous fracturing fluids: Opportunities and challenges for supercritical co2. *Applied Energy*, 147:500–509, 2015.

[2] Karsten Pruess. Enhanced geothermal systems (egs) using co2 as working fluid—a novel approach for generating renewable energy with simultaneous sequestration of carbon. *Geothermics*, 35(4):351–367, 2006.

[3] Chao Dong, Chen Change Loy, Kaiming He, and Xiaoou Tang. Image super-resolution using deep convolutional networks. *IEEE transactions on pattern analysis and machine intelligence*, 38(2):295–307, 2015.

[4] Yukai Wang, Qizhi Teng, Xiaohai He, Junxi Feng, and Tingrong Zhang. Ct-image of rock samples super resolution using 3d convolutional neural network. *Computers & Geosciences*, 133:104314, 2019.

[5] Ying Da Wang, Ryan T Armstrong, and Peyman Mostaghimi. Enhancing resolution of digital rock images with super resolution convolutional neural networks. *Journal of Petroleum Science and Engineering*, 182:106261, 2019.

[6] Honggang Chen, Xiaohai He, Qizhi Teng, Raymond E Sheriff, Junxi Feng, and Shuhua Xiong. Super-resolution of real-world rock microcomputed tomography images using cycle-consistent generative adversarial networks. *Physical Review E*, 101(2):023305, 2020.

[7] Jiwon Kim, Jung Kwon Lee, and Kyoung Mu Lee. Accurate image super-resolution using very deep convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1646–1654, 2016.

[8] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.

[9] Christian Ledig, Lucas Theis, Ferenc Huszár, Jose Caballero, Andrew Cunningham, Alejandro Acosta, Andrew Aitken, Alykhan Tejani, Johannes Totz, Zehan Wang, et al. Photo-realistic single image super-resolution using a generative adversarial network. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4681–4690, 2017.

[10] Justin Johnson, Alexandre Alahi, and Li Fei-Fei. Perceptual losses for real-time style transfer and super-resolution. In *European conference on computer vision*, pages 694–711. Springer, 2016.

[11] Timothy I Anderson, Bolivia Vega, and Anthony R Kovscek. Multimodal imaging and machine learning to enhance microscope images of shale. *Computers & Geosciences*, 145:104593, 2020.

[12] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A Efros. Image-to-image translation with conditional adversarial networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1125–1134, 2017.