

Rock Image Enhancement Using Super-Resolution Neural Networks

Bo Gong* and Paul Duke

Chevron, 1500 Louisiana St, Houston, TX 77002, USA

Abstract. Pore-scale rock properties can be estimated from core images through image-based calculation or numerical simulation. However, accuracy of the estimation is directly limited by the resolution of the applied imaging technique, leading to uncertainties in interpretations. Imaging at an extremely high resolution can be time-consuming and expensive, and hence is usually done on selected samples, but the imaged area or volume is often not large enough to represent the formation heterogeneity. The method presented in this work helps resolve the trade-off between image scale and resolution in heterogeneous rocks. By training and applying a super-resolution convolutional neural network, pore-scale details can be effectively learned and reconstructed using relatively low-resolution images acquired in large rock volumes. Backscattered electron microscopy images and microCT images were acquired at multiple resolutions, and used for training and testing the models. Results showed that a properly trained model could increase image resolution by up to eight times, even when the training images appeared to be significantly different from the test images. Compared to bicubic interpolation techniques, the presented method resulted in more realistic visualization, as it better preserved the sharp edges of pore spaces.

1 Introduction

Rock imaging techniques, like X-ray computed tomography (CT), microtomography (microCT), scanning electron microscopy (SEM), and confocal microscopy, allow earth scientists to visualize and analyze rock samples in extremely fine detail. The ability to characterize micrometer- and nanometer-scale features, such as pore size, grain surface roughness, and mineral composition, is critical for understanding fluid flow behaviors in subsurface reservoirs, which in turn impacts reservoir production forecast and development decisions. With digital rock physics techniques, numerical simulations can be conducted on 2D or 3D images to study static and dynamic properties of the imaged rock systems.

Typically, the physical size of rock samples being imaged is directly linked to the resolution of the image: higher-resolution images are usually acquired on samples with smaller sizes, as they are relatively more expensive and time-consuming to acquire on large samples. Lower-resolution data (coarser images) are more available on larger volumes, which may be more statistically representative for the properties of interest, but accuracy of the resulted interpretations could be compromised because of the loss of fine details.

In heterogeneous rocks with complex pore systems, properties estimated with any image-processing-based methods could be especially sensitive to the image resolution. An example is demonstrated in **Fig. 1**. A backscattered scanning electron microscopy (BSEM) image of a carbonate rock sample was down-sampled by 2 times (2x), 4 times (4x), 8 times (8x) and 16 times (16x).

The original image has a resolution of 1 μm per pixel. The reduction of image quality by down-sampling can be seen from the comparison of **Fig. 1(a)** and **Fig. 1(b)**, where a section of the original (1x) image and the 16x down-sampled image are shown.

Pore size distribution was estimated on the original and all four down-sampled images using the maximum inscribed sphere method (modified for 2D images) described in [1]. As shown in **Fig. 1(c)**, pore size distribution is strongly impacted by the reduced resolution. The pore sizes estimated from the 8x and 16x down-sampled images failed to maintain the bi-modal distribution seen on the original image, whereas the 4x down-sampled image showed a shifted left peak, because small pores close to or below the degraded resolution were clustered into larger, artificial “pores”. This exercise suggests that even the same rock sample imaged at different resolutions could result in very different interpretations on pore space geometries, introducing high uncertainties into reservoir quality prediction. This effect was further discussed in [2] and [3].

Numerous methods have been explored in the past to improve the perceptual image quality [4]. Standard interpolation methods can reduce the pixelation effect on the low-resolution images, but they do not preserve high-frequency features very well and result in blurry images. Statistical methods like [5] and [6] aim to improve the preservation of sharp details, but often rely on image priors like gradient profiles and total variation, the extraction of which may be time-consuming.

To help resolve the trade-off between image scale and resolution in heterogeneous rocks, we present a new workflow to digitally enhance image resolution with a

* Corresponding author: bo.gong@chevron.com

minimal amount of acquisition. By training and applying a super-resolution convolutional neural network, high-resolution images can be reconstructed from low-resolution images, while preserving details like pore throats and angular corners of pore spaces. We first apply the method to synthetically generated low-resolution BSEM images, and then test it on manually acquired BSEM and microCT images. Besides using the traditional metric, means squared error (MSE), to compare the pixel-by-pixel similarity of the predicted and the target images [7], we also calculate pore size distribution to demonstrate the impact of the enhancement on petrophysical interpretations.

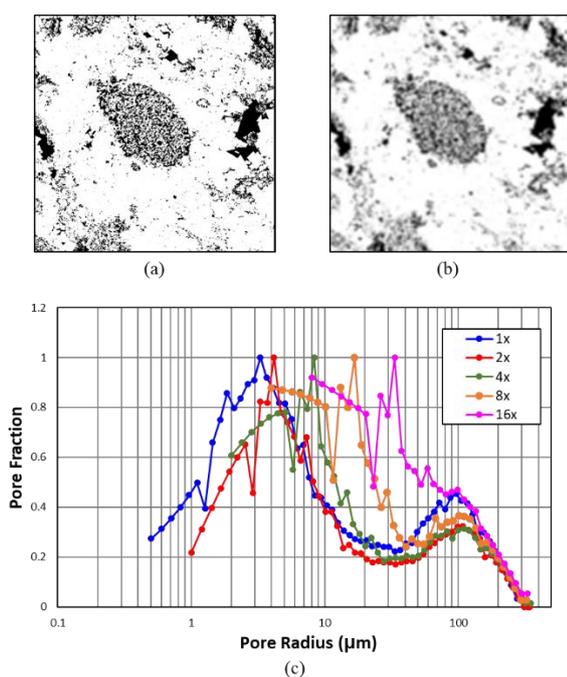


Fig. 1. Pore size distribution estimated from images of different resolutions show very different characteristics. (a) A section of the original image (1x); (b) The same section of the original image down-sampled by 16 times (16x). (c) Pore size distribution computed from the original image and four down-sampled images. Microporosity (left peak) is no longer captured on 8x and 16x images.

2 Image super-resolution using deep learning

Super-resolution refers to the process of reconstructing a high-resolution image based on one or more low-resolution images [4]. The objective is to recover details that have been lost during data compression, transmission, or low-resolution acquisition. Due to the non-unique nature of the problem, multiple solutions exist for the same low-resolution image. Therefore, the reconstruction requires the guidance of certain a-priori knowledge, which could be obtained either through feature extraction beforehand, or through example-based learning. Deep learning algorithms have made automatic learning possible, so a trained model could approximate the mapping from low-resolution images to high-resolution with minimal user intervention.

Convolutional neural network (CNN) is a type of deep-learning algorithms that are commonly used for image recognition and classification. Designed to handle 2D data, CNNs can effectively optimize the mapping from input images to output images by changing the weights between different layers of nodes (“filters”), based on the tasks they are trained on. Super-Resolution Convolutional Neural Network (SRCNN) [8] was specifically designed for image restoration using the basic architecture of CNNs. With low-resolution images paired with the corresponding high-resolution images, the model could learn the non-linear relationship between inputs and targets, and establish the rules to enrich image quality without creating anything inconsistent with the context. Information in the high-resolution images helps constrain the reconstructed rock texture, so that the enhanced images resemble the sampled rock facies. Different variants of SRCNN models were applied to synthetically down-sampled microCT images in [9], which showed high enhancement accuracy at a scale factor of 4.

U-Net is a CNN model type originally developed for medical image segmentation [10]. One challenge with plain CNNs is that certain information gets lost as it goes deeper into the network, e.g., the spatial relationship among different edges and corners is not maintained, so it may encounter issues with feature localization. Such problems are overcome in U-Net through the use of shortcut connections between corresponding layers, preserving spatial information as much as possible. As a result, U-Net is capable of delivering good performance with a relatively small training data set. This is especially beneficial for core image analysis, where data acquisition is usually costly. [11] shows that U-Net can be modified for image super-resolution with better performance than a simpler SRCNN.

In this study, we experimented with both SRCNN and U-Net for core image enhancement and compared their performances.

3 Super resolution on synthetically generated low-resolution images

3.1 Data preparation

We first used synthetically generated images to explore the enhancement limit, since arbitrary resolutions can be created for the training data.

Two BSEM images of a carbonate rock, S-01 and S-02, were selected for building and evaluating the models, as shown in **Fig. 2**. Each image has $32,768 \times 32,768$ pixels, with a resolution of $0.25 \mu\text{m}$. We used only S-01 to create the training set, and reserved S-02 as a blind test image.

S-01 was first down-sampled, or coarsened, to lower-resolution images. Four levels of low-resolution images (2x, 4x, 8x and 16x) were created as input images of the training data. Each low-resolution image was then up-sampled to create a blurred image which matches the number of pixels with the original high-resolution image using bicubic interpolation. Because of the large size of the BSEM images, the blurred images and the original image were sectioned into sub-images, each having

128×128 pixels. This process and the resulted sub-images are illustrated in **Fig. 3** and **Fig. 4**. Each level yielded 65,536 sub-images.

3.2 Model training and results

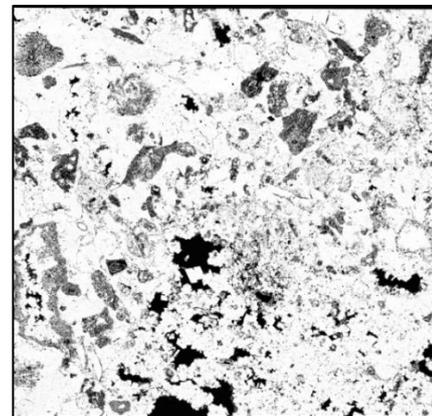
Two approaches were taken with the SRCNN method. In the first scenario, we built separate, fit-for-purpose models to enhance specific resolutions. Each model was exclusively trained on images of a certain resolution. Those models were then tested on the corresponding low-resolution images they were designed for. In the second scenario, a single model was trained with all levels of input images, this is referred to as the “mixed model”. In practice, a mixed model would be more desirable than separate models, because it offers flexibility in the resolution of input images. This is essential when images are acquired in different batches or stages, and the resolutions are inconsistent. With a mixed model, one does not need to train a new network each time when new images of a different resolution become available.

Considering the practicality and efficiency of the mixed-model approach, we trained two mixed models, one with SRCNN, the other with U-Net. The MSEs of the reconstructed images for the different models are plotted in **Fig. 5**. A bicubic baseline was also included for reference. With the architecture of SRCNN, although the separate models performed slightly better than the mixed models on each level, the latter produced promising results, with a much higher accuracy than bicubic interpolation. The U-Net mixed models outperformed the SRCNN separate models on all levels except for 2x, where they produced similar results.

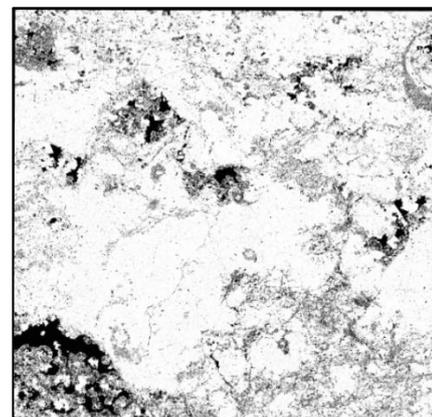
The advantage of the U-Net mixed model can also be seen in **Fig. 6**, where the 16x enhanced images from both mixed models are shown. Even though many of the details are lost beyond the recovery limit because of the high scale factor, the U-Net model seems to have better preserved the edges of the pore spaces than the SRCNN model (see features in the red boxes).

With the U-Net mixed model, we investigated the limit for image recovery with CNNs. **Fig. 7** shows the enhancement results of the U-Net mixed model from different levels of down-sampled images. For the 2x and 4x enhancement, the reconstructed images are very close to the original, with little details missing. On the 8x enhanced image, some small pores start to disappear (red

boxes) or cluster with nearby pore spaces (blue boxes), but most of the edges and corners of larger pore spaces are still well preserved. On the 16x enhanced image, pore spaces are further rounded, and the microporosity seen on the original image is visibly lost. An 8x enhancement appears to be the limit of reliable resolution recovery with CNNs, based on visual assessment. Beyond that level, small features like pore throats and angular pore spaces are smoothed out to an unacceptable degree, which would lead to biased interpretations.



(a)



(b)

Fig. 2. Two visually different BSEM images were selected for building and evaluating the models. (a) S-01, image used to create the training set; (b) S-02, image reserved for out-of-sample testing.

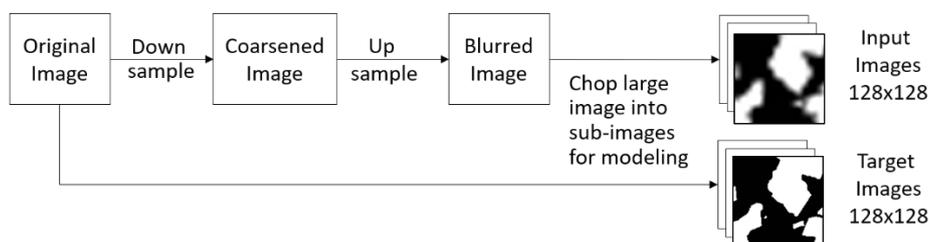


Fig. 3. Creating input and target images for the training set by degrading the original image.

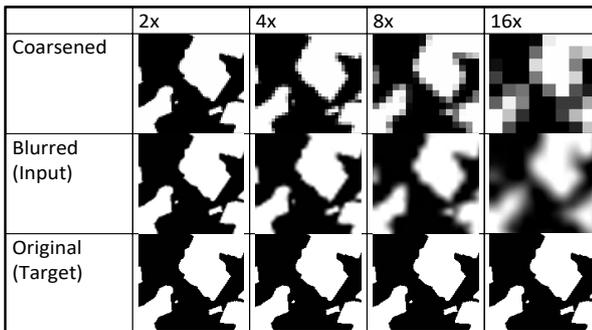


Fig. 4. Sub-images (128×128 pixels) created for model training.

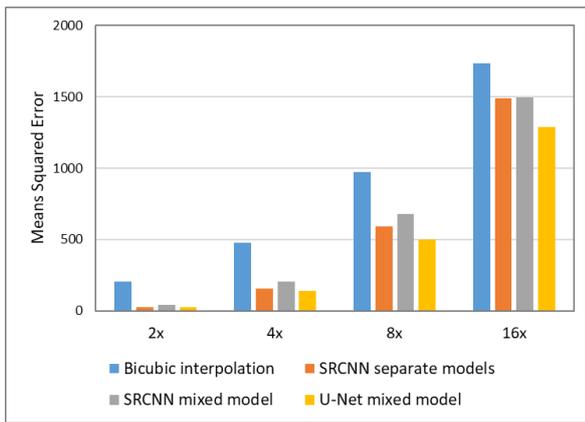


Fig. 5. Mean squared error (MSE) of enhanced images from different models, implemented on test image S-02.

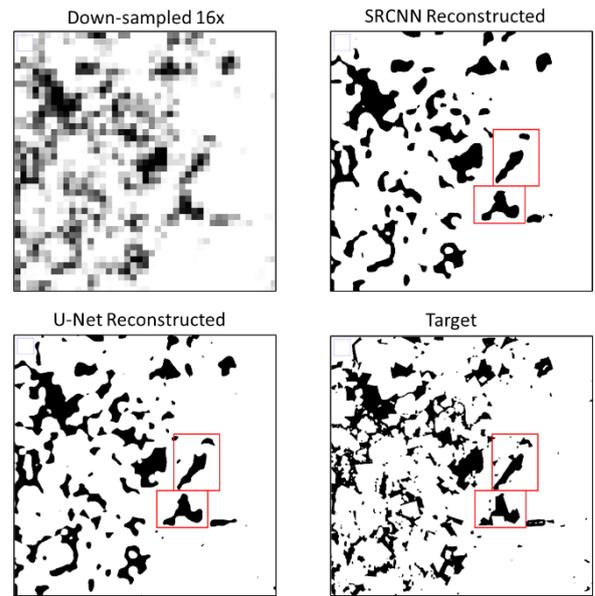


Fig. 6. Comparison of images enhanced with SRCNN and U-Net mixed models.

Besides the perceptual image quality, it is also important to examine the impact of the enhancement on the properties derived from the images, because the ultimate goal for core image enhancement is not pixel-by-pixel recreation, but the maximum preservation of the rock properties of interest. As an example of petrophysical properties, pore size distribution was calculated to evaluate the effectiveness of the feature reconstruction, using the method described in [1]. The results are plotted in (a)8.

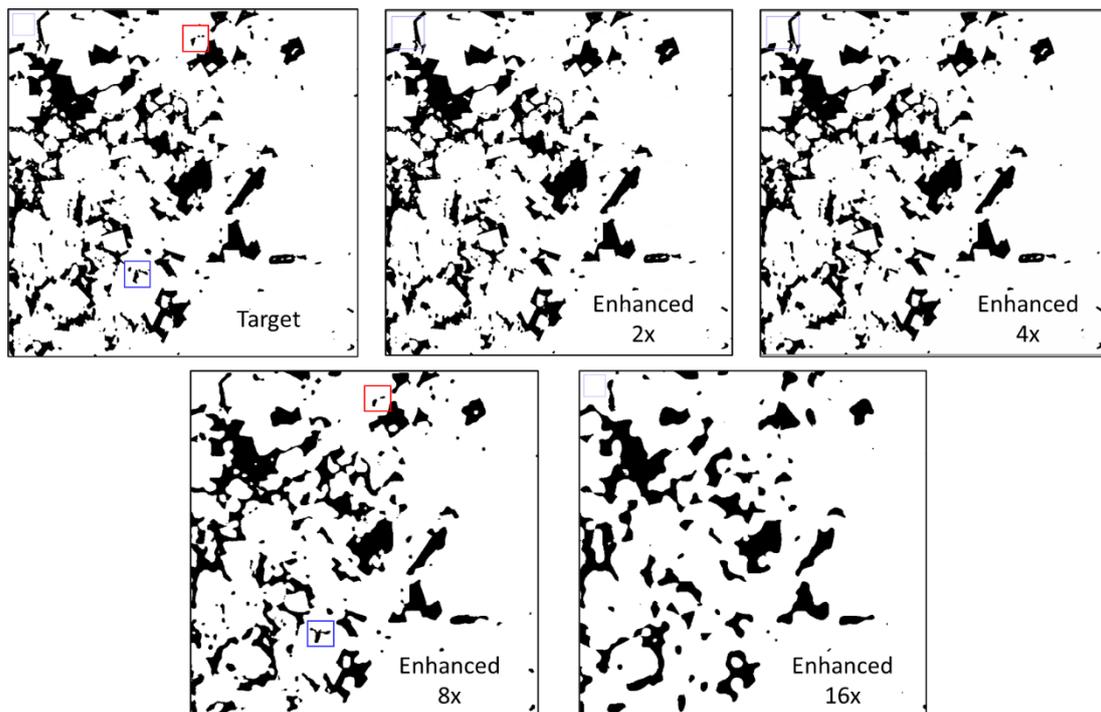


Fig. 7. A 170 $\mu\text{m} \times 170 \mu\text{m}$ section of target image, with the same area of four enhanced images at 2x, 4x, 8x and 16x. The enhanced images were reconstructed from down-sampled images with a U-Net mixed model.

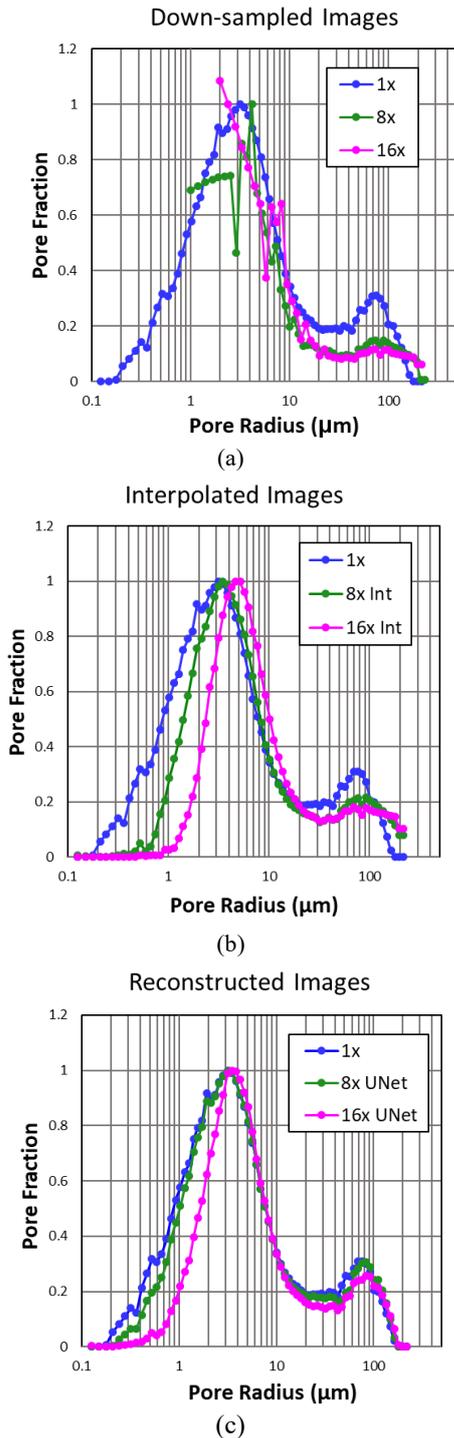


Fig. 8. Comparison of pore size distribution calculated with down-sampled, interpolated, and U-Net reconstructed images, compared with the original (1x). (a) Images down-sampled by 8x and 16x, and segmented by thresholding. (b) Images down-sampled by 8x and 16x, and interpolated to match the original image size. (c) Images down-sampled by 8x and 16x, and reconstructed with the U-Net mixed model.

(a) shows the pore size distributions calculated on two down-sampled images (8x and 16x), compared with the original (1x). Impacted by the image degradation, not only a large amount of the microporosity (left peak) was missing on the down-sampled images, but the secondary porosity (right peak) was also smoothed out due to the

resolution reduction. (b) shows the results calculated from images that were first down-sampled by 8x and 16x, and then interpolated to match the original image size. The 8x interpolated image recovered the location of the left peak, but failed to capture a large portion of the pores smaller than 2 μm (radius smaller than 1 μm). A larger amount of microporosity is lost on the 16x interpolated image. Moreover, the location of the left peak is shifted to the right, with the peak radius almost doubled. Both the 8x and 16x interpolated images resolved the right peak better than the down-sampled images, but a smoothed-out effect can still be seen.

In contrast, the U-Net reconstructed images showed a better performance, as shown in Fig. 8). The 8x enhanced image preserved the bi-modal pore size distribution very well, with only a small difference on the pores that have radius less than 1 μm. This indicates that the statistical impact on pore sizes is insignificant. The 16x enhanced image did not fully recover the left peak, but its location is almost identical to the original. Both the 8x and the 16x enhanced images produced the right peak with a good resolution.

For core image interpreters, it is essential to determine what level of recovery is truly necessary, since different applications call for different precisions. In this example, if it is known that pores smaller than 2 μm play a critical role in the studied rock, the 8x enhancement would be preferable to the 16x. However, if a task only requires the identification of the local maxima, a 16x enhancement would be sufficient. These varying criteria to judge the enhancement performance distinguish the image super-resolution applications in digital rock physics from some of the other fields, which may have lower tolerance for errors.

4 Super resolution on manually acquired low-resolution images

4.1 Application on BSEM images

We then tested the workflow with BSEM images manually acquired at different levels.

Two levels of BSEM images were acquired on multiple carbonate rock samples, with one resolution roughly 5 times as high as the other. For each low-resolution image, a high-resolution image was positioned in its centre for the ease of registration, as shown in Fig. 9. Consequently, the high-resolution images covered a much smaller area of the rock samples. Because of the limitation of the imaging speed, it would not be practical to scan the entire surface of the cores at the high resolution.

Images of the two resolutions were then registered with an internally developed software. The low-resolution images were trimmed to match the areas covered by the high-resolution patches. Once aligned, each image was divided into sub-images with 128x128 pixels, and pairs of low-resolution/high-resolution sub-images were used as training data, containing 86,400 examples. Images from different rock samples outside of the training data were reserved as a testing dataset of the same size.

* Corresponding author: bo.gong@chevron.com

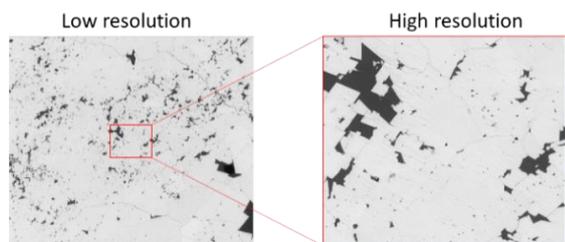


Fig. 9. BSEM images acquired at two resolutions of roughly 5 times difference. The high-resolution image was acquired in the central section of the low-resolution image.

A new U-Net-based model was trained and applied to enhance the low-resolution BSEM images. An example of the results is shown in **Fig. 10**. The enhancement effectively sharpened the edges of the pore spaces, and preserved small pores and pore throats well. A reduction in noise was also seen as a by-product of the enhancement. The White grain boundaries seen on the high-resolution images were not captured by the enhancement, likely due to their low color contrast compared with the background.

Fig. 6 shows the pore size distribution calculated on an image before and after a 5x enhancement. A bi-modal distribution can be clearly seen from the reconstructed image, but was not captured by the low-resolution image due to the resolution limit. The existence of the microporosity peak was further confirmed by mercury injection capillary pressure (MICP) data, measured on the same rock sample (**Fig. 12**).

This exercise confirms that the presented workflow not only applies to synthetically generated low-resolution/high-resolution image pairs, but can also be trained with images actually acquired at two different resolutions.

4.2 Application on microCT images

The same workflow was trained with and applied to microCT images on a slice-by-slice basis. Two levels of

images were acquired on the same rock sample at 22.5 μm per voxel and 6.6 μm per voxel. The training was done in a similar manner to the BSEM super-resolution model, with low-resolution and high-resolution images registered and paired. The training dataset contains about 135,000 samples, and the testing dataset has about 20,000. All the sub-images are 128x128 pixels.

The enhancement results are shown in **Fig. 8**. Compared with the high-resolution image, many details were recovered from the low-resolution image. Noise level was also suppressed during the enhancement.

5 Conclusion

Super-resolution convolutional neural networks provide a low-cost solution to increase the accuracy of image-based rock property evaluation with a minimal amount of high-resolution image acquisition. With images digitally enhanced on samples of larger sizes, interpretations would be more representative of the rock heterogeneity, which helps resolve the incompatibility between imaging resolution and scale.

With synthetically generated images, we demonstrated that U-Net based models are more suitable for the super-resolution task than traditional SRCNNs. One could train a single mixed model with images of multiple resolutions instead of multiple fit-for-purpose models, since the sacrifice in performance is insignificant.

Results have shown that pore structures can be well preserved for images coarsened by up to 8x, whereas local maxima of the pore size distribution could be correctly identified from a 16x enhancement. We also confirmed that the presented workflow is equally applicable to training data consisting exclusively of manually acquired images, including BSEM and microCT images. The enhanced image quality could be beneficial for any following digital rock workflows, leading to more reliable observations of pore-scale rock properties and less uncertainties in rock texture analysis.

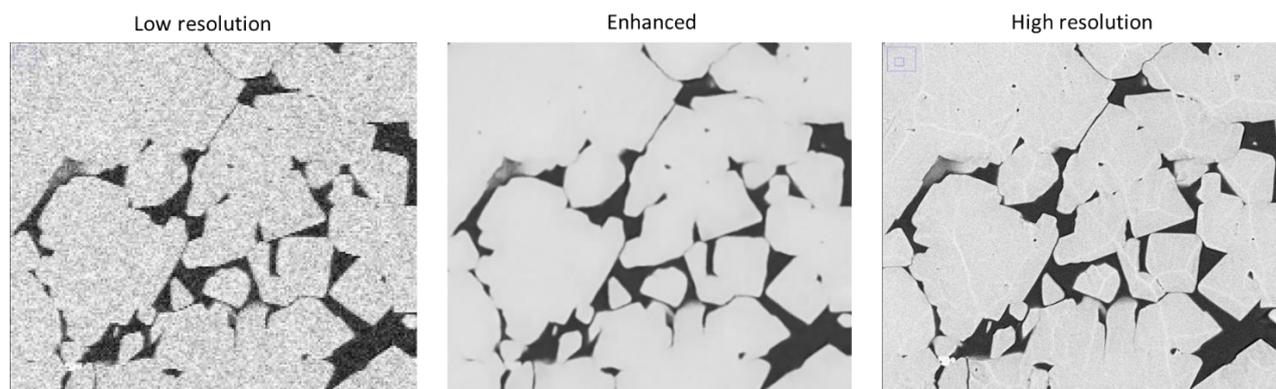


Fig. 5. Comparison of an enhanced image (middle) with manually acquired low-resolution and high-resolution BSEM images.

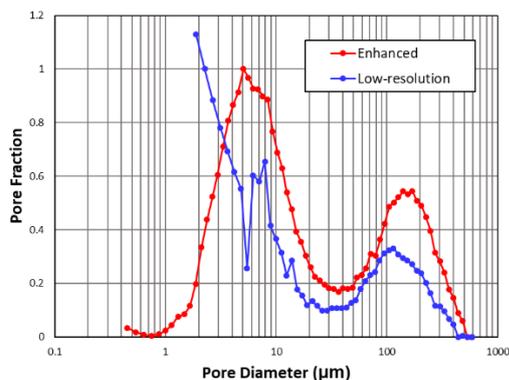


Fig. 6. Pore size distribution of a low-resolution image and its enhanced image. The enhanced image exhibited a bi-modal distribution.

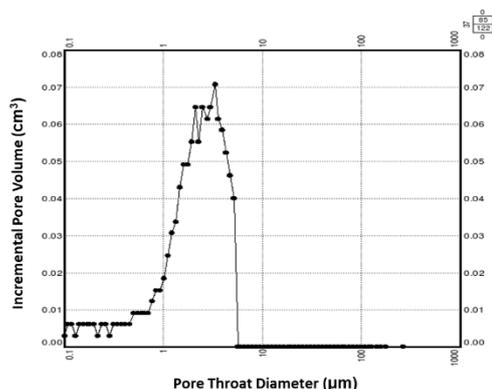


Fig. 7. Mercury injection capillary pressure (MICP) data confirms the presence of the microporosity revealed from enhanced image.

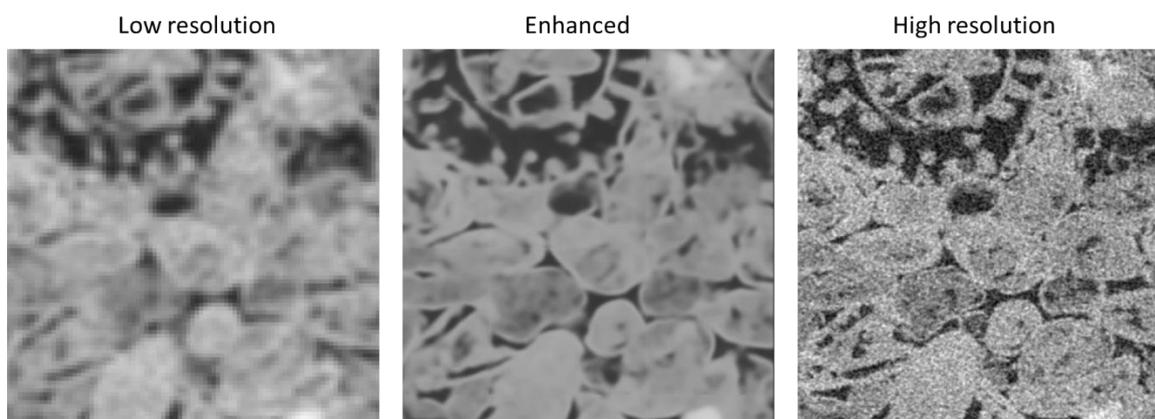


Fig. 8. Comparison of an enhanced image (middle) with manually acquired low-resolution and high-resolution microCT images.

References

1. D. Silin, & T. Patzek, Pore space morphology analysis using maximal inscribed spheres. *Physica A: Statistical Mechanics and its Applications*, **371**, 336-360 (2006)
2. Y. Bazaikin, B. Gurevich, S. Iglauer, T. Khachkova, D. Kolyukhin, M. Lebedev, . . . G. Reshetova, Effect of CT image size and resolution on the accuracy of rock property estimates. *Journal of Geophysical Research: Solid Earth*, **122**(5), 3635-3647 (2017)
3. N. Alyafei, A. Q. Raeini, A. Paluszny, & M. J. Blunt, A Sensitivity Study of the Effect of Image Resolution on Predicted Petrophysical Properties. *Transport in Porous Media*, **110**(1), 157-169 (2015)
4. W. Yang, X. Zhang, Y. Tian, W. Wang, & J. Xue, Deep Learning for Single Image Super-Resolution: A Brief Review. *arXiv:1808.03344 [cs.CV]* (2019)
5. A. Marquina, & S. Osher, Image super-resolution by tv-regularization and bregman iteration. *J. Sci. Comput.*, **37**(3), 367-382 (2008)
6. J. Sun, Z. Xu, & H. Shum. Image super-resolution using gradient profile prior. *2008 IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (2008)
7. Z. Wang, A. C. Bovik, H. R. Sheikh & E. P. Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*, **13**(4), 600-612 (2004)
8. C. Dong, C. C. Loy, K. He, & X. Tang, Image Super-Resolution Using Deep Convolutional Networks. *arXiv:1501.00092 [cs.CV]* (2015)
9. Y. D. Wang, R. T. Armstrong, & P. Mostaghimi, Enhancing Resolution of Digital Rock Images with Super Resolution Convolutional Neural Networks. *Journal of Petroleum Science and Engineering*, **182**, 106261 (2019)
10. O. Ronneberger, P. Fischer, & T. Brox, U-Net: Convolutional Networks for Biomedical Image Segmentation. *arXiv:1505.04597v1 [cs.CV]* (2015)
11. Z. Lu, & Y. Chen, Single Image Super Resolution based on a Modified U-net with Mixed Gradient Loss. *arXiv:1911.09428 [eess.IV]* (2019)