Restoration of Material Pore Structure Image Using Transformer Architecture

Jianwei Pan[1], Yi Yin[1*], Yuanbing Li[2], Shujing Li[2], Wei Wang[1], Zhen Cai[2], Xin Xu[1]
[1]School of Compter Science and Technology, Wuhan University of Science and Technology, Wuhan, PR China
[2]The State Key Laboratory of Refractories and Metallurgy, Wuhan University of Science and Technology, Wuhan, PR China
{yinyi, panjianwei, wangwei8, xuxin}@wust.edu.cn, lybref2002@126.com, lsjwust@163.com, 2456107293@qq.com

***Abstract*—The latest Transformer architecture has shown significant progress in the field of image restoration. However, research on the application of the Transformer architecture in material scanning electron microscopy (SEM) image restoration is still lacking. Material SEM images are typically influenced by factors such as noise, exposure, and distortion, resulting in low-resolution (LR) image and clarity. The self-attention mechanism and global information interaction capabilities of the Transformer architecture make it perform well in image restoration tasks. This paper uses a Transformer-based algorithm to study the restoration of sandstone SEM images. The experimental results demonstrate that the Transformer-based algorithm is capable of restore the details and edge information of the sandstone SEM images, improve their resolution and clarity, and enhance the overall image quality. Moreover, the algorithm is able to learn the distribution of the images. The PSNR and SSIM of this method reach a maximum of 34.59dB and 0.8893, which are improved compared to other SOTA algorithms.**

***Index Terms*—Image Restoration, Transformer Architecture, Sandstone SEM Image, Pore Structure**

I. Introduction

Image restoration refers to the process of repairing damaged or degraded images in order to restore the quality and characteristics of the original images as closely as possible [1]. These images may have been affected by noise, blurriness, distortion, or other forms of damage. The goal of image restoration techniques is to enhance the clarity and details of images through methods such as denoising, deblurring, and image super-resolution (SR), making them appear more realistic and accurate [2]. Currently, image restoration is typically implemented using deep learning models, which analyze and repair specific attributes of the images to improve their quality and usability. Applications of image restoration techniques include denoising, deblurring, image SR enhancement, and color correction, among others [3].

At present, there have been studies applying deep learning methods such as CNN [4, 5] and SRGAN [6, 7] to the super-resolution imaging of material porosity structures. By using scanning electron microscopy (SEM) or other microscopy techniques to obtain images, an appropriate deep learning architecture is selected to learn the features and structures of the material images for restoration. Image restoration techniques are used to enhance the contrast and details of the images, making the crystal grain boundaries and grain boundaries more clearly visible, helping researchers to more accurately analyze the microscopic structures of materials and provide more precise data support for materials design and development [8, 9]. Lukas Mosser and colleagues reconstructed the three-dimensional digital images of material's porous structure based on GANs. They applied the training and image generation process based on GANs to non-segmented grayscale datasets of micrometer-scale computed tomography (μCT) of a chalky limestone [10]. Yuzhu Wang and colleagues have proposed a method for reconstructing high-resolution (HR) porous structures using a convolutional neural network based on LR μCT images and HR SEM images. They applied numerical reconstruction methods to enhance the resolution of μCT images [11]. Ramin Bostanabad introduced an efficient and novel method based on transfer learning to achieve extrapolation reconstruction of virtual three-dimensional samples from two-dimensional images. This method is widely applicable to microstructures, including alloys, porous media, and polycrystals. Viewing the reconstruction task as an optimization problem, random three-dimensional images are iteratively refined to match their microstructural features with canonical features [12]. Reza Shams and the team have developed an adaptive framework that effectively reconstructs heterogeneous porous media using conditional GANs [13]. Hua Zhang and colleagues have developed an efficient and accurate method for analyzing the pore structure of permeable concrete based on deep learning. They utilized permeable concrete CT slices as a dataset and employed an improved Mask R-CNN algorithm as the model training framework to train a pore identification model. The results indicate that the improved algorithm shows better convergence and higher accuracy in pore segmentation and identification compared to Mask R-CNN [14]. Bo Yuan and the team argue that traditional semantic segmentation networks require HR input images to improve performance, and can only generate masks that are of the same resolution as the input images. Moreover, these traditional networks demand high-performance devices, leading to increased computational and equipment costs. In response, they have proposed a method that utilizes a SR

*Corresponding author
This work was supported by National Natural Science Foundation of China (Project No. 62376201), and the Open Fund of the State Key Laboratory of Refractories and Metallurgy (Wuhan University of Science and Technology) (Project No. G202410).

979-8-3503-5409-6/24/$31.00 ©2024 IEEE
DOI 10.1109/CAI59869.2024.00216

neural network to enhance the recognition of digital rock images [15]. Yinghan Zhao and the team have conducted research utilizing a variational autoencoder neural network model to characterize the 3D structural information of porous materials and represent them using low-dimensional latent variables. They have also applied a structure-linkage combined with Bayesian optimization method to model the structure-performance relationship and address inverse problems in the process [16]. Our team also accumulated relevant knowledge and data on electron microscopy images of foam ceramics in the early stage [17-22].

It is show that, for the restoration of SEM images, the application of image SR methods aims to reconstruct high-quality images that have the same distribution as the original low-quality images from degraded low-quality images.

However, it can also be seen that material images often need expensive experimental equipment and personnel with professional knowledge to obtain, the available data sets are relatively small. At the same time, the pore structure of materials is often very complex, including pores of various sizes and shapes, and pore structure images may be affected by noise and artifacts, these uncertainties can lead to erroneous estimations and generalization ability in the restoration process of the model[23]. But the current methods are not yet sufficient to effectively solve these problems.

At present, methods based on Transformer architecture are increasingly used in restoring traditional images and have achieved good results [24]. However, there is no application of Transformer architecture for material electron microscopy image restoration.

II. Dataset and Methodology

A. Dataset

This paper use Bentheimer Sandstone electron microscopy images from the DRSRD1_2D dataset[26]. In this dataset, 800×800 HR unsegmented Bentheimer sandstone images are selected. These images have been shuffled and divided into training, validation, and test sets in a ratio of 8:1:1. This paper uses average pooling and bicubic interpolation methods to downsample the image by 2X and 4X respectively, forming a LR image. Then, image SR reconstruction is performed to increase the details and clarity of the image during the restoration process, and to ensure that the structural distribution of the image is consistent with the original image. As shown in Fig. 1.

The reason for choosing this dataset for training is as follows: (1) Dataset format. The organization of this dataset is similar to the DIV2K dataset, and it is the first dataset used for benchmarking SR algorithms on digital rock images. It has been utilized to test several SR convolutional neural networks (SRCNN)[27] so far. This indicates that the dataset has the same file organization structure as some mainstream SR algorithms and has been effectively validated by the first SRCNN algorithmused for SR. (2) Dataset scale. The dataset consists of 800×800 undivided HR images of Bentheimer sandstone (3.8 micrometers). The training set contains 800 images each of quality and accuracy of the restored image. The research demonstrates the effectiveness and advantages of the CNN+ Transformer architecture-based SR imaging method in image restoration of material SEM images, and it is Bentheimer Sandstone. (3) Dataset variety. The samples in the dataset are diverse, representing both simple resolved sandstone images and complex underresolved images of carbonate microporosity. The objective of this dataset is to analyze pore shape parameter distribution characteristics based on the real microstructure of digital rock cores. Additionally, the good transferability based on the Transformer architecture aligns with the purpose of this paper, which applies the Transformer structure for image restoration. Overall, the organization, scale, and diversity of the dataset, along with its alignment with the objectives of the study and the transferability of the CNN+Transformer architecture, make it a suitable choice for training the model.

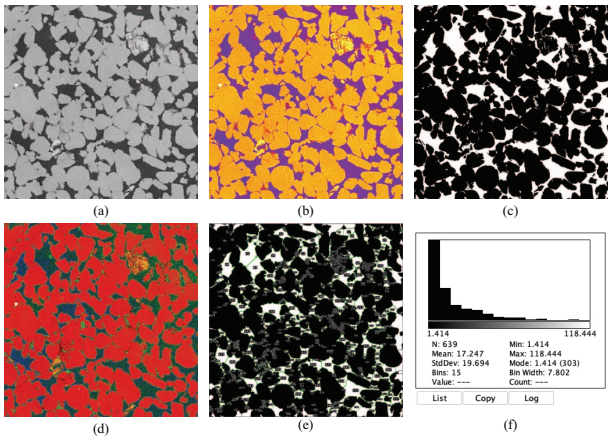

Fig. 1. Sandstone SEM image and microstructure feature. (a) Original image. (b) pore pattern. The orange regions are grains, while the purple regions are pores. (c) Binary processing of grains and pores. (d) Grain and pore separation. The red regions are grains, the other part are pores. The darker the pore means larger pore depths. (e) Number of pores. (f) Pore size distribution.

B. Method

ESRT is the first attempt to combine lightweight CNN and Transformer for image SR direction, namely the combination of Lightweight CNN Backbone (LCB) and Lightweight Transformer Backbone (LTB) [28]. LCB is able to dynamically adjust the size of the feature maps to extract deep features with low computational cost. LTB is composed of a series of Transformer modules and combined with efficient multi-head attention to reduce memory usage. The algorithm applied to LR SEM images is called ES-DR.

C. Data downsampling

The experiment in this paper uses average pooling and bicubic interpolation methods to downsample the image by 2X and 4X respectively, forming a LR image. LR images may lose some detailed information in the original image,

resulting in a blurry effect because it combines multiple pixels from the original image into a new pixel. Average pooling downsampling can result in the loss of some detailed information when taking the average value. Due to the average pixel values in each region, the image resolution may decrease, which may result in information loss and smoothing effects in the details. Average pooling downsampling results in a smoothing effect on the image. Taking the average of the pixels within a region blurs the details of the image, making the image look smoother. Bicubic interpolation downsampling aims to preserve more fine details during the downsampling process. It uses interpolation techniques to estimate pixels, taking into account the values of surrounding pixels, and calculates the weighted average of new pixel values. However, details such as noise, edges, and textures in the image may interfere with the interpolation process, resulting in artifacts in the interpolated results.

III. Results and Discussion

In this section, the performance of images generated by the ES-DR algorithm and state-of-the-art methods is compared, and the representation of the pore structure in the image is tested.

A. Performance

According to Table 1, no matter what downsampling method and multiple are used, ES-DR obtains higher PSNR and SSIM compared to the other four algorithms. Evaluation indicators of five algorithms.

TABLE I
Quantitative comparison of five SR algorithm.

Method	Downsampling	Scale	Sandstone	
			PSNR(dB)	SSIM
Bicubic interpolation[29]	average pooling	×2	32.23	0.8613
SRCNN[27]			32.51	0.8346
SRGAN[30]			33.28	0.8660
IMDN[31]			34.58	0.8890
ES-DR			34.59	0.8892
Bicubic interpolation	Bicubic interpolation		32.23	0.8612
SRCNN			32.73	0.8487
SRGAN			33.08	0.8635
IMDN			34.59	0.8901
ES-DR			34.59	0.8893
Bicubic interpolation	average pooling	×4	26.12	0.6517
SRCNN			27.82	0.6962
SRGAN			28.58	0.6931
IMDN			29.70	0.7308
ES-DR			29.78	0.7322
Bicubic interpolation	Bicubic interpolation		26.12	0.6518
SRCNN			27.88	0.6988
SRGAN			28.58	0.6947
IMDN			29.71	0.7312
ES-DR			29.79	0.7323

It can be seen that almost all SR images obtained using ES-DR are significantly improved, especially when the initial value is twice the downsampling, the improvement effect is most obvious. This also shows that the distribution of the image generated by ES-DR is closest to the original image, illustrating the ability of ES-DR to accurately restore the details and structures of HR images.

B. Visual effect comparison

As shown in Fig. 2, it can be seen that regardless of whether using average pooling or Bicubic interpolation for down-sampling, the SR images generated by bicubic interpolation, SRCNN, and SRGAN are blurrier. In the down-sampling mode of average pooling, the image generated by IMDN is relatively clear, but the edges are obviously sharpened. At the same time, compared with the original slice, although the pore structure in the generated images is clearer, the interior of the grains is obviously blurred. Moreover, under the downsampling method using Bicubic interpolation, the images generated by IMDN are blurrier than those generated using average pooling, which shows that the universality of image generated by IMDN is not good. It can also be seen that ES-DR is closer to the original slice image in visual presentation. By comparing the enlarged portions of the images, it can be observed that the visually appealing results obtained by ES-DR outperformed the other four algorithms in terms of image restoration. Therefore, ES-DR is more effective in improving the accuracy of restoring pore structure. This also indicates that the ES-DR algorithm is able to learn the distribution of sandstone pore structure more accurately.

C. Microstructure feature analysis

Accurate material image structure models are crucial for optimizing material performance, preparation control, and functional design. The model needs to accurately represent the pore space, including the number of pores, pore size distribution, pore shape, circumference and roundness for evaluation. By referencing the original images in the dataset, compare multiple representation values obtained from five algorithms with the original images, and analyze the accuracy of the algorithms.

The SR image obtained by selecting the same algorithm for downsampling using different methods and multiples is compared with the original image. The results of the number of pores, pore area, pore perimeter and circularity are shown in Fig. 3. Specifically, there are significant differences between Bicubic interpolation, SRCNN, and SRGAN compared to IMDN and ES-DR. The representation values of the IMDN and ES-DR are significantly closer to the original values. Specifically, there is not much difference between IMDN and ES-DR in terms of number of pores, pore size distribution and pore shape. However, in terms of circularity, especially in the results of double SR images, the representation values of the SR images obtained by ES-DR are closer to the original image. On the other hand, the IMDN algorithm shows better stability, with relatively consistent average values.

In Fig. 4, a more detailed comparison is made between ES-DR and IMDN in terms of the similarity of four representation values (number of pores, pore perimeter, pore circularity, and feret diameter) with the original image under the initial condition of bicubic interpolation with double downsampling. It can also be seen that ES-DR is closer to the original in these representation value than IMDN, which reflects that ES-

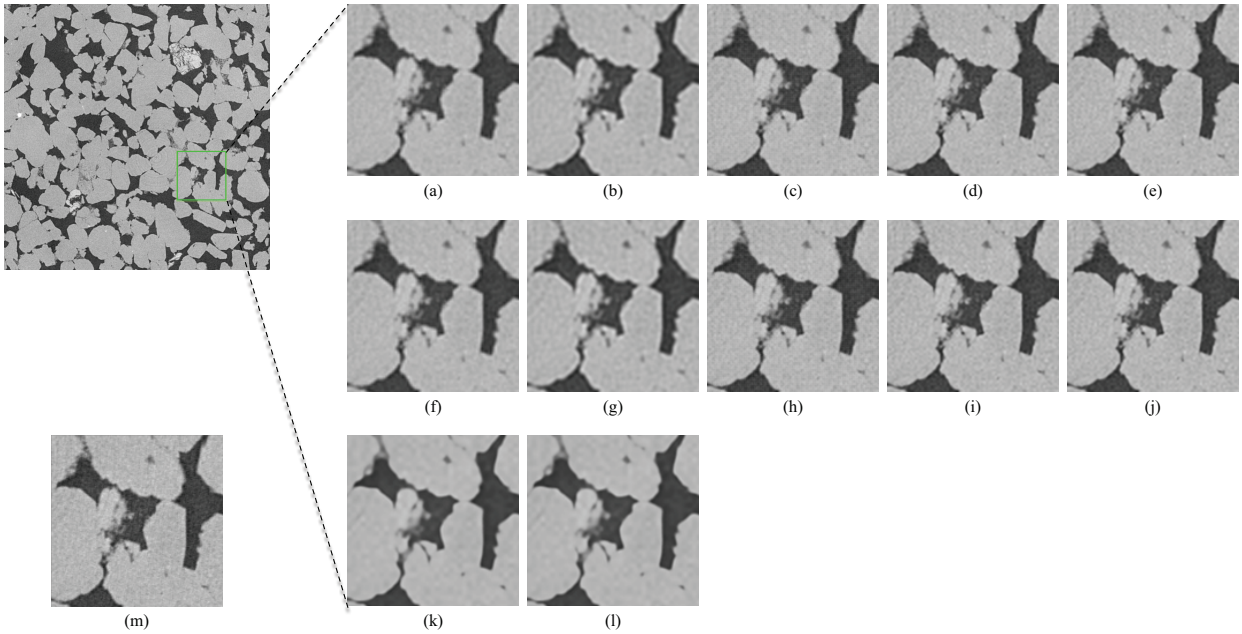

Fig. 2. Enlarged view of the sliced part of the original image obtained using five SR algorithms. (a) - (e) and (f) - (j) are respectively enlarged views of slices super-resolved by Bicubic interpolation, SRCNN, SRGAN, IMDN and ES-DR after average pooling and bicubic interpolation downsampling. (a) - (j) are all initialized double downsampling. (k) and (l) are the SR image slices of ES-DR under four times down sampling initialization of average pooling and bicubic interpolation respectively. (m) is the original slice. The slice image is 150*150 pixels.

DR is better able to learn the distribution characteristics of microstructural features in the original image.

D. Conclusion

Deep learning methods, especially CNN and GAN, have shown great potential in the SR of the pore structure in material images, and have achieved some impressive results. On the other hand, the latest Transformer architecture has made significant progress in the field of natural language processing and image processing, particularly in image restoration. Its self-attention mechanism and global information interaction capabilities make it perform exceptionally well in handling image tasks, especially in extracting global features and reconstructing details.

However, due to data limitations, computational complexity and the dominance of traditional methods, the research and application of Transformer architecture in the field of material image restoration have not been fully explored. In the field of material image restoration, the Transformer architecture has advantages such as modeling long-range dependencies, parallel computation, flexibility, scalability, and multimodal processing. These advantages provide a new and promising solution for material image restoration tasks.

This paper applies the SR imaging method based on CNN+Transformer architecture to the image restoration of materials SEM images. By experiments and analysis, it is demonstrated that this method can improve the resolution and details of the image, improve the visualization of the material's microstructure, and show the advantage of learning the distribution information of the original image. Next, with the effectiveness and advantages of this image restoration method, it is expected to provide optimization and guidance in the design and preparation of the "microstructure-performance" of materials. By restoring the microstructure of materials more clearly and accurately, researchers and engineers can better understand the properties and behavior of materials, and make more accurate adjustments and optimizations in the design and preparation process.

REFERENCES

[1] Banham M R, Katsaggelos A K. "Digital image restoration". IEEE signal processing magazine, 1997, 14(2): 24-41.
[2] Lehtinen J, Munkberg J, Hasselgren J, et al. "Noise2Noise: Learning image restoration without clean data". arXiv preprint arXiv:1803.04189, 2018.
[3] Zamir S W, Arora A, Khan S, et al. "Multi-stage progressive image restoration". Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2021: 14821-14831.
[4] LeCun Y, Bottou L, Bengio Y, et al. "Gradient-based learning applied to document recognition". Proceedings of the IEEE, 1998, 86(11): 2278-2324.
[5] Dong C, Loy C C, He K, et al. "Image super-resolution using deep convolutional networks". IEEE transactions on pattern analysis and machine intelligence, 2015, 38(2): 295-307.
[6] Goodfellow I, Pouget-Abadie J, Mirza M, et al. "Generative adversarial nets". Advances in neural information processing systems, 2014, 27.
[7] Ledig C, Theis L, Huszár F, et al. "Photo-realistic single image super-resolution using a generative adversarial network". Proceedings of the IEEE conference on computer vision and pattern recognition. 2017: 4681-4690.

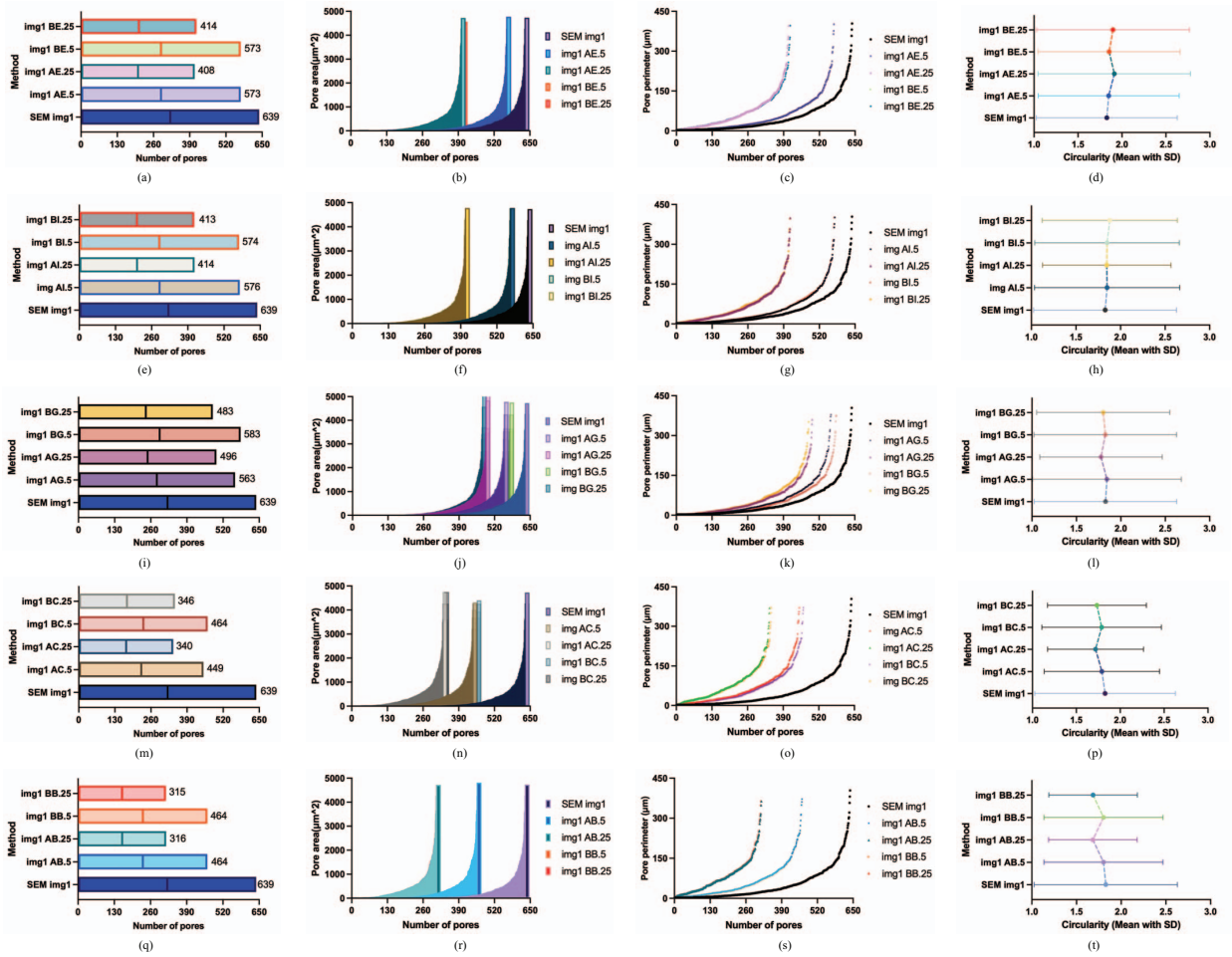

Fig. 3. Compare the feature values of the same algorithm with the original image under different downsampling conditions. A: average pooling; B: Bicubic interpolation; E: ES-DR; I: IMDN; G: SRGAN; C: SRCNN; .5: Double downsampling; .25: Quadruple downsampling.

[8] A. Jain, S. Ong, G. Hautier, W. Chen, W. Richards, S. Dacek, S. Cholia, D. Gunter, D. Skinner, G. Ceder, K. Persson, "Commentary: The materials project: A materials genome approach to accelerating materials innovation". APL Mater. 1 (2013), 011002.

[9] Zhou Y T, Chellappa R, Vaid A, et al. "Image restoration using a neural network". IEEE transactions on acoustics, speech, and signal processing, 1988, 36(7): 1141-1151.

[10] Mosser L, Dubrule O, Blunt M J. "Stochastic reconstruction of an oolitic limestone by generative adversarial networks". Transport in Porous Media, 2018, 125(1): 81-103.

[11] Wang Y, Arns C H, Rahman S S, et al. "Porous structure reconstruction using convolutional neural networks". Mathematical Geosciences, 2018, 50: 781-799.

[12] Bostanabad R. "Reconstruction of 3D microstructures from 2D images via transfer learning". Computer-Aided Design, 2020, 128: 102906.

[13] Shams R, Masihi M, Boozarjomehry R B, et al. "A hybrid of statistical and conditional generative adversarial neural network approaches for reconstruction of 3D porous media (ST-CGAN)". Advances in Water Resources, 2021, 158: 104064.

[14] Zhang H, Zhang R, Sun D, et al. "Analyzing the pore structure of pervious concrete based on the deep learning framework of Mask R-CNN". Construction and Building Materials, 2022, 318: 125987.

[15] Yuan B, Li H, Du Q. "Enhancing identification of digital rock images using super-resolution deep neural network". Geoenergy Science and Engineering, 2023, 229: 212130.

[16] Zhao Y, Altschuh P, Santoki J, et al. "Characterization of porous membranes using artificial neural networks". Acta Materialia, 2023, 253: 118922.

[17] Li Y, Liu J, Yin B, et al. "Novel fibrous/nano-Al2O3 insulation composites produced using sol-gel impregnation for energy-saving". Journal of Sol-Gel Science and Technology, 2023: 1-10.

[18] Chen P, Li Y, Li S, et al. "Thermally insulating Bi6B10+2xO24+3x (x=−2, −1, 0, 1, 2) ceramics with neutron/gamma shielding capability". Journal of the Australian Ceramic Society, 2023, 59(2): 291-301.

[19] Liu J, Li Y, Yin B, et al. "Thermally insulating magnesium borate foams with controllable structures". Ceramics International, 2022, 48(17): 25506-25512.

[20] Liu J, Li Y, Yin B, et al. "Novel magnesium borate ceramic matrix composites with glass fiber reinforcement". Ceramics International, 2023, 49(7): 11197-11203.

[21] Chen P, Li Y, Yin B, et al. "New design of bismuth borate ceramic/epoxy composites with excellent fracture toughness and radiation shielding capabilities". Materials Today Communications, 2023, 35: 106102.

[22] Li X, Li S, Li Y, et al. "Fabrication of novel BPO4 ceramic foams using the combination of the direct foaming method and freeze-drying techniques". International Journal of Applied Ceramic Technology, 2023, 20(6): 3565-3575.

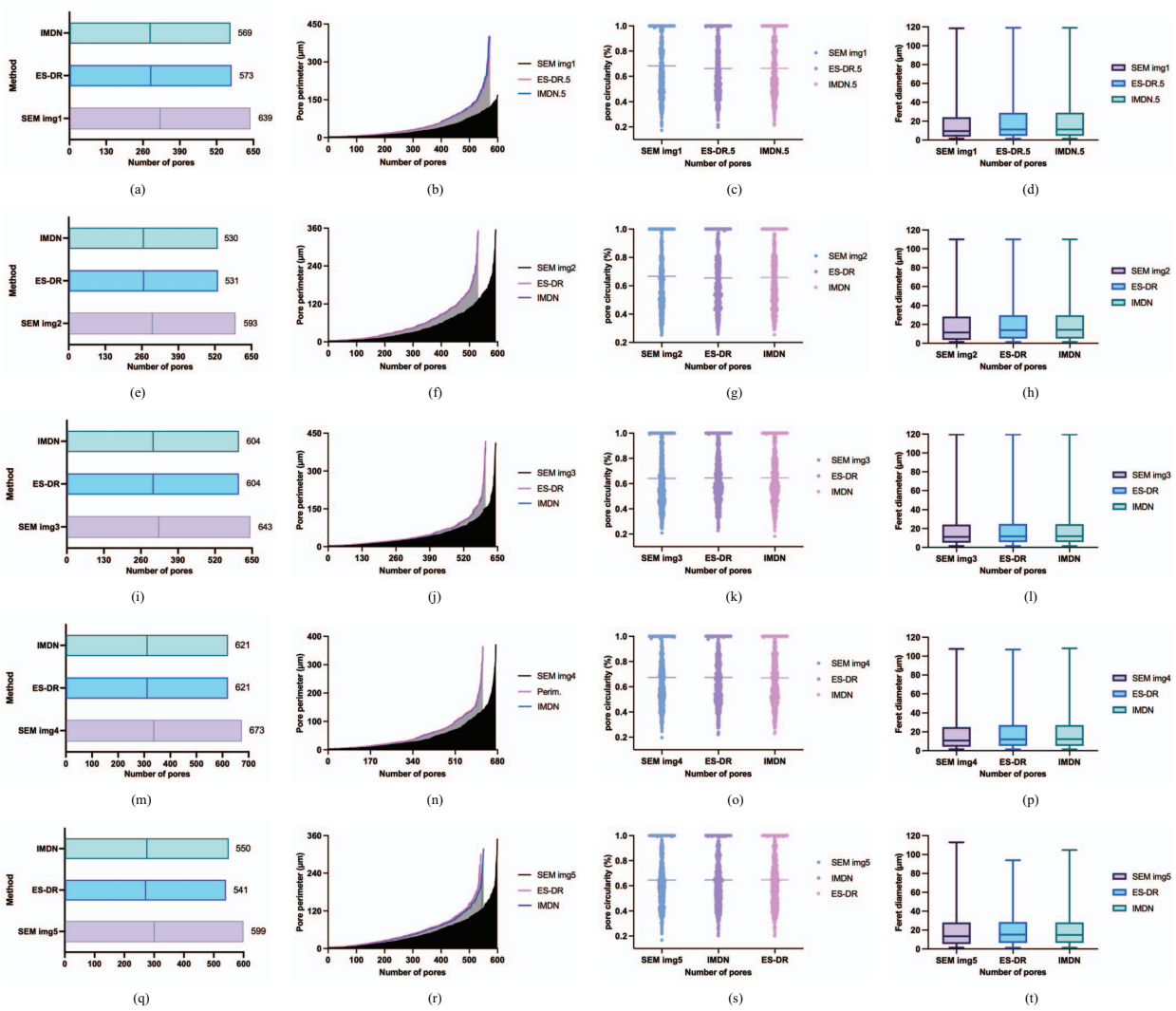

Fig. 4. Comparing the feature values of ES-DR and IMDN with the original image by bicubic interpolation with double downsampling.

[23] Kamrava S, Tahmasebi P, Sahimi M. "Enhancing images of shale formations by a hybrid stochastic and deep learning algorithm". Neural Networks, 2019, 118: 310-320.

[24] Xie Y, Zhang J, Shen C, et al. "Cotr: Efficiently bridging cnn and transformer for 3d medical image segmentation". Medical Image Computing and Computer Assisted Intervention–MICCAI 2021: 24th International Conference, Strasbourg, France, September 27–October 1, 2021, Proceedings, Part III 24. Springer International Publishing, 2021: 171-180.

[25] Vaswani A, Shazeer N, Parmar N, et al. "Attention is all you need". Advances in neural information processing systems, 2017, 30.

[26] Y. Wang, P. Mostaghimi, and R. Armstrong, "A Super Resolution Dataset of Digital Rocks (DRSRD1): Sandstone and Carbonate", Digital Rocks Portal, 2019. [Online]. Available: http://www.digitalrocksportal.org.

[27] Dong C, Loy C C, Tang X. "Accelerating the super-resolution convolutional neural network". Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11-14, 2016, Proceedings, Part II 14. Springer International Publishing, 2016: 391-407.

[28] Lu Z, Li J, Liu H, et al. "Transformer for single image super-resolution". Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2022: 457-466.

[29] Keys R G ."Cubic convolution interpolation for digital image processing". IEEE Transactions on Acoustics, Speech, and Signal Processing, 2003, 29.

[30] Ledig C, Theis L, Huszár F, et al. "Photo-realistic single image super-resolution using a generative adversarial network". Proceedings of the IEEE conference on computer vision and pattern recognition. 2017: 4681-4690.

[31] Hui Z, Gao X, Yang Y, et al. "Lightweight image super-resolution with information multi-distillation network". Proceedings of the 27th acm international conference on multimedia. 2019: 2024-2032.