Invited review

Advances in the application of deep learning methods to digital rock technology

Xiaobin Li¹, Bingke Li¹, Fangzhou Liu², Tingting Li³, Xin Nie¹,4*¹

¹School of Geophysics and Petroleum Resources, Yangtze University, Wuhan 430100, P. R. China.
²Department of Civil and Environmental Engineering, University of Alberta, Edmonton AB T6G 2R3, Canada
³School of Earth Sciences, Northeast Petroleum University, Daqing 163318, P. R. China
⁴Key Laboratory of Exploration Technologies for Oil and Gas Resources, Yangtze University, Wuhan 430100, P. R. China

Keywords:
Digital rock reconstruction
super resolution
segmentation
parameter prediction
deep learning

Abstract:
Digital rock technology is becoming essential in reservoir engineering and petrophysics. Three-dimensional digital rock reconstruction, image resolution enhancement, image segmentation, and rock parameters prediction are all crucial steps in enabling the overall analysis of digital rocks to overcome the shortcomings and limitations of traditional methods. Artificial intelligence technology, which has started to play a significant role in many different fields, may provide a new direction for the development of digital rock technology. This work presents a systematic review of the deep learning methods that are being applied to tasks within digital rock analysis, including the reconstruction of digital rocks, high-resolution image acquisition, grayscale image segmentation, and parameter prediction. The results of these applications prove that state-of-the-art deep learning methods can help advance and provide a new approach to scientific knowledge in the field of digital rocks. This work also discusses future research and developments on the application of deep learning methods to digital rock technology.

1. Introduction

In recent years, the demand for oil and for greater efficiency and benefits in oil exploration has steadily increased. However, conventional reservoir exploration has been unable to meet the vast requirements of the market. Therefore, unconventional reservoir exploration has received increasing attention, and digital rock technology has become increasingly important for its realization (Wang et al., 2021a; Yang, 2022). Digital rock images can be used to describe pore and grain morphology (Blunt et al., 2013; Xia et al., 2019). Moreover, physical properties of rocks, such as permeability, resistivity and elasticity, can be obtained through numerical simulations based on three-dimensional (3D) digital rocks (Nie et al., 2016a, 2016b; Zhu and Shan, 2016; Zhu et al., 2019; Andhumoudine et al., 2021). Establishing 3D digital rock models, converting from low- to high-resolution rock images, segmenting grayscale rock images, and obtaining rock properties from digital rocks are essential procedures in digital rock technology. However, these procedures are time-consuming and expensive. For example, the widely used Lattice Boltzmann method (LBM) is highly accurate in permeability calculation, the size of the sample is limited due to the method's high computational cost (Okabe and Blunt, 2004; Wu et al., 2006; Liu et al., 2022).

The rapid development of artificial intelligence over the last decade has made the application of deep learning methods a promising solution to these intractable problems in digital rock technology (Cai et al., 2020; Wang et al., 2021c; Xiao, 2022). This work presents some successful examples of the use of deep learning methods in the reconstruction of digital rocks, acquisition of high-resolution digital rock images, automatic
2. 3D digital rock reconstruction

Digital rocks are the images of rocks. There are three traditional digital rock reconstruction methods: physical experiments, numerical reconstruction, and hybrid modelling (Lin et al., 2018; Zhao et al., 2020). Physical experiments mainly include scanning electron microscopy (SEM), focused ion beam scanning (FIB-SEM), and X-ray computed tomography (CT). All these physical procedures require expensive experimental equipment, take a long time to complete even when scanning a small-sized core, and cannot guarantee both large imaging volume and high resolution (Yang et al., 2021). Numerical reconstruction methods include process-based modeling and stochastic methods including simulated annealing method, Markov chain Monte Carlo method, truncated Gaussian random field method, multiple-point statistics, and so on (Wang et al., 2013; Yao et al., 2013; Yang et al., 2015; Yao et al., 2018). Although the 3D digital rocks obtained from numerical reconstruction methods have good pore structures, the reconstruction results are controlled by different constraints. The more the conditions, the better the reconstruction results and the higher the calculation cost. The hybrid modelling method combines the advantages of the first two methods, obtaining two-dimensional (2D) slice data of rocks via SEM and other physical methods, and then reconstructing the 3D digital rocks through numerical reconstruction. However, this reconstruction method is often aimed at highly homogeneous rocks, making it unsuitable for unconventional reservoirs with complex and heterogeneous pore structures, such as shale (Cao et al., 2022).

Deep learning algorithms that can be used to generate images are primarily composed of generative adversarial networks (GANs) and variational autoencoders (VAEs) (Cang et al., 2018; Zhang et al., 2021a). Compared to VAEs, GANs can generate more realistic images. Therefore, various variants of GANs are primarily used to reconstruct digital cores. The GANs have powerful image-generation abilities (Goodfellow et al., 2014), which are mainly composed of generator and discriminator networks. The generator is used to learn the distribution characteristics of real data samples and generate fake samples that are similar to the real ones. The goal of the discriminator is to accurately distinguish the input samples from real or fake training samples. As the number of training epochs increases, the capabilities of the generator and discriminator continue to improve, as shown in Fig. 1. Once the training process is completed, the generator can be used directly to generate realistic images.

Although GANs have achieved remarkable results in generating high-quality images, they have problems such as difficult training, only generate a single image, and the lack of indicators to monitor the training progress (Hu et al., 2022). In response to these problems, many researchers have improved GANs by developing different variants. Table 1 compares the advantages and disadvantages of several common GANs variants, including Deep Convolutional Generative Adversarial Networks (DCGANs), Least Square Generative Adversarial Networks (LSGANs), Wasserstein Generative Adversarial Networks (WGANs), Wasserstein Generative Adversarial Networks with Gradient Penalty (WGANs-GP), Large-scale Generative Adversarial Networks (BigGANs), Cycle-Consistent Generative Adversarial Networks (CycleGANs). As can be seen, the emphasis on reconstruction models varies between different GANs variants. In practice, additional experimental requirements are set to choose a suitable GANs that provides adequate results according to our needs.

Many scholars have studied the use of GANs for digital rock reconstruction. Mosser et al. (2017) creatively used DCGANs to achieve the rapid reconstruction of a bead pack, Berea sandstone, and Ketton limestone. The synthetic rocks were compared to the original ones using two-point statistics,
Table 1. Comparison of common GANs variants.

<table>
<thead>
<tr>
<th>Variant name</th>
<th>Advantage</th>
<th>Disadvantage</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>DCGANs</td>
<td>Fast training and small memory occupied</td>
<td>Chessboard effect resulting in poor picture quality</td>
<td>Radford et al. (2015)</td>
</tr>
<tr>
<td>LSGANs</td>
<td>Less gradient disappearance and high generative quality</td>
<td>Not good enough for generating large-scale images</td>
<td>Mao et al. (2017)</td>
</tr>
<tr>
<td>WGANs</td>
<td>High training stability</td>
<td>Producing bad samples and complex convergence</td>
<td>Arjovsky et al. (2017)</td>
</tr>
<tr>
<td>WGANs-GP</td>
<td>Short convergence time and stable training</td>
<td>Unable to generate high-resolution images</td>
<td>Gulrajani et al. (2017)</td>
</tr>
<tr>
<td>BigGANs</td>
<td>High resolution and good quality</td>
<td>Complex network structure and high computational cost</td>
<td>Brock et al. (2018)</td>
</tr>
<tr>
<td>CycleGANs</td>
<td>High realism and diversity of the results</td>
<td>Complex network structure and training process</td>
<td>Zhu et al. (2017a)</td>
</tr>
</tbody>
</table>

Fig. 2. Original shale digital rock (Yang et al., 2021).

Fig. 3. Schematic diagram of the generator structure (Yang et al., 2021).

morphological characteristics, and single permeability. The results showed good consistency between the synthetic and original cores. This work used 3D digital rocks composed of pores and grains, whereas posterior studies focused on the reconstruction of 3D grayscale rocks based on 2D grayscale slices. Based on previous work, Mosser et al. (2018) reconstructed Oolitic limestone in grayscale. Likewise, Zha et al. (2020) used WGANs to generate realistic 2D grayscale images of shale, demonstrating that GANs can also achieve good results in the reconstruction of highly heterogeneous. Valsecchi et al. (2020), Feng et al. (2020) and You et al. (2021) realized the reconstruction of 3D digital rocks from 2D slices. In addition, the time consumption and average use of CPU, GPU, and memory by GANs and traditional methods were compared, proving that the reconstruction speed of GANs was faster (Feng et al., 2020; Zhang et al., 2021b).

The application of GANs to digital rock reconstruction were illustrated through the 3D digital rock reconstruction of shale by Yang et al. (2021). They obtained large-size training images based on three-dimensional FIB-SEM images of real shale cores, grouped the four mineral components of the skeleton phases of the original core images into one phase, and recombined them with pore phases into a new three-dimensional volume data (Fig. 2). Then, a six-layer deep convolution neural network was used as the generator (Fig. 3) and discriminator.
Fig. 4. 3D digital rocks and corresponding 2D slices produced by the generator (Yang et al., 2021).

Fig. 5. Structure of SR-Resnet.

Partial displays of the 3D cores and 2D slices are shown in Fig. 4. The digital rock synthesized by the generator can not only restore the pores of the original core but also capture their spatial distribution, which gives a good description of the pore space of the original core. These results show that the trained GANs model can generate digital shale cores that satisfy the given pore structure information.

3. Image resolution enhancement

Microcomputed tomography (micro-CT) is widely used to describe the microstructure of conventional rocks (Flannery et al., 1987; Coenen et al., 2004; Shan et al., 2022). However, owing to the inherent limitations of micro-CT, the field of view of high-resolution (HR) images is small and the resolution of images with a large field of view is low (Li et al., 2017).

To overcome the limitations of micro-CT, a super-high-resolution (SR) algorithm was proposed in the 1960s. As an effective method to overcome the tradeoff between the field of view and image resolution, the SR algorithm can reconstruct rock micro-CT images from low-resolution (LR) to HR images. However, in most cases, existing SR methods cannot satisfactorily produce HR images, lack flexibility in the generation stage, and have problems such as a high computational cost (Yang et al., 2008). Some deep learning methods, such as super-resolution convolutional neural network (SRCNN) (Wang et al., 2019a), super-high-resolution cycle-consistent generative adversarial networks (SR-CycleGAN) (Wang et al., 2019a), and hybrid spatiotemporal deep learning (HSDL) (Kamrava et al., 2019), have been applied in continuous exploration to effectively improve the image resolution.

Dong et al. (2015) proposed an SR with convolutional neural network SRCNN. Through mapping and reconstruction, the algorithm can convert LR images into SR images. Wang et al. (2019a) applied the SRCNN technique to the trained HR source and twice to the LR source to generate high-resolution images of sandstone and carbonate rocks. Compared with the bicubic interpolation, the experimental results show that the image quality improved and the relative error reduced by 50%-70%, indicating that the SRCNN can generate high-quality, high-resolution images by processing sandstone and carbonate images. Compared with the traditional method, the recovery quality significantly improved, indicating that the SRCNN method can be used as a feasible processing step in the digital rock workflow. Three models were subsequently developed from the SRCNN (Wang et al., 2019a): SR-Resnet, enhanced deep SR (EDSR), and wide-activation deep SR (WDSR), all of which had similar structures (Figs. 5-7). LR Image: Bentheimer (50 × 50), 15.2 micron resolution, 0.76mm × 0.76mm; SR Image: Bentheimer (200 × 200), 3.8 micron resolution 0.76mm × 0.76mm).

The application of GANs to high resolution image generation can also achieve good results. Ledig et al. (2017) proposed to apply the GANs to image super resolution, namely SRGAN. Zhu and Zheng (2022) used SRGAN to carry out research on super-resolution reconstruction of rock micro-CT images, and achieved good results. Training SRGAN requires a large number of paired data, which is usually difficult to achieve. Therefore, the Cycle-GAN network is proposed to
solve this problem (Zhu et al., 2017b). Chen et al. (2020a) proposed a simple rock micro-CT image reconstruction based on an SR-CycleGAN (Fig. 8). The SR-CycleGAN transcends the limitations of imaging systems in terms of field of view and resolution and can simultaneously obtain a large field of view and HR rock micro-CT images. This method consists of two stages: offline training and online testing (Fig. 9). The offline training stage uses a set of unpaired rock micro-CT images to train the network. In the online testing phase, the mapping between micro-CT images is modeled and the SR-CycleGAN improves the resolution of the LR input by learning the mapping. Compared with the HR results generated by LR rock images and bicubic interpolation (Fig. 10), the experimental results show that the SR-CycleGAN algorithm can significantly improve the quality of the simulated and real rock micro-CT images.

The problem of lacking training image, which limit the usage of deep learning algorithms for image super resolution. To address this issue, Kamrava et al. (2019) proposed using an HSDL algorithm for generating a large number of plausible shale. This method uses very few input images to train a deep learning stochastic convolutional network at a meager cost while improving image resolution (Kamrava et al., 2019; Wang et al., 2019a; Chen et al., 2020a). In addition, Kamrava et al. (2019) used the HSDL algorithm to analyze and model complex shale formations with irregular pores. The results showed that the accuracy of the HSDL algorithm is higher than that of the conventional deep learning algorithm without reinforcement training. The frequency distribution of the images generated by the HSDL algorithm was close to that of the reference images, and the enhanced images generated by the HSDL algorithm were consistent with the original HR images (Fig. 11).

4. Image segmentation

In the field of digital rock, image segmentation usually only needs to divide the pore space and solid particle space. However, the binary images containing pore and solid grain are not applicable in the case of P- and S- wave speed simulation (Andrä et al., 2013), hybrid moisturizing flow simulation (Akai et al., 2019), and non-response transport simulation (Liu et al., 2018). It is necessary to segment the grayscale image of rock into the image containing multiple mineral components. However, traditional image segmentation methods (Seo et al., 2020), such as multi threshold segmentation, edge detection, clustering segmentation, are difficult to accurately segment.
Fig. 8. The architecture of the SR-CycleGAN which comprises (a) two generators ($G_X : Y \rightarrow X$ and $G_Y : X \rightarrow Y$) and two associated discriminators ($D_X$ and $D_Y$), (b) forward cycle consistency: $x \approx G_X (G_Y (x))$. (c) Backward cycle consistency: $y \approx G_Y (G_X (y))$.

Fig. 9. The off-line training phase (bottom) and the on-line testing phase (top) of the SR-CycleGAN (Chen et al., 2020a).

Fig. 10. LR rock images: (top) HR results produced through bicubic interpolation, (middle) HR results produced by the SR-CycleGAN, and (bottom) ground truth (Chen et al., 2020a).
rock images. To address the lack of traditional image segmentation, some scholars have used convolution neural network (CNN) for image semantic segmentation (Ning et al., 2005; Cireşan et al., 2012; Farabet et al., 2012; Ganin and Lempitsky, 2014; Gupta et al., 2014; Pinheiro and Collobert, 2014). Deep learning methods such as fully convolutional networks (FCN) (Long et al., 2015), U-Net (Ronneberger et al., 2015), DeepLab (Chen et al., 2014), and SegNet (Badrinarayanan et al., 2017) have been applied in semantic segmentation. Since the U-Net network was proposed, the U-shaped network structure of encoders and decoders connected by skip connection has been widely used in the field of image segmentation and the improvement structures were proposed, such as SegNet, U-Net++ (Badrinarayanan et al., 2017; Zhou et al., 2018). The U-Net structure is illustrated in Fig. 12, the left side of the network is the encoder, the right side is the decoder, and the two sides are connected using the skip connection layer.

SegNet uses the pooling index obtained by the maximum pooling step of the encoder corresponding to the decoder for nonlinear samples (Fig. 13), which not only improves the image resolution but also reduces the need for the learning about up-sampling. For different aspects and the outlook re-

---

**Fig. 11.** Comparison of image resolution enhancement methods: (a) reference image, (b) low-resolution input image, (c) regular deep learning image, (d) bicubic interpolation image and (e) HSDL-generated image (Kamrava et al., 2019).

**Fig. 12.** U-Net structure (Ronneberger et al., 2015).
Fig. 13. The max pooling index is used for up-sampling low-resolution graphs in SegNet (Badrinarayanan et al., 2017).

Fig. 14. The basic internal structure of SegNet (Karimpouli and Tahmasebi, 2019a).

Table 2. Prediction of performance of different learning models (Li et al., 2021).

<table>
<thead>
<tr>
<th>Models</th>
<th>F1-score</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Micro</td>
</tr>
<tr>
<td>Logistic regression</td>
<td>0.8705</td>
</tr>
<tr>
<td>Linear SVM</td>
<td>0.7980</td>
</tr>
<tr>
<td>K-Nearest neighbors</td>
<td>0.9150</td>
</tr>
<tr>
<td>Random forest</td>
<td>0.9238</td>
</tr>
<tr>
<td>ANN</td>
<td>0.8869</td>
</tr>
<tr>
<td>U-Net</td>
<td>0.8832</td>
</tr>
</tbody>
</table>

et al., 2017), the internal structure of the SegNet is shown in Fig. 14. Karimpouli and Tahmasebi (2019a) used the SegNet to realize the automatic segmentation of rock image of Berea sandstone, which can achieve more accurate results compared with the multi-threshold segmentation method.

Previous research shows that the U-Net network can achieve good results in segmenting digital rock images. Wang et al. (2021b) realized automatic segmentation of 2D and 3D sandstone CT images using 2D and 3D U-Net network models. Experiments have proved that with revised weight function the U-Net architecture can effectively distinguish clay sets of mixed matrix mineral particles and organic matter. Chen et al. (2020b) used the U-Net architecture with modified the weight function to segment the scanning electron microscope images of Duvernay shale samples. The experimental results (Fig. 15) show that there is an obvious separation between the clay collection and matrix mineral particles, and the boundary is acceptable. This method proves that extracting deep-learning features based on texture is feasible, economical, and timely.

In addition, Li et al. (2021) used a variety of image segmentation methods, including support vector machine, k-nearest neighbor, random forest, artificial neural network and U-Net network model to perform multi-component segmentation of shale SEM images. The image segmentation effects of different methods can be seen in Table 2, and the U-Net model can achieve better results. Meanwhile, compared with the machine learning method used in this paper, the U-Net model uses the whole image as the input to achieve end-to-end image segmentation, without manually extracting image features, but also considering the neighborhood information of pixels. As for multi-component segmentation of shale CT images, Li et al. (2022) used U-Net network model to obtain a 3D shale
core model with multiple mineral components by combining shale CT images and QEMSCAN images containing mineral information. It is challenging to segment the small targets and the pixels near the boundary accurately in the rock image. To solve this problem, Wang et al. (2022) used the U-Net++ structure to carry out image segmentation experiments, compared the segmentation results of U-Net++ with those of the commonly used U-Net and wide U-Net model, and found that U-Net++ can achieve satisfactory results in terms of pixel-wise and physics-based evaluation metrics. The network structure of U-Net++ (Fig. 16) contains four conventional U-Net networks, the blocks and skip connections shown in the black are for the fourth U-Net. Comparison of segmentation results of different segmentation methods is shown in Fig. 17.

5. Digital rock parameter prediction

Rock parameters prediction is essential for formation evaluation (Wang et al., 2020). For example, permeability and porosity, which are inherent properties of the rocks whose values depend only on the pore structure, are important parameters used to characterize the heterogeneity and anisotropy of reservoirs (Haagsma et al., 2021; Ishola and Vilcaez, 2022). Numerical simulation has been used to calculate rock’s parameters. Zhu et al. (2008) used the LBM to study the seepage characteristics of porous rocks. Song et al. (2015) used a structured pore network model of the rock samples to predict their permeability. Since the calculations involved are very time-consuming, the application of numerical simulation methods tends to be limited to small domains depending on the computational resources available. Therefore, the rapid prediction of core parameters has attracted great research interest.

Achievements have also been made in the parameter prediction of digital rock using deep learning algorithm. Alqahtani et al. (2020) used CNNs to predict the porosity, specific surface, and average pore size of binary and grayscale 2D X-ray images of sandstone. The results showed that the parameters predicted were consistent with ground truth, with a relative
Fig. 17. Comparison of segmentation results of different segmentation methods (Wang et al., 2022).

Fig. 18. Schematic of the CNN network for the prediction of parameters.

error of less than 7% for both binary and grayscale CT images. Misbahuddin (2020) utilized CNNs to predict properties from grayscale SEM images of shale. The relative errors in the predicted porosity and average pore radius were 0.4% and 1%, which are considered negligible. For 3D digital rocks, methods based on CNNs can also achieve excellent results (Srisutthiyakorn, 2016; Alqahtani et al., 2018; Yang et al., 2018; Karimpouli and Tahmasebi, 2019b). The schematic of the CNN network for the prediction of rock’s parameters in shown in Fig. 18. Karimpouli and Tahmasebi (2019b) used CNNs to estimate P- and S- wave velocities from digital rock images, the estimated properties were compared with the numerical simulation results, indicating that CNNs perform outstanding in predicting the physical parameters. Tembely et al. (2021) used CNNs to predict the porosity, formation factor, and permeability of 3D CT images with high accuracy. Rabbani et al. (2020) proposed a workflow based on CNNs to estimate a wide range of morphological, hydraulic, and electrical properties for binarized 3D CT images. Compared with traditional methods, the workflow proposed by Rabbani is compatible with any physical size. Zhang et al. (2022) proposed to predict the permeability of porous media from low-resolution images and achieved outstanding results.

A specific example was illustrated regarding the use of CNNs to predict the parameters of digital rocks. Wang et al. (2019b) established a 3D pore network model based on the OpenFOAM framework and calculated the porosity and permeability of pore network model. Then the 3D CT images
and corresponding porosity and permeability are as training dataset for the CNNs model training. The actual data exhibits a strong matching relationship between pores and permeability (Fig. 19). This experiment shows that the fast prediction digital rock model established based on the 3D CNN has excellent potential and generalization ability for digital rock feature extraction. The prediction time of a good deep learning model was only 0.03 s. In contrast, the average time of the OpenFOAM numerical simulation method was 1.58 h, thus demonstrating a significant improvement in the calculation speed and providing an effective solution for the rapid prediction of digital rock permeability. In summary, deep learning methods can predict parameters of digital rocks accurately and improve prediction speed in one order of magnitude compared with traditional numerical simulation methods, effectively reducing computational costs, and significantly improving work efficiency.

6. Conclusions

This work provides an overview of the application of deep learning methods to 3D digital rock reconstruction, image resolution enhancement, image segmentation, and digital rock parameter prediction. Although digital rock technology has been developed for decades, many research challenges are yet to be addressed. The methods mentioned in this study have partially overcome the challenges posed by reconstruction, resolution enhancement, segmentation, and parameter prediction tasks. However, these methods are still unable to simultaneously consider the training speed, image size, and modeling accuracy. Therefore, the application of artificial intelligence methods in the field of digital rocks should be more comprehensively developed. The reconstruction of the digital rocks should be constrained by physical properties to make sure the reality and diversity of the generated samples. The component segmentation procedures by now are mostly based on 2D slices, which cannot assure the continuity of the components in all directions. Therefore, the ortho-slice segmentation should be considered. Parameter prediction should not only based on the image itself, but also the physical properties such as porosity and pore space distribution. In addition, the accuracy of segmentation can be further improved and attempts to predict more rock parameters are needed.

Moreover, as the future development of intelligent digital oilfields becomes a general trend, researchers are suggested to make full use of the powerful capabilities of deep learning and other artificial intelligence methods to continuously learn and update the collected core data and properties. Unlike other common tasks for machine learning such as number or animal identification, its application in digital rock area is relatively new and lack of solid data. Therefore, it is necessary to build an open source and renewable database containing rocks’ digital images and their physical properties. This should enable the combination of geological and geophysical data to allow a comprehensive and systematic development of reliable strategies for integrating microscopic and local digital rock technology into macroscopic and overall exploration and development processes.

Acknowledgements

This research was supported by the National Natural Science Foundation of China (No. 41504094), the Undergraduate Training Program of the Yangtze University for Innovation and Entrepreneurship (No. YZ2020042), and Heilongjiang Provincial Natural Science Foundation of China (No. LH2020D008).

Conflict of interest

The authors declare no competing interest.

Open Access This article is distributed under the terms and conditions of the Creative Commons Attribution (CC BY-NC-ND) license, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

References


Badrinarayanan, V., Kendall, A., Cipolla, R. SegNet: A deep


Li, Z., Teng, Q., He, X., et al. Sparse representation-based


