SPE-195834-MS

Incorporated Artificial Intelligence and Digital Imaging System for Unconventional Reservoirs Characterization

Huafeng Sun, China University of Petroleum-Beijing; Hadi Belhaj, Khalifa University of Science and Technology

This paper was prepared for presentation at the SPE Annual Technical Conference and Exhibition held in Calgary, Alberta, Canada, 30 Sep - 2 October 2019.

Abstract

The unconventional reservoirs (UCRs) play a key role in global oil and gas supply. However, their reservoir characterization is difficult because of complex pore structure and low permeability-viscosity ratio. Usually, traditional techniques hardly can be used for determination of pore structure and estimation of reservoir properties. In this case, digital rock analysis (DRA) shows the potential for capturing detailed pore structure information and simulating rock properties, such as porosity, permeability, electrical properties and elastic properties. Recently, artificial intelligence (AI) techniques have presented an ever-increasing trend in a wide variety of research and commercial fields. Many AI applications can free man from the labor of complicated works in some way. Machine leaning (ML), which is a subdivision of AI, has attracted researchers' attention and been widely used in geoscience and reservoir characterization, such as feature extracting, rock type prediction and reservoir property estimation. The incorporation of AI and DRA is becoming an inevitable development trend for future reservoir study. In this paper, firstly, DRA workflow for reservoir characterization is introduced; secondly, the commonly used ML algorithms in DRA study is reviewed; finally, a case study of characterization of a tight carbonate reservoir with ML algorithm and DRA is presented. The analysis shows that ML can be applied in any part of DRA progress such as image segmentation, feature detection, rock image classification, numerical simulation and result analysis. Compared with traditional DRA algorithm, ML-based methods can reduce manual operation that has greatly impact on the results. The combination of ML and DRA provides a new insight in UCRs characterization and outlook the future opportunities of AI to solve the oilfield problems.

Introduction

The unconventional reservoirs (UCRs) include reservoirs such as tight gas sands, gas and oil shales, coalbed methane, heavy oil and tar sands, and gas-hydrate deposits, which have become a vital component of global oil and gas supply now and production of unconventionals account for nearly 10% of the total global yield (Zou et al., 2013). However, traditional techniques are not applicable to industrial production because unconventional petroleum resources refer to continuous or quasi-continuous accumulations of petroleum resources. Figure 1 shows a simple chart that can explain the physical properties of rock and the nature of the fluids in petroleum resources. In this figure, the X-axis and Y-axis are both logarithmic. It can be seen that

the conventional resources fall into the bottom-right quadrant of the chart and all unconventional resources fall outside the bottom-right quadrant. Therefore, for unconventional resources, special recovery solutions that can overcome economic constraints of enhanced oil recovery are needed because of their relatively low permeability-viscosity ratios (Zou, 2017).

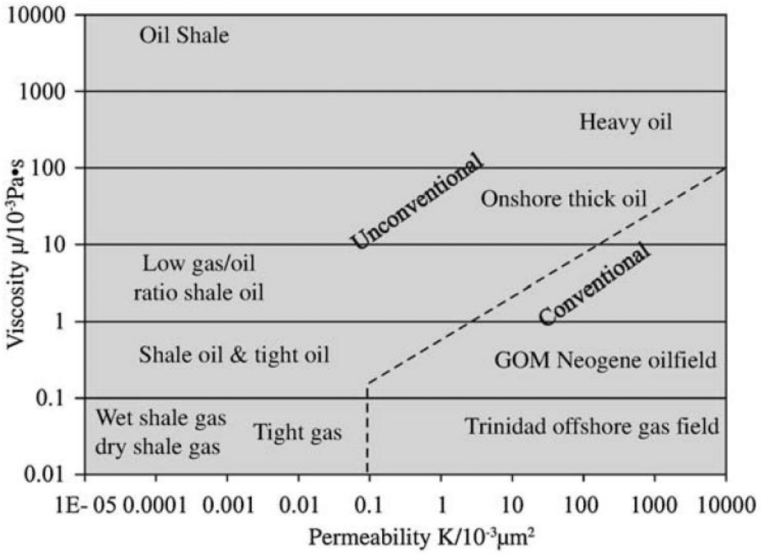

Figure 1—Viscosity-permeability relationship of petroleum resources (Zou, 2017).

To help improve well economics in unconventional resources, it is important to better characterize their reservoirs. The most effective way is to evaluate rock properties with routine core analysis (RCA), special core analysis (SCAL) and digital rock analysis (DRA). Among these methods, DRA has more advantages such as non-destructivity, repeatability and intuitive visualization. This technique can accurately image rock pore structure from nano scale to millimetre scale and can simulate rock properties, such as porosity, permeability, electrical properties, and elastic properties, which provides a new insight in reservoir characterization (Sun et al., 2019). For UCRs study, DRA also has a good application. For example, Fathi et al. (2012), Zhang et al. (2014), Chen et al. (2015), Sun et al. (2015) and Sun et al. (2017a) used DRA to study the flow properties of shale gas in the nanometer pores of kerogen. Liu et al. (2018) used this technique to critically study the tight sandstone reservoir. Sun et al. (2017b) characterized the gas transport behaviors in tight gas reservoir by DRA. In addition, low permeability and tight carbonate reservoir were also studied by using this technique (Mishra and Akbar, 2015; Sun et al., 2019).

In recent years, artificial intelligence (AI) techniques have shown an ever-increasing development trend and affected a wide variety of research and commercial fields (McCoy and Auret, 2019). A number of highly-publicized successes have been made in different area, such as significant improvements to language translation, self-driving cars and buses, voice and facial recognition, online shopping, smart phone and furniture, video streaming and social media platforms (Wu et al., 2016; Silver et al., 2016 & 2017; Sutton and Barto, 2018). Machine leaning (ML) is a subdivision of AI based on the biological learning process. It is an effective approach for both regression and/or classification of nonlinear systems involving a few or literally thousands of variables (Lary et al., 2016). The commonly used ML algorithms include artificial neural networks (ANN), decision trees, fuzzy c-means (FCM), self-organizing map (SOM), support vector machines (SVM), random forests, case-based reasoning, neuro-fuzzy, genetic algorithm, and multivariate adaptive regression splines (Shahin et al., 2001; Das and Basudhar, 2008; Samui, 2008; Azamathulla and Wu, 2011; Garg et al., 2014a, b, c). Many publications have recently demonstrated significant advances of ML applications in many fields such as medical contexts (Criminisi, 2016; Calhoun, 2018), finance (Sarlin

and Björk, 2017; in't Hout et al., 2018), environmental science (Gibert et al., 2018a, b), and data mining (Boixader, 2017).

Some researchers tried to use ML methods for DRA study. Srisutthiyakorn (2016) used deep learning (DL) methods for predicting permeability from two-dimensional (2D) and three-dimensional (3D) digital rock images of the Fontainebleau and Berea sandstone. Chauhan et al. (2016a, b) evaluated the performance and accuracy of ML techniques to segment rock grains, matrix and pore voxels from 3D X-ray computed tomography (XCT) images. The segmentation and classification capability of k-means, FCM, SOM, ANN, least-squares SVM, bragging and boosting were tested using XCT images of different rock samples, such as andesite volcanic rock, Berea sandstone, Rotliegend sandstone and a synthetic sample. Budennyy et al. (2017) presented a methodology of petrographic thin section analysis, combing the image processing and statistical learning methods. Saad et al. (2018) used ML methods for rock mechanical properties characterization. Tian et al. (2018) used ML algorithms for feature detection with digital rock images.

Similarly, many researchers are trying to apply ML algorithms in reservoir study. For example, Helmy et al. (2010) utilized the capabilities of data mining and computational intelligence in the prediction of porosity and permeability for reservoir characterization, based on the hybridization of Fuzzy Logic, SVM, and Functional Networks, using several real-life well-logs. Al-Anazi and Gates (2010) examined the potential of SVM to recognize electrofacies and predict permeability in a heterogeneous sandstone reservoir. Patel and Chatterjee (2016) presented a computer vision-based rock-type classification algorithm for fast and reliable identification. Viswanathan and Samui (2016) determined the rock depth using three ML techniques, which are Gaussian process regression, least squares SVM and extreme learning machine. Xu et al. (2018) proposed a ML-based method for borehole resistivity measurement modeling and they applied this technique to the logging-while-drilling measurements modeling problem. Likewise, some ML-based methods for UCRs study have been used by some researchers. For example, Esmaili and Mohaghegh (2016) focused on full field reservoir modeling of shale assets by a method of advanced data-driven analytics. Yu et al. (2017) proposed a new method, based on Gaussian Process Regression bridging geostatistics and ML technique, for TOC estimation in tight shale gas reservoirs. Tahmasebi et al. (2017) used a combination method of data mining and ML for identifying sweet spots in shale reservoirs. Ma et al. (2018) presented a preliminary attempt in correlating stochastic shale parameters with observable features in production time-series data using an ANN model. Wang and Chen (2019) developed a comprehensive data mining process to evaluate well production performance in UCRs based on ML modeling. Tian and Daigle (2019) presented a completely automatic ML method for quantifying the preferential mineral-microfracture relationships in intact and deformed shales. All these works show a good application of ML methods in reservoir study.

The above examples demonstrate the applicability of ML-based methods in DRA and open a new area of digital rock research for reservoir characterization. This paper starts with an introduction of DRA workflow in reservoir characterization (Section 2); then, some ML-based methods in DRA study are reviewed in Section 3; Section 4 presents a case study for reservoir characterizations of a complex tight carbonate reservoir; Last section (Section 5) summarizes this paper and discusses some future prospects.

DRA Workflow

DRA includes rock image acquisition, image processing, numerical simulation, and result analysis. This technique has become a reliable method for obtaining accurate digital rock images and rock properties from micro-scale physical processes by using numerical methods. Figure 2 presents a DRA workflow for reservoir study. This workflow makes full use of 2D and 3D rock images for observation, processing, simulation and analysis, which can better extract useful information for reservoir characterization.

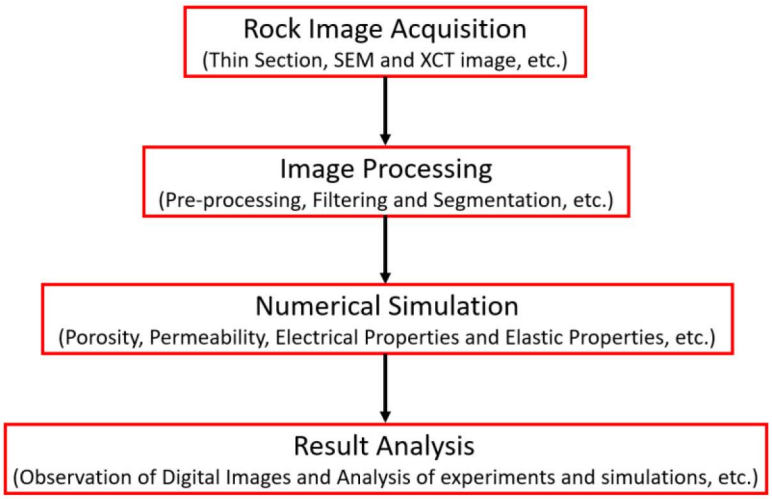

Figure 2—DRA workflow for reservoir characterization.

The commonly used image acquisition tool is micro-computed tomography (MCT). Figure 3 shows a schematic diagram of X-ray MCT scanner. It can be seen that, after scanning a real rock sample by rotating for 360 degree, the 3D rock image can be obtained in the computer (Dong and Blunt, 2009). This technique has captured the internal formation of rock sample in detail. In addition, other imaging techniques including thin section, scanning electron microscope (SEM), and focused ion beam- scanning electron microscope can also provide rock images for DRA progress.

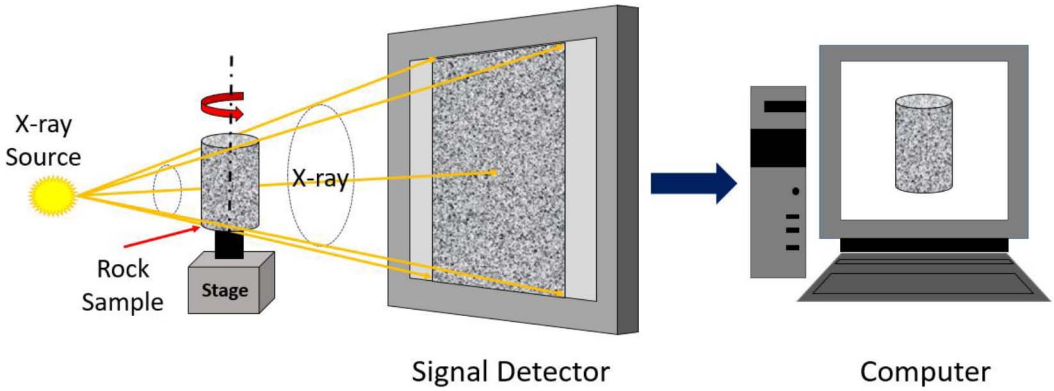

Figure 3—Schematic diagram of 3D rock image acquisition by MCT scanner (Dong, 2007).

The image processing includes image pre-processing, image filtering and image segmentation. The purpose of image pre-processing is to select a representative subsample from the whole image. This process is usually called representative element volume selection. Image filtering can remove the noise from the rock image and enhance the image. Image segmentation can separate the pores, the grains and other minerals. Figure 4 shows an example of image processing in DRA study. It can be seen that the digital rock model with separated pores and solid phase is built for simulation and analysis after image processing. Some phases containing some minerals and cement are defined as unresolved phase (green color in Figure 4 (d)), because these phases hardly can be clearly segmented.

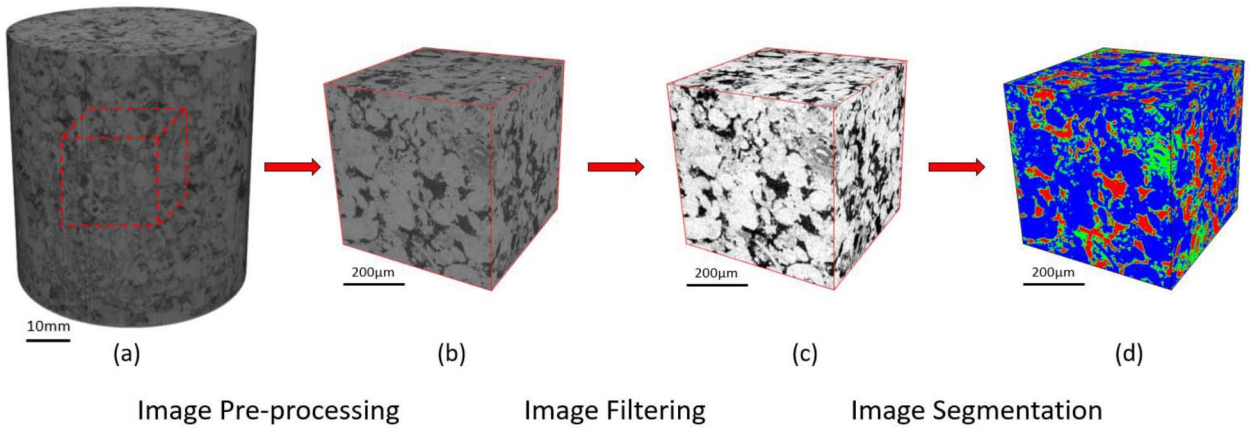

Figure 4—Example of image processing in DRA progress: (a) is the original rock image with size of 1000×1000×1000 voxels and image resolution of 1μm/voxel, (b) is the subsample with size of 400×400×400 voxels, which is selected from (a), (c) is the filtered image of (b), and (d) is the segmented image on which red color represents pores, blue color represents solid phase and green color represents unresolved phase.

With a digital rock model, many numerical methods can be used to do simulations, such as finite element method, finite difference method, finite volume method and lattice Boltzmann methods (LBM) (Sun et al., 2019). The rock properties, including porosity, permeability, electrical properties, elastic properties and multiphase flow properties can be obtained by simulations. Figure 5 shows an example of permeability calculation by LBM. The fluid flow situation in the connected pores of rock sample is clearly observed in this figure. According to Darcy's Law, the permeability can be then calculated with the velocity field.

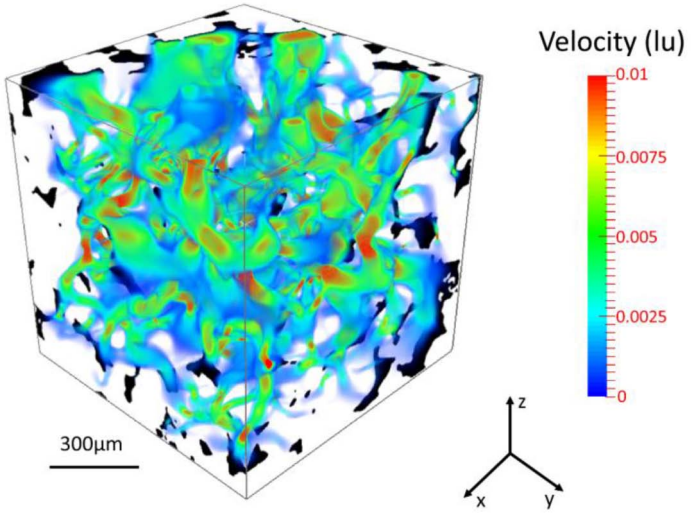

Figure 5—Example of fluid flow simulation with LBM for permeability calculation. The color bar shows the velocity with lattice unit. The image size is 300×300×300 voxels and image resolution is 4.8μm/voxel (Sun et al., 2019).

Result analysis includes digital image observations and analysis of simulated results. The combination of digital rock images, experimental data and numerical simulations provides accurate micro-scale pore structure information and macro-scale rock properties for reservoir study. Figure 6 shows an example of reservoir characterization by analyzing rock images, experiments and simulations. In this figure, the cemented fractures and the dissolved vugs can be clearly observed from thin sections (Figure 6 (a) and (c)), and their properties can be simulated from XCT images (Figure 6 (b) and (d)). Therefore, the fractures and vugs in this carbonate reservoir are well characterized by comprehensive analysis.

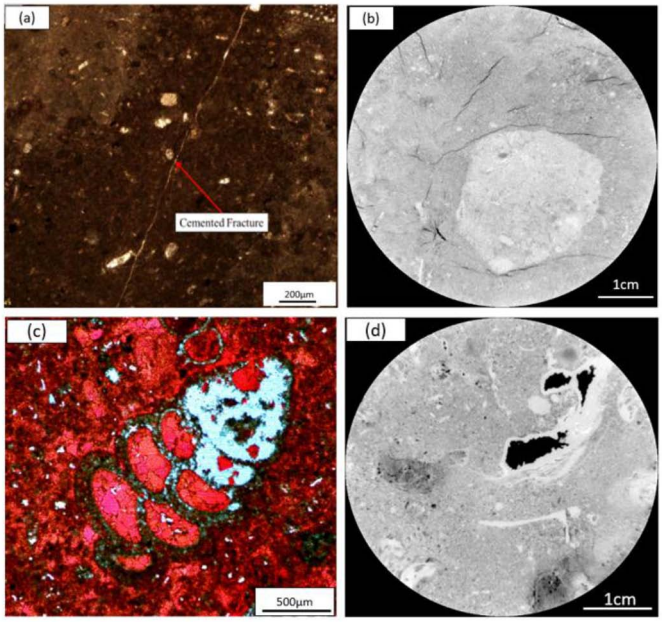

Figure 6—Example of combining observations of rock images, experiments and simulations for fracture and vug characterization of a carbonate reservoir: (a) and (b) are the thin section and XCT slice showing some cemented fractues; (c) and (d) are the thin section and XCT slice showing some dissolved vugs (Sun et al., 2018).

ML Methods in DRA

ML applications in DRA mainly focus on image segmentation, feature detection, rock image classification, and rock property prediction, and in general, the ML methods can be classified into three categories of unsupervised algorithms, supervised algorithms and ensemble classifiers (Chauhan et al., 2016a, b; Cheng and Guo, 2017; Tian et al., 2018; Sudakov et al., 2019).

The k-means clustering algorithm is one of unsupervised ML algorithms. This method is suitable for segmenting pores, grains and other phases from digital rock images (Sun et al., 2019). The basic idea is to initialize k centroid centers for dividing the dataset (pixel/voxel in digital rock image) into different cluster. K-means firstly assigns each datapoint to its closest centroid center; then, is iteratively updates each centroid to the mean of the constituent datapoints. Finally, the algorithm converges when there is no further change in the assignment of datapoints to the centroids. In other words, the convergence is obtained when the objective function $J_{k\text{-}mean}$ is minimized. $J_{k\text{-}mean}$ can be given by (MacQueen, 1967):

$$J_{k-means} = \sum_{i=1}^{m}\left(\sum_{j=1}^{n} \| x_i - c_j \|^2\right) \quad (1)$$

where x_i is the dataset, c_j is the centroid center, $\| x_i - c_j \|$ is the Euclidean distance between x_i and c_j.

The FCM algorithm is an extension of k-means. In this method, each datapoint can be member of multiple clusters with a membership value. FCM also involves minimizing the objective function (Jain et al., 1999; Jain, 2010):

$$J_{fcm} = \sum_{j=1}^{n}\sum_{i=1}^{m}\left(\mu_j(x_i)\right)^l \| x_i - c_j \|^2 \quad (2)$$

where c_j is the fuzzy cluster center, l is the fuzziness parameter (if l=1, FCM becomes k-means), $\mu_j(x_i)$ is the membership function, which represents the grade of membership of object x_i to cluster c_j and typically, μ_j values are between zero and one.

ANN is an information-processing paradigm that mimics the behavior of human brain (Chauhan et al., 2016b). This algorithm consists of input layer, hidden layer (one or more), and output layer, as shown in

Figure 7. In this figure, each circle represents a neuron that is connected to all neurons of adjacent layer. Each neuron represents a certain output function (activation function), and each connection between two neurons represents a weight value. ANN is suitable for modeling complex patterns and prediction problems.

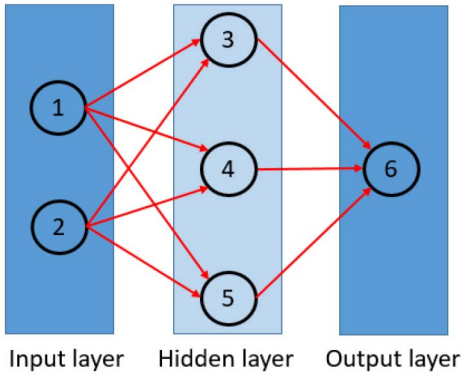

Figure 7—Schematic diagram of ANN.

The SOM algorithm is based on ANN but only consists of input layer and output layer (computational layer). It is another unsupervised ML algorithm. In this method, a competitive learning process is used. Specifically speaking, SOM arranges the neurons in one of the desired topologies and then uses a distance formula to determine the neuron position in this topology. By using a competitive learning process iteratively, a winner node will emerge. All the neurons in a defined neighborhood around the winner node are defined as a cluster (Kohonen, 1990). The winner node can be determined by the minimum Euclidean distance, which is obtained from the equations below (Chauhan et al., 2016b):

$$x(t) - m_c(t) = min_i \| x(t) - m_i(t) \| \tag{3}$$

$$m_i(t+1) = m_i(t) + \alpha(t)[x(t) - m_i(t)], \quad i \in N_c \tag{4}$$

$$m_i(t+1) = m_i(t), \quad i \notin N_c \tag{5}$$

where t is the current learning iteration, x is the input data, m_c is the winner, α is the time-varying learning rate, and N_c is the neighborhood of the winner.

A feedforward nueral network (FFNN) is one of the simplest ANNs. FFNN with one hidden layer can be used to solve many finite input-to-output mapping problems such as facial recognition and image segmentation (Hopfield, 1982). With more hidden layers, FFNN is suitable for complex clustering. The back propagation (BP) neural network algorithm is a multi-layer FFNN and is one of the most widely applied ANN models (Egmont-Petersen et al., 2002). This method keeps modifying all network parameters until a sufficiently well-fitting input–output function is found. To achieve the minimum error sum of square, BP algorithm uses the steepest descent method in which the back propagation is adopted for regulating the weight value.

The SVM is one of supervised ML algorithm based on the statistical learning theory. This method is suitable for classification and regression analysis. With non-linear kernel function, SVM is able to map the input dataset into the high-dimensional space for non-linear classification. The commonly used SVM classification model is linear separating hyperplane that is the simplest model in this family. Its function and support vectors around the separating hyperplane can be given by (Li et al., 2017):

$$w^T \cdot \mathrm{X} + b = 0 \tag{6}$$

$$\left| w^T \cdot \mathrm{X} + b \right| = 1 \tag{7}$$

where X is the input dataset, w is N-dimensional vector and b is the offset. w and b can be selected to maximize the margin around the separating hyperplane, as shown in Figure 8.

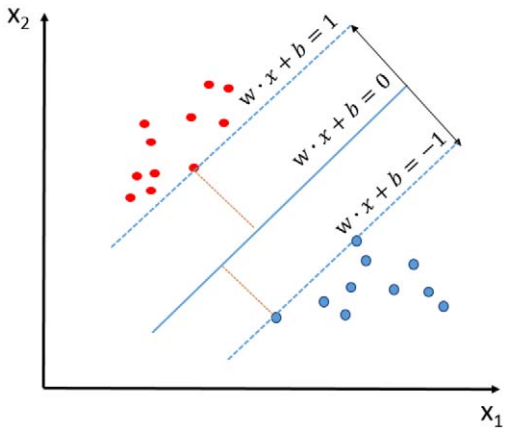

Figure 8—Schematic diagram of SVM hyperplane. The solid line (optimal hyperplane) in the middle has the maximum margin of separation between the two groups (red and blue points). The area dedicated by the black arrow is the maximum margin between two groups (Tian et al., 2018).

Figure 8 shows an optimal hypeplane illustration of SVM. In this figure, SVM determines the optimal hyperplane and the support vectors during model training and uses them to distinguish different groups.

Bragging and boosting are the two most frequently used ensemble classification methods. The bragging method is proposed based on the tree algorithm (Breinman 1996; Kohavi and Quinlan, 2002). This method, firstly, breaks down a large training dataset into many small datasets; then randomly selects samples from the training sets and uses them to train classifiers. During this process, a certain amount of misclassification will emerge, because different training set is used to train different classifier. The misclassified instances are collected in the next step, and the previous step is repeated until the misclassification errors are minimized. Freund and Schapire (1997) proposed the boosting concept, and Schapire et al. (1998) developed an AdaBoost algorithm that uses the whole dataset to train classifiers sequentially. The individual weights are then assigned to accurate and inaccurate classifiers. The inaccurate classifiers are then retrained until the weighted error is minimized.

Convolution neural network (CNN) is an important DL algorithm that is a subdivision of ML. CNNs have a similar structure to a regular ANN because they share similar weights and biases arranged in multiple hidden layers. However, CNNs contain the convolutional layers where many filters are iterated over the input tensor for learning. The features can be extracted automatically and accurately by using this method. CNNs have achieved a wide range of applications such as plant classification, facial recognition, and handwritten Chinese character recognition (Cheng and Guo, 2017). The convolutional layers (feature maps) are alternately connected with a number of pooling layers and can effectively characterize the essential features of the dataset. The feature map F_k can be obtained by (Karimpouli and Tahmasebi, 2019a, b):

$$F_k = (\sum_i W_{ki} * X_i) + d_k \tag{8}$$

where W_{ki} is the sub-kernel for ith channel, X_i is the ith input channel, * represents the convolutional operations that include continuous convolutions and discrete convolutions, and d_k is a bias term.

Case Study

Tight carbonate reservoirs, which are naturally fractured and vuggy, play an important role in the world's oil and gas production (Sun et al., 2018; Wang and Chen, 2019). Fractures, fissures, cracks and vugs, to certain extent, can enhance the flow potential of the reservoir, because they can improve the connectivity

within the matrix domain, but may reduce the overall productivity of the reservoir due to their low storage capacity (Belhaj et al., 2003a, b). It is hard to capture their characteristics, features and parameters in details (size, length, aperture, orientation, delineation, extension, etc.), from visual macroscopic observation of core samples only. Figure 9 shows images of rock samples from a tight and complex carbonate reservoir along with their XCT images. The averaged porosity of these rock samples is 10.83% and averaged permeability is 1.22mD. Most permeability values are less than 1mD, which makes these rock samples ideal representative of very tight UCRs.

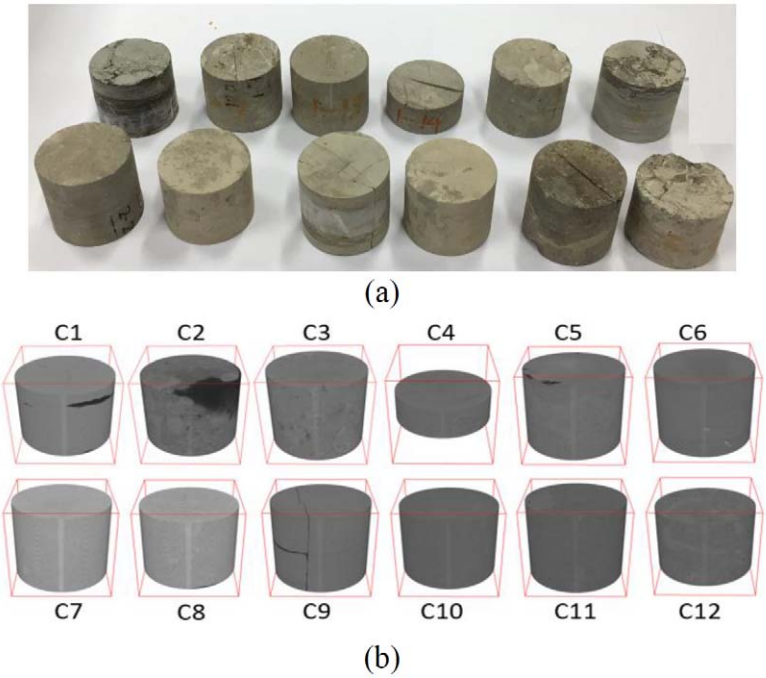

Figure 9—Photo of rock samples (a) and their XCT images (b).

Traditionally, fractures, fissures, cracks and vugs can be determined physically in the laboratory from retrieved rock samples through direct observation of the full core or cut plugs. The set back of this process is being destructive and does not tell anything about the subsurface characteristics of these very important properties of rock samples deeming it useless. However, DRA techniques provide details and analysis of the internal 3D structure and heterogeneity of the rock samples. In addition, this method is non-destructive and easier to perform. Figure 10 demonstrates the differences between the traditional observatory procedure and the state-of-the-art DRA method proposed in this study. In this figure, each image size is 1000×1000 pixels, and pixel size is 40μm. The features of fractures and vugs are observable and diagnosable by the naked eye.

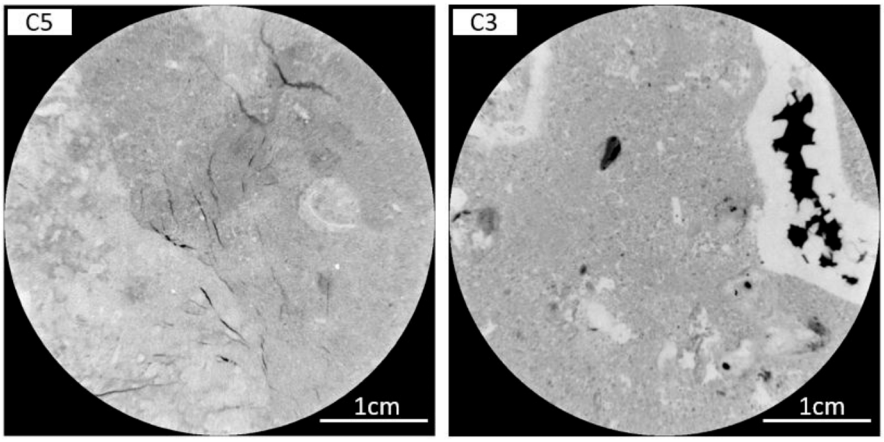

Figure 10—Image slices of two rock samples C5 (left) and C3 (right).

However, although fractures and vugs can be easily distinguished from each other with visual observation, their quantization is still a challenge. Especially, if there are thousands of digital rock images scanned from hundreds of core samples, it will be time-consuming and may result in many uncertainties. The better way is to use an automatic numerical method to extract the fractures and vugs and analyse their characteristics. In this case study, in order to characterise the fractures and the vugs qualitatively and quantitatively, the combination of DRA and ML is used. Figure 11 shows the workflow. For the image processing in this workflow, Mean filtering and k-means method are used for removing noise and image segmentation. Figure 12 shows an example of image processing. After image filtering and segmentation, the pores (red colour), the solid phase (blue colour) and some unresolved phase (green colour) are clearly resolved.

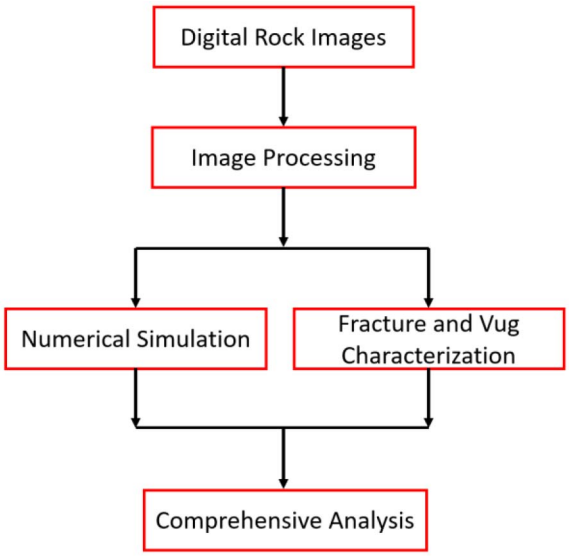

Figure 11—Workflow for fracture and/or vug characterization.

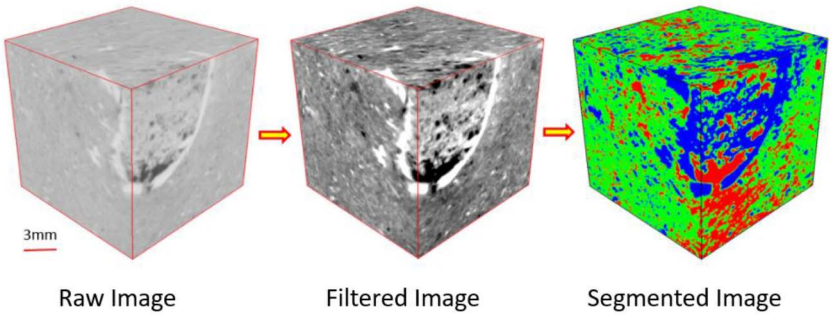

Figure 12—An example of image processing with mean filter and k-means method.

For numerical simulation, the porosity φ_s can be calculated by (Sun et al., 2018):

$$\varphi_s = \varphi_{XCT} + \varphi_{SEM} \cdot \frac{N_{unresolve}}{N_{total}} \tag{9}$$

where φ_{XCT} is the porosity calculated from XCT image, φ_{SEM} is the micro-porosity obtained from SEM image, and $\frac{N_{unresolve}}{N_{total}}$ is the ratio of unresolved phase volume to total volume of XCT image.

In order to distinguish fractures from vugs rock content, a SVM algorithm is used. The detailed background of this algorithm is available in the reference (Li et al., 2017). Figure 13 shows one example of identification of fractures and vugs in samples C1 and C3. In this figure, the red colour demonstrates the extracted fractures and the blue colour is the vugs representation.

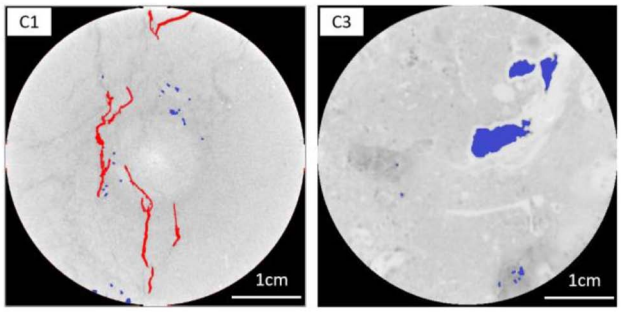

Figure 13—An example of identification of fractures and vugs with SVM algorithm

In order to classify these TZ rocks quantitatively, an improved carbonate rock index (ICRI) is used here (Li et al., 2017; Sun et al., 2018):

$$ICRI = \frac{(\frac{L_{fmax}}{L_{core}} + \frac{L_{fsum}}{D_{core}} + \frac{V_{vsum}}{V_{core}} + \varphi)}{(1 - S_{fr} \times S_{vr})} \times 100 \tag{10}$$

where $\mathbf{L_{fmax}}$ is the longest fracture in the longitudinal CT images; $\mathbf{L_{core}}$ is the core length; $\mathbf{L_{fsum}}$ is the sum of effective length larger than 5 mm; $\mathbf{D_{core}}$ is the core diameter; $\mathbf{V_{vsum}}$ is the volume of all vugs in the CT images; $\mathbf{V_{core}}$ is the total volume of the core; $\boldsymbol{\varphi}$ is the core porosity; $\mathbf{S_{fr}}$ is the area percent of fracture; $\mathbf{S_{vr}}$ is the area percent of vug.

Table 1 lists the results of porosity calculation and rock domain classification with ICRI. In this table, Φ_m is the measured porosity, Φ_s is the simulated porosity, and Φ_d is the relative difference between Φ_m and Φ_d. It can be seen that the simulated porosity are close to the measured porosity, which shows DRA accuracy. These rocks are classified into three types: fractured (ICRI > 10), vuggy (1≤ICRI≤10) and matrix (0 < ICRI < 1).

Table 1—Results of porosity calculation and rock domain classification

Sample	Porosity			Rock Domain Classification	
	Φ_m (%)	Φ_s (%)	Φ_d (%)	ICRI	Rock Domain
C1	1.30	1.80	38.46	15.53	Fractured
C2	14.60	13.80	5.48	2.57	Vuggy
C3	15.20	14.90	1.97	3.15	Vuggy
C4	14.40	13.70	4.86	12.86	Fractured
C5	9.20	9.10	1.09	16.84	Fractured
C6	11.40	10.90	4.39	0.34	Matrix
C7	13.20	12.80	3.03	0.56	Matrix
C8	7.90	7.30	7.59	0.48	Matrix
C9	13.10	12.10	7.63	0.37	Matrix
C10	15.40	14.30	7.14	0.39	Matrix
C11	12.20	11.30	7.38	2.76	Vuggy
C12	2.10	1.50	28.57	0.27	Matrix

It's worth saying that this method can be used in the same manner of porosity simulation described above to characterize other rock properties such as remaining RCA or SCAL properties.

Conclusion and Outlook

DRA is a very powerful tool for UCRs characterization, because it provides detailed information of pore structure and helps to study reservoirs by image observation and comprehensive analysis. A case study has been used to prove the validity of this promising technology and describes the procedure to use it. Indeed, it carries a great potential to be the ultimate solution for reservoir characterization with the support and complement of other efforts. AI techniques, in general, can be applied in DRA and reservoir characterization and they can help to free man from unnecessary mental and physical burdens. The combination of DRA and AI provides new insight for UCRs characterization. It's expected that this technology will grow fast and provide more applications in oil and gas exploration and production field.

Acknowledgments

The authors would like to thank Abu Dhabi National Oil Company (ADNOC) for sponsoring this project and providing the core data and the permission for publication.

References

Al-Anazi A. and Gates I. D. 2010. A Support Vector Machine Algorithm to Classify Lithofacies and Model Permeability in Heterogeneous Reservoirs. *Engineering Geology*, **114**, pp. 267–227.

Azamathulla H. M. and Wu F. C. 2011. Support Vector Machine Approach for Longitudinal Dispersion Coefficients in Natural Streams. *Applied Soft Computing*, **11** (2), PP. 2902–2905.

Belhaj, H.A., Agha, K.R., Nouri, A.M., Butt, S.D., and Islam, M.R. 2003a. Numerical and Experimental Modeling of Non-Darcy Flow in Porous Media. Proceedings of the SPE Latin American and Caribbean Petroleum Engineering Conference, Port-of-Spain, Trinidad, West Indies, 27-30 April.

Belhaj, H.A., Agha, K.R., Nouri, A.M., Vaziri, H.F. and Islam, M.R. 2003b. Numerical Simulation of Non-Darcy Flow Utilizing the New Forchheimer's Diffusivity Equation. Proceedings of the Middle East Oil Show, pp. 647–656.

Boixader D. 2017. Special Issue on Pattern Recognition Techniques in Data Mining. *Pattern Recognition Letters*, **93**, pp. 1–2.

Breiman L. 1996. Bagging Predictors. *Machine Learning*, **24** (2), PP. 123–140.

Budennyy S., Pachezhertsev A., Bukharev A., Erofeev A., and Mitrushkin D. 2017. Image Processing and Machine Learning Approaches for Petrographic Thin Section Analysis. Paper Presented at the SPE Russian Petroleum Technology Conference held in Moscow, Russia, 16-18 October.

Calhoun V. 2018. Editorial on Special Issue: Machine Learning on MCI. *Journal of Neuroscience Methods*, **302**, pp. 1–2.

Chauhan S., Rühaak W., Anbergen H., Kabdenov A., Freise M., Wille T., and Sass I. 2016a. Phase Segmentation of X-ray Computer Tomography Rock Images Using Machine Learning Techniques: An Accuracy and Performance Study. *Solid Earth*, **7**, PP. 1125–1139.

Chauhan S., Rühaak W., Khan F., Enzmann F., Mielke P., Kersten M., Sass I. 2016b. Processing of Rock Core Microtomography Images: Using Seven Different Machine Learning Algorithms. *Computers & Geosciences*, **86**, pp. 120–128.

Chen L., Zhang L., Kang Q. J., Yao J., Tao W. Q. 2015. Nanoscale Simulation of Shale Transport Properties Using the Lattice Boltzmann Method Permeability and Diffusivity. *Scientific Reports 5, Article Number: 8089*.

Cheng G. and Guo W. 2017. Rock Images Classification by using Deep Convolution Neural Network. Journal of Physics: Conference Series, **887**, 012089.

Criminisi A. 2016. Machine Learning for Medical Images Analysis. *Medical Image Analysis*, **33**, pp. 91–93.

Das S. K. and Basudhar P. K. 2008. Prediction of Residual Friction Angle of Clays Using Artifical Neural Network. *Engineering Geology*, **100** (3-4), pp. 142–145.

Dong H. 2007. *Micro-CT Imaging and Pore Network Extraction*. PhD Thesis, London: Imperial College, pp. 1–80.

Dong H., and Blunt M. 2009. Pore-network Extraction from Micro-Computerized-Tomography Images. *Physical Review E*, **80**, 036307.

Egmont-Petersen M., deRidder D., Handels H. 2002. Image Processing with Neural Networks a Review. *Pattern Recognition*, **35**(10), PP. 2279–2301.

Esmaili S. and Mohaghegh S. D. 2016. Full field reservoir modeling of shale assets using advanced data-driven analytics. *Geoscience Frontiers*, **7** (1), pp. 11–20.

Fathi E., Tinni A., Akkutlu I.Y. 2012. Correction to Klinkenberg slip theory for gas glow in nano-capillaries. *International Journal of Coal Geology*, **103**, pp.51–59.

Freund Y. and Schapire R.E. 1997. A Desicion-Theoretic Generalization of On-Line Learning and an Application to Boosting. *Journal of Computer and System Sciences*, **55** (1), PP. 23–37.

Garg A., Garg A., Tai K., and Sreedeep S. 2014a. An Integrated SRM-multi-gene Genetic Programming Approach for Prediction of Factor of Safety of 3-D Soil Nailed Slopes. *Engineering Applications of Artificial Intelligence*, **30**, pp. 30–40.

Garg A., Tai K., Savalani M. M. 2014b. State-of-the-Art in Empirical Modeling of Rapid Prototyping Processes. *Rapid Prototyping Journal*, **20** (2), pp. 164–178.

Garg A. Tai K., Gupta A. K. 2014c. A Modified Multi-gene Genetic Programming Approach for Modelling True Stress of Dynamic Strain Aging Regime of Austenitic Stainless Steel 304. *Meccanica*, **49** (5), pp. 1193–1209.

Gibert K., Horsburgh J. S., Athanasiadis I. N., Holmes G. 2018a. Preface to the Thematic Issue on Environmental Data Science. Applications to air quality and water cycle. *Environmental Modelling & Software*, **106**, pp. 1–3.

Gibert K., Horsburgh J. S., Athanasiadis I. N., Holmes G. 2018b. Environmental Data Science. *Environmental Modelling & Software*, **106**, pp. 4–12.

Helmy T., Fatai A., and Faisal Kanaan. 2010. Hybrid Computational Models for the Characterization of Oil and Gas Reservoirs. *Expert Systems with Applications*, **37**, pp. 5353–5363.

Hopfield J. J. 1982. Neural Networks and Physical Systems with Emergent Collective Computational Abilities. Proceedings of the National Academy of Sciences of the United States of America, **79**(8), PP. 2554–2558.

in't Hout K., Itkin A., Sydow L. V., Toivanen J., 2018. Special Issue—Computational and Algorithmic Finance. *Journal of Computer Science*, **24**, pp. 180–181.

Jain A. K., 2010. Data Clustering: 50 Years beyond K-means. *Pattern Recognition Letters*, **31**(8), pp. 651–666.

Jain A. K., Murty M. N., and Flynn P.J. 1999. Data Clustering: A Review. *ACM Computing Surveys*, **31**(3), pp. 264–323.

Karimpouli S. and Tahmasebi P. 2019a. Image-based Velocity Estimation of Rock Using Convolutional Neural Networks. *Neural Networks*, **111**, 89–97.

Karimpouli S. and Tahmasebi P. 2019b. Segmentation of Digital Rock Images Using Deep Convolutional Autoencoder Networks. *Computers and Geosciences*, **126**, pp. 142–150.

Kohavi R. and Quinlan J. R. 2002. *Handbook of Data Mining and Knowledge Discovery*. Oxford University Press, Inc., NewYork, NY, USA, pp. 1–100.

Kohonen T. 1990. The Self-Organizing Map. Proceedings of the IEEE, **78**(9), 1464–1480.

Lary D. J., Alavi A. H., Gandomi A. H., and Walker A. L. 2016. Machine learning in geosciences and remote sensing. *Geoscience Frontiers*, **7**, pp. 3–10.

Li B., Tan X., Wang F., Lian P., Gao W., and Li Y. 2017. Fracture and vug characterization and carbonate rock type automatic classification using X-ray CT images. *Journal of Petroleum Science and Engineering*, **153**, pp. 88–96.

Liu T, Jin X and Wang M. 2018. Critical Resolution and Sample Size of Digital Rock Analysis for Unconventional Reservoirs. *Energies*, **11**(7), pp. 1–15.

Ma Z., Leung J. Y., and Zanon S. 2018. Integration of Artificial Intelligence and Production Data Analysis for Shale Heterogeneity Characterization in Steam-Assisted Gravity-Drainage Reservoirs. *Journal of Petroleum Science and Engineering*, **163**, PP. 139–155.

MacQueen, J. 1967. Some Methods for Classification and Analysis of Multi-variate Observations. Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Volume **1**: Statistics, pp. 281—297.

McCoy J. T. and Auret L. 2019. Machine Learning Applications in Minerals Processing: A Review. *Minerals Engineering*, **132**, pp. 95–109.

Mishra P K. and Akbar B. H. 2015. Application of Digital Rock Physics and FIBSEM Imaging to Unconventional Reservoir Characterization in the Najmah Formation. Paper Presented at the Third EAGE Workshop on Rock Physics held in Istanbul, Turkey, 15-18 November.

Patel A. K. and Chatterjee S. 2016. Computer Vision-Based Limestone Rock-Type Classification Using Probabilistic Neural Network. *Geoscience Frontiers*, **7** (1), pp. 53–60.

Saad B., Negara A., and Ali S.S. 2018. Digital Rock Physics Combined with Machine Learning for Rock Mechanical Properties Characterization. Paper Presented at the Abu Dhabi International Petroleum Exhibition & Conference held in Abu Dhabi, UAE, 12-15 November.

Samui P. 2008. Support Vector Machine Applied to Settlement of Shallow Foundations on Cohesionless Soils. *Computers and Geotechnics*, **35** (3), pp. 419–427.

Sarlin P., Björk K. M. 2017. Machine Learning in Finance—Guest Editorial. *Neurocomputing*, **264**, pp. 1.

Schapire R. E., Freund Y., Bartlett P., and Lee W.S. 1998. Boosting the Margin: A New Explanation for the Effectiveness of Voting Methods. *Annals of Statistics*, **26** (5), pp. 1651–1686.

Shahin M. A., Jaksa M. B., Maier H. R. 2001. Artificial Neural Network Applications in Geotechnical Engineering. *Australian Geomechanics*, **36** (1), pp. 49–62.

Silver D., Huang A., Maddison C. J., Guez A., Sifre L., van den Driessche G., Schrittwieser J., Antonoglou I., Panneershelvam V., Lanctot M., Dieleman S., Grewe D., Nham J., Kalchbrenner N., Sutskever I., Lillicrap T., Leach M., Kavukcuoglu K., Graepel T., Hassabis D. 2016. Mastering the game of Go with deep neural networks and tree search. *Nature*, **529**, pp. 484–489.

Silver D., Schrittwieser J., Simonyan K., Antonoglou I., Huang A., Guez A., Hubert T., Baker L., Lai M., Bolton A., Chen Y., Lillicrap T., Hui F., Sifre L., van den Driessche G., Graepel T., Hassabis D. 2017. Mastering the Game of Go without Human Knowledge. *Nature*, **550**, pp. 354–359.

Srisutthiyakorn N. 2016. Permeability Prediction of 3-D Binary Segmented Images using Neural Networks. *SEG Technical Program Expanded Abstracts*, https://doi.org/10.1190/segam2016-13972613.1.

Sudakov O., Burnaev E., and Koroteev D. 2019. Driving Digital Rock towards Machine Learning: Predicting Permeability with Gradient Boosting and Deep Neural Networks. *Computers and Geosciences*, **127**, 91–98.

Sun H., Belhaj H., Bera A., Liu L., Tao G., and Nizamuddin S. 2018. Digital Rock Analysis: A Case Study on Characterization of Capillary Transition Zones in Carbonate Reservoirs. Paper Presented at 80th EAGE Conference & Exhibition, Copenhagen, Denmark, 11-14 June 2018.

Sun H., Belhaj H., Tao G., Vega S., and Liu L. 2019. Rock Properties Evaluation for Carbonate Reservoir Characterization with Multi-Scale Digital Rock Images. *Journal of Petroleum Science and Engineering*, **175**, pp. 654–664.

Sun H., Tao G., Vega S., and Al-Suwaidi A. 2017a. Simulation of Gas Flow in Organic-Rich Mudrocks Using Digital Rock Physics. *Journal of Natural Gas Science and Engineering*, **41**, pp. 17–29.

Sun H., Vega S., Tao G. 2015. Simulation of Shale Gas Flow in Nano Pores with Parallel Lattice Boltzmann Method.Paper presented at the 77th EAGE Conference & Exhibition, IFEMA Madrid, Spain, 1-4 June 2015.

Sun H.; Yao J.; Cao Y.C., Fan D. Y.; Zhang L. 2017b. Characterization of Gas Transport Behaviors in Shale Gas and Tight Gas Reservoirs by Digital Rock Analysis. *International Journal of Heat and Mass Transfer*, **104**, pp. 227–239.

Sutton R.S. and Barto A.G. 2018. *Reinforcement Learning: An Introduction (Second Edition)*, MIT Press, Cambridge, MA, pp. 1–60.

Tahmasebi P., Javadpour F., and Sahimi M. 2017. Data Mining and Machine Learning for Identifying Sweet Spots in Shale Reservoirs. *Expert Systems With Applications*, **88**, PP. 435–447.

Tian X. and Daigle H. 2019. Preferential Mineral-Microfracture Association in Intact and Deformed Shales Detected by Machine Learning Object Detection. *Journal of Natural Gas Science and Engineering*, **63**, pp. 27–37.

Tian X., Daigle H., and Jiang H. 2018. Feature Detection for Digital Images Using Machine Learning Algorithms and Image Processing. Paper Presented at the Unconventional Resources Technology Conference held in Houston, Texas, USA, 23-25 July. URTeC: 2886325.

Viswanathan R. and Samui P. 2016. Determination of rock depth using artificial intelligence techniques. *Geoscience Frontiers*, **7** (1), pp. 61–66.

Wang S. and Chen S. 2019. Insights to Fracture Stimulation Design in Unconventional Reservoirs Based on Machine Learning Modeling. *Journal of Petroleum Science and Engineering*, **174**, PP. 682–695.

Wu Y., Schuster M., Chen Z., Le Q. V., Norouzi M., Macherey W., Krikun M., Cao Y., Gao Q., Macherey K., Klingner J., Shah A., Johnson M., Liu X., Kaiser Å., Gouws S., Kato Y., Kudo T., Kazawa H., Stevens K., Kurian G., Patil N., Wang W., Young C., Smith J., Riesa J., Rudnick A., Vinyals O., Corrado G., Hughes M., Dean J. 2016.*Google's Neural Machine Translation System: Bridging the Gap between Human and M*

Xu Y., Sun K., Xie H., Zhong X., Mirto E., Feng Y., and Hong X. 2018. Borehole Resistivity Measurement Modeling Using Machine-Learning Techniques. *Petrophysics*, **59** (6), pp. 778–785.

Yu H., Rezaee R., Wang Z., Han T., Zhang Y., Arif M., and Johnson L. 2017. A New Method for TOC Estimation in Tight Shale Gas Reservoirs. *International Journal of Coal Geology*, **179**, pp. 269–277.

Zhang X. L., Xiao L. Z., Shan X.W., Guo L. 2014. Lattice Boltzmann Simulation of Shale Gas Transport in Organic Nano-Pores. *Scientific Reports*, **4**, Article Number: 4843.

Zou C., Zhang G., Yang Z., Tao S., Hou L., Zhu R., Yuan X., Ran Q., Li D., and Wang Z. 2013. Concepts, characteristics, potential and technology of unconventional hydrocarbons: On unconventional petroleum geology. *Petroleum Exploration and Development*, **40**(4), pp. 413–428.

Zou C. 2017. *Unconventional Petroleum Geology (Second Edition)*. Petroleum Industry Press, Beijing, pp. 1–480.