Computational Geosciences (2020) 24:571–580
https://doi.org/10.1007/s10596-019-09886-9

ORIGINAL PAPER

Deep learning–driven permeability estimation from 2D images

Mauricio Araya-Polo[1] · Faruk O. Alpak[1] · Sander Hunter[1] · Ronny Hofmann[1] · Nishank Saxena[1]

Received: 13 November 2018 / Accepted: 22 August 2019 / Published online: 18 November 2019

Abstract

Current micro-CT image resolution is limited to 1–2 microns. A recent study has identified that at least 10 image voxels are needed to resolve pore throats, which limits the applicability of direct simulations using the digital rock (DR) technology to medium-to-coarse–grained rocks (i.e., rocks with permeability $>$ 100 mD). On the other hand, 2D high-resolution colored images such as the ones obtained from transmitted light microscopy delivers a much higher resolution (approximately 0.6 microns). However, reliable and efficient workflows to jointly utilize full-size 2D images, measured 3D core-plug permeabilities, and 2D direct pore-scale ow simulations on 2D images within a predictive framework for permeability estimation are lacking. In order to close this gap, we have developed a state-of-the-art deep learning (DL) algorithm for the direct prediction of permeability from 2D images. We take advantage of the computing graphics processing units (GPUs) in our implementation of this algorithm. The trained DL model predicts properties accurately within seconds, and therefore, provide a significant speeding up simulation workflow. A real-life dataset is used to demonstrate the applicability and versatility of the proposed method.

Keywords Deep learning · Machine learning · Permeability · Thin-section · Digital rock · Sandstone · Reservoir properties

1 Introduction

One of the critical properties of a reservoir is permeability. The permeability of a rock is one of the key controls on the rate at which oil can be produced, thus impacting decisions ranging from the economic viability of a prospect to the field development planning and facilities sizing. Obviously, having an accurate information of brine or relative permeability of a reservoir is crucial. However, such measurements are often relatively time consuming to make in a laboratory environment (see Fig. 1), and they are not always possible. For instance, re-visiting legacy samples that have already been impregnated and made into thin sections. All above makes the proposed approach even more valuable, at least as first-order approximation.

In an effort to reduce the time between sample acquisition and having the data necessary to decisions, researchers have turned to computational simulations, collectively referred to as the digital rock physics (DRP). DRP focuses on using image volumes, generally microcomputed tomography (μCT), of rock samples to simulate accurate pore-scale flow physics. These simulations require several inputs that constrain the applicability of DRP. First, an accurate depiction of the 3D geometry of the sample of interest is required (e.g., Fig. 1). Although this is a logical requirement, in practice, even high-resolution μCT images can omit important details. These details may include sub-resolution pore throats, or flow pathways, or sub-resolution clays obstructing flow pathways. The implications of this fact are that while DRP relies on μCT images with this limited resolution, a methodology for addressing the issue of image resolution is necessary to derive an accurate answer. Second, DRP relies on algorithms to describe and replicate the physical behavior of fluids moving through a rock. Although these algorithms are rigorously tested and benchmarked, assumptions and boundary conditions are still necessary, which may impact the accuracy of the simulation. DRP has shown significant promise. However, as users push the bounds of current technology, the limits of μCT imaging are reached, while computational times grow.

Current micro-CT image resolution is limited to $\approx$ 1–2 microns. A recent study has identified that at least 10 image voxels are needed to resolve pore throats, which limits the applicability of direct simulations using the digital

Dr. Mauricio Araya-Polo is now with Total R&T US.

✉ Mauricio Araya-Polo
mauricio.araya@shell.com

[1] Shell International Exploration & Production Inc., 3333 Highway 6 South, Houston, Texas, 77082, USA

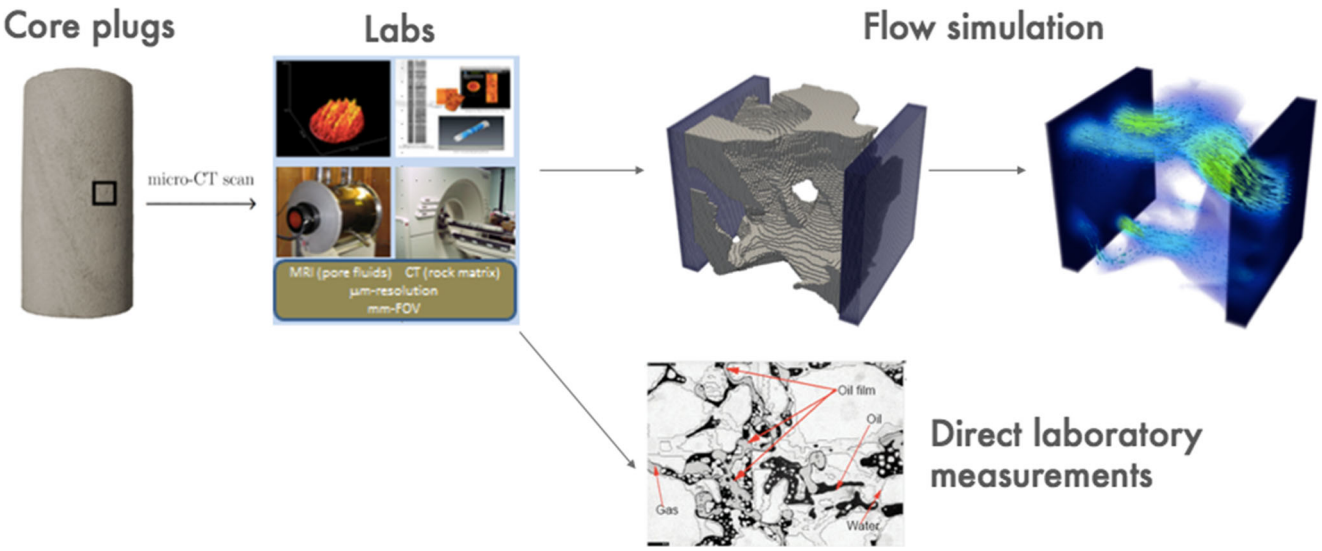

Fig. 1 From core plugs the current properties estimation workflow can take the laboratory path and/or the simulation path, both with advantages and shortcomings. Major concern are the time needed to compute the property, and accuracy of it

rock (DR) technology to medium-to-coarse–grained rocks (i.e., rocks with permeability $\geq$ 100 mD)(see [1]). On the other hand, 2D high-resolution colored images such as the ones obtained from transmitted light microscopy delivers a much higher resolution ($\approx$ 0.6 microns). However, reliable and efficient workflows to jointly utilize full-size SEM images, measured 3D core-plug permeabilities and 2D direct pore-scale flow simulations on SEM images within a predictive framework for permeability estimation are lacking. In order to close this gap, we introduce a deep learning (DL) algorithm for the direct prediction of permeability from 2D, plane-polarized light photomicrograph images of thin-sectioned rock samples with permeabilities in a range between 50 and 1100 millidarcies (mD).

The ultimate vision driving this work is to minimize the turn around between sample acquisition and derivation of rock properties. In the case of permeability, we envision imaging thin-sectioned cutting samples to produce a pseudo-permeability log over the target intervals. The advantages of this approach are two-fold. First, sample acquisition is generally a time-consuming and expensive proposition, involving an additional trip into the hole with a new tool. Second, cuttings are not only an inevitable byproduct of the drilling process but also they are generally retained in storage long after well completion. In instances where detailed information regarding rock properties is of interest after the completion of a well, the proposed workflow offers a simple and cost-effective means to this end.

In this paper, we describe in detail, the first step in achieving the vision described above by testing the viability of using 2D thin section images paired with permeability measurements to estimate permeability on other rock samples via DL. To our knowledge, this is the first work where laboratory-measured permeability data (on the same samples from which thin-section images are taken) for 135 different rock samples from real-life producing reservoirs (11) are used as label. Also, this is the first time the input to the DL systems is not a segmented image, but rather a high-resolution image, thus, allowing the DL system to deduce the relevant features by itself, which is an idea at the core of the DL paradigm. Nonetheless, two works from the literature standout in terms of prediction accuracy and similarly innovative approaches. First, in Sudakov et al. [2] work permeability is computed by a Pore Network Model [3] on sub-sample images in order to be used as labels, from only 1 rock sample (Berea sandstone) model, which makes us wonder about how generalizable are their permeability predictions for other rocks. Basically [2] work purposes is to compare off-the-shelf DL architectures, to that at the end they use a custom metric. Second, in Srisutthiyakorn's work [4] permeability for labels is computed by direct numerical simulation with a Lattice Boltzmann-based method solver) on sub-samples of 2 rock samples (Fontainebleau and Berea sandstones). Srisutthiyakorn [4] work is primarily devoted to 3D aspects of the problem. Also, in both [2, 4] the input data set is a segmented version of the rock, wherein [2] Minkowski functionals are used as features.

In our approach, physics is contained in the measurement dataset with which we train the deep-learning algorithm. Therefore, when putting together the dataset used in this work, we paid special attention such that permeability measurements (that correspond to the thin-section images in the training dataset) were either performed with a consistent protocol or corrected to be consistent with the mainstream measurement protocol used for the majority

of the datapoints. The measurement protocol includes provisions for ensuring that the permeability measurements were free of strong rock-brine interactions (salinity effects). Moreover, if air permeability was measured (in a small subset of the dataset) instead of brine permeability, we made sure that these measurements were carefully corrected for the non-Darcy flow effects (e.g., Klinkenberg correction) such that these measurements are on equal footing with the rest of the dataset. The images and measured permeability data contain sufficient information to resolve the absolute (single-phase laminar flow) permeability with reasonably good accuracy as it will be demonstrated later in the paper.

Our initial results show that over the lithologies and permeabilities used, this approach is a viable means for quickly determining permeability. The trained DL model predicts properties accurately within seconds, and therefore, provide a significant speeding up simulation workflows.

The outline of the paper is as follows: First we introduce our machine learning approach, starting with the data labeling and then the deep learning method. After an outline of the implementation, we introduce prediction results for rock samples. The paper concludes with a summary of the main results, as well as an outlook for future work.

1.1 Geologic description methodology and sample description

Samples used in this study come from 11 clastic reservoirs across Shell's asset portfolio and cover a range of lithofacies that might be encountered in clastic basins. All samples are sandstones, and range encompass a range of grain sizes from very fine sand to coarse sand (see Fig. 2). Similarly, this dataset encompasses samples across a spectrum of sorting from well sorted to very poorly sorted. Finally, the sample set was chosen such that a range of permeabilities was covered from between $\approx$ 50 and 1000 mD.

Samples were prepared into thin slices of rock, called thin sections, that are primarily 1 inch in diameter and 30 microns in thickness. To preserve texture and structure, each rock sample is impregnated with a blue-dyed epoxy prior to cutting, grinding, and polishing. This is the standard sample preparation technique, and there is no evidence that this approach damages or significantly alters the texture within the sample. After cutting and polishing the samples, they were imaged using a transmitted plane-polarized light microscope at 10×, resulting in images with a resolution of 0.65 microns per pixel. A scanning stage and stitching software were used to create images of the entire thin section. From these images, two 9000×9000 pixel areas deemed representatives of the sampled lithology were cut from the larger image. One image from each sample is coupled with the sample permeability measured in the lab, and the second is withheld for blind testing. Care was taken to ensure that samples included in the study were homogeneous on the scale of the thin section, and showed no signs of damage during sample preparation (see Fig. 3).

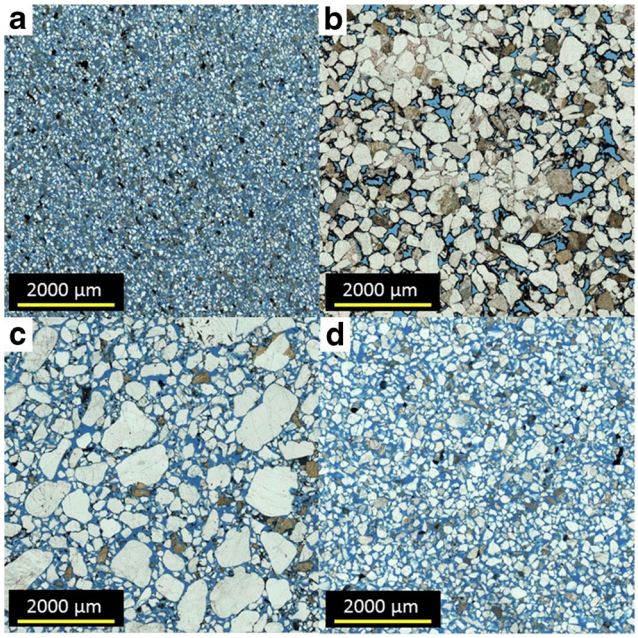

Fig. 2 Plane polarized photomicrographs of samples showing a variety of textures included in the training dataset. In these images, blue is epoxy filling the inter-granular pore space. Samples range in grain size from very fine sand in (**a**), to very coarse sand (**c**). Porosities and permeabilities are as follows: **a** porosity = 30.5 porosity units (pu), permeability = 70 mD, **b** porosity = 21.3 pu, permeability = 1189 mD, **c** porosity = 22.7 pu, permeability = 929 mD, **d** porosity = 31 pu, permeability = 980 mD

2 Method

Deep learning is part of a new wave of powerful machine learning (ML) algorithms [6, 7], which provide state-of-the-art results in numerous computer vision, speech processing, and artificial intelligence problems. In particular, deep neural networks (DNN) provide excellent results for imaging inverse problems such as compressed sensing [8], and X-ray computed tomography [9, 10] or seismic tomography [11]. In addition, according to the universal approximation theorem [12], DNNs can be used to approximate any arbitrary continuous function up to a specified accuracy. For these reasons, there is great promise in using this approach to approximate complex functions that are highly non-linear. One key dependency of the DL approach is the availability of labeled data, i.e., data for which the expected result is known. To fulfill the data requirement Shell legacy collection of thin sections provide a rich dataset. Not all images from thin sections are labeled, that is a time consuming task usually carried out in the

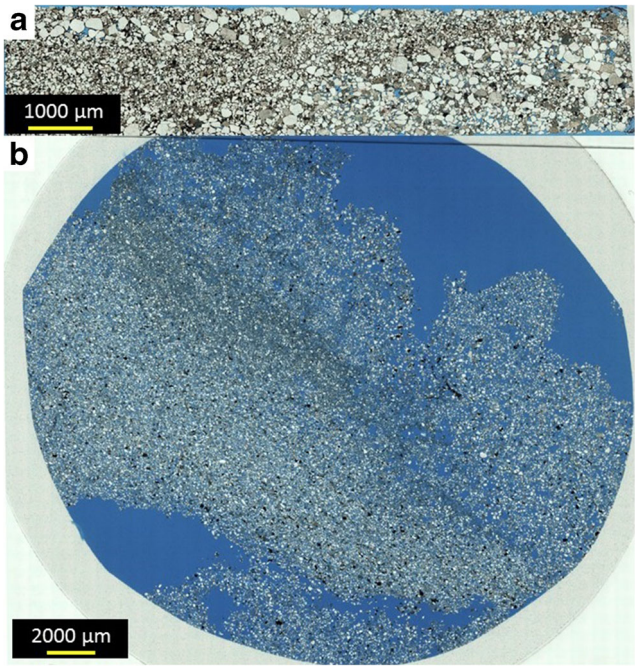

Fig. 3 **a** A plane-polarized light photomicrograph of a sample with significant heterogeneity in both grain size and sorting of grains. Both of these attributes have a strong impact on the permeability of a rock, which means that the lab permeability measurement for this sample is a combination of the permeability of the fine-grained, moderately well-sorted sand, and the coarse-grained, poorly sorted sand. The relative contributions of each are unclear, and thus not suitable for including in the deep learning training dataset. **b** A plane-polarized light photomicrograph of a sample of unconsolidated sand where the original textures of the sand were not retained during sample acquisition and preparation

laboratory (see Fig. 4). Alternatively, permeability labels can be generated by using direct flow simulation.

2.1 Machine learning

The machine learning approach is described here. First, we describe the training process. Our approach is supervised learning, thus input and labels needs to be available at training time. In order to fulfill that condition, we prepare a set of high-resolution images from thin sections for which we have laboratory measurements Fig. 4. Second, we describe the prediction or inference process, which is similar to what a regular user experience would be.

ML classification is not the task we need to solve to tackle the problem at hand, since permeability is a property that can not be derived by just identifying groups of grains. Further, segmentation will not provide the answer since identifying shapes of pore throats or grain sorting is not enough to estimate permeability. Fundamentally, a mapping in a continuous manifold between the thin-sectioned images and the labels needs to be learned. Therefore, our approach is regression (as depicted in Fig. 5).

For very simple cases, analytical expressions for permeability calculation can be derived (for instance, see Fig. 6), but the complexity of real rock images (3) makes any analytical approach a futile endeavor.

The following expression describe our regression approach as is commonly expressed in the machine learning community:

$$\hat{\theta} = \arg\min_{\theta} \frac{1}{N} \sum_{i=1}^{N} L(l_i, P(X_i, \theta)), \tag{1}$$

where $P(X, \theta)$ is the *permeability* operator, parameterized by the coefficients vector θ, X is the input to the permeability operator, and its output is the predicted permeability $\hat{p}$. In machine learning terminology, X is known as the input feature and l is known as the label. The loss function $L(l_i, \hat{p}_i)$ measures the difference between the ground truth permeability and its predicted value. The loss function employed in this work is the squared error $L(l_i, \hat{p}_i) = (l_i - \hat{p}_i)^2$, which is frequently used in regression problems. Replacing the generic loss function with the squared-error loss, we can express the optimization problem in (1) as

$$\hat{\theta} = \arg\min_{\theta} \frac{1}{N} \sum_{i=1}^{N} (l_i - P(X_i, \theta))^2. \tag{2}$$

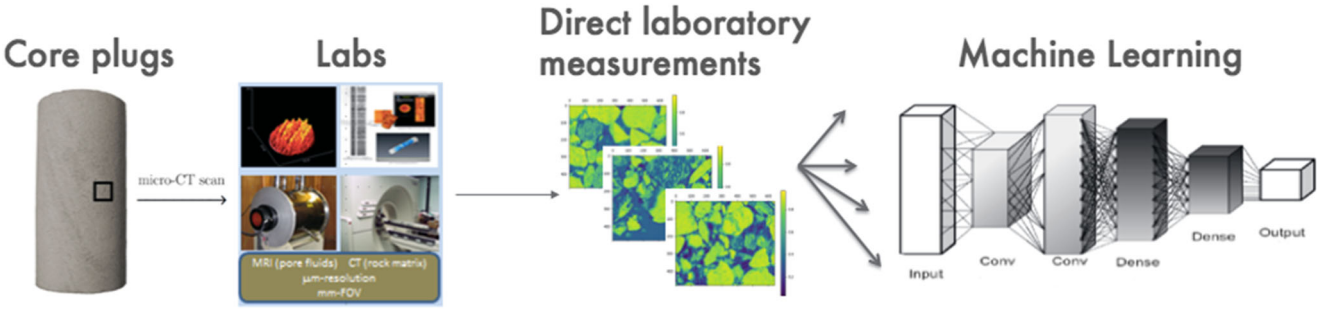

Fig. 4 Using laboratory measurements as labels for the training process and as ground-truth for inference performance. The latter will be described in the results section

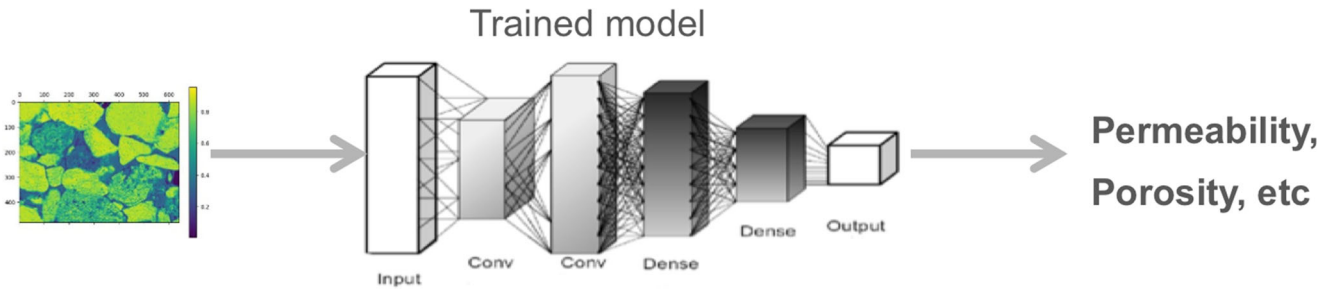

Fig. 5 General depiction of the proposed machine learning approach. Where the input is a set of thin-section images and the output is permeability prediction per image. Notice that the porosity prediction or any other property is not covered in this work

This familiar regression problem is usually solved with a stochastic gradient descent approach which iteratively updates the coefficients of θ. Expression 2 can be seen as an inverse problem, but the solution of the system is reached by an approach very different wrt to the conventional one. The latter is a deterministic optimization problem, the ML approach profits from a learning process in which a statistical mechanism helps to minimize the loss function by exploiting the training dataset. Therefore, the Achilles heel of this ML approach is the quality and labeling of the data used for training.

In this work, the permeability predictor $P(X, \theta)$ is implemented as a convolutional neural network (CNN) composed of *layers* of weighted nodes parameterized by θ. The input to the network is connected to the input layer which is followed by a varying number of convolutional layers. The output scalar is a permeability prediction given a input feature vector X, in our case, a thin-section image. It is worth noticing that the output can be a vector that holds simultaneous predictions of different parameters given the corresponding labels; in this work, we will focus on permeability. Neural networks are trained with examples per the statistical learning approach in which the correct output (label) is known for a given input, and the weight parameters in the nodes of the network update due to the minimization of the error between the prediction and true value. Figure 7 shows the workflow used to train the weights of the CNN.

Once the DL model is fully trained, it can be use to predict permeability when expose to unseen high-resolution images. This step is straightforward, and it costs very little in terms of computing capacity, which is the grand promise of this whole approach wrt laboratory work or simulation.

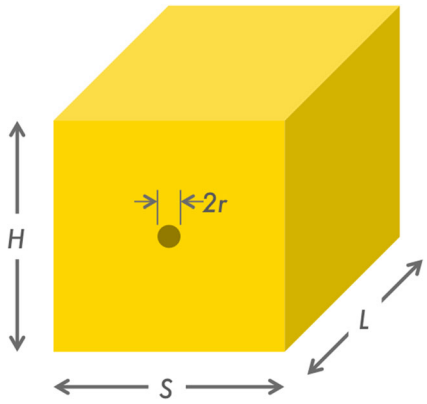

Fig. 6 For this simple case, the analytical expression for permeability is known as $k = \frac{\pi \times r^2}{8 \times H \times s}$. And k varies widely (orders of magnitude) given small r changes. For instance, if diameter is 0.8 mm then $k =$ 10 mD, but if diameter is 2.5 mm then $k = 1000$ mD, two orders of magnitude change

3 Results

In this section, we will first introduce implementation details, then actual training and testing results are presented and analyzed. We will cover two main cases. In the first case, all samples used in training and testing belong to the same thin section. In the second case, which is closer to the expected deployment mode, thin-sections from 11 different reservoirs from Gulf of Mexico fields are used.

3.1 Implementation

The most important aspect of the implementation is the neural network architecture. We built two different architectures, one for each case. The first network is composed of seven convolutional layers followed by two dense layers. The second network is inspired by the U-Net architecture [13]; therefore, after a series of downscaling convolutional layers (contraction), the architecture have a symmetrical series of upscaling convolutional layers (expansion) that end in a couple of dense layers. The concentration section share intermediate feature maps with the expansion section (see Fig. 8.(top-left)).

Both networks share the following characteristics:

- Relatively small networks therefore in the order of few millions of parameters to learn, good for rapid investigation and hyperparameter tuning.

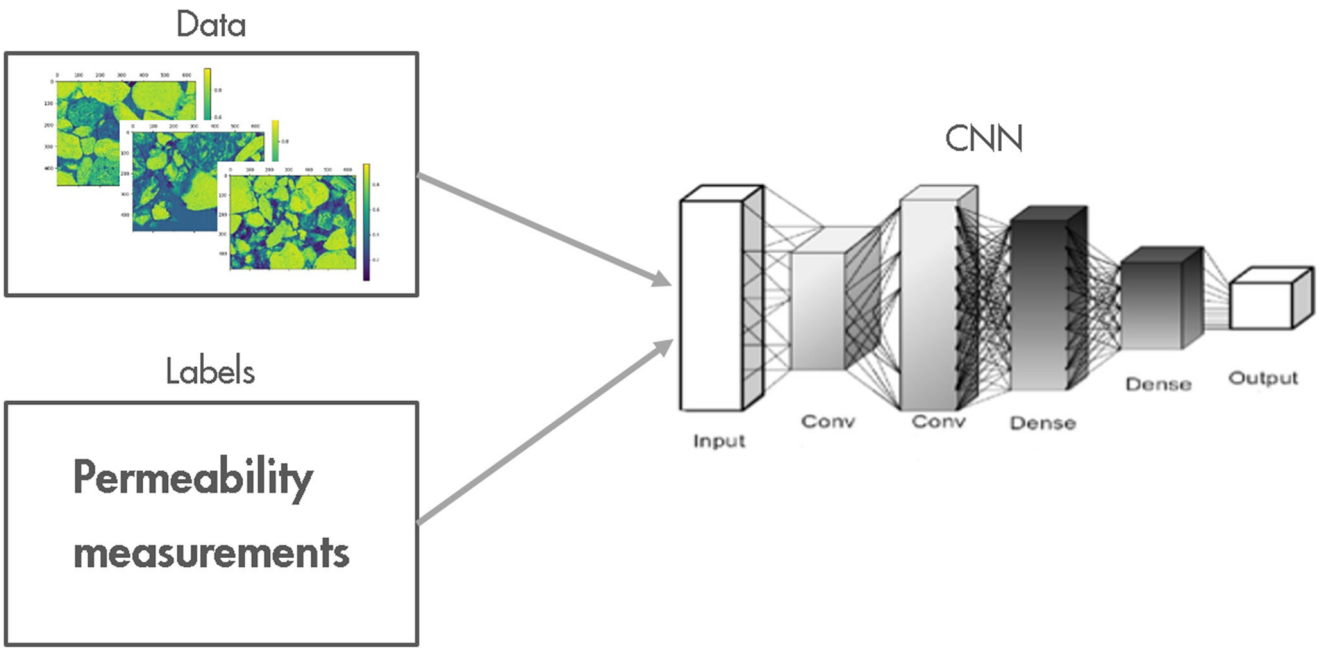

Fig. 7 General depiction of the training workflow

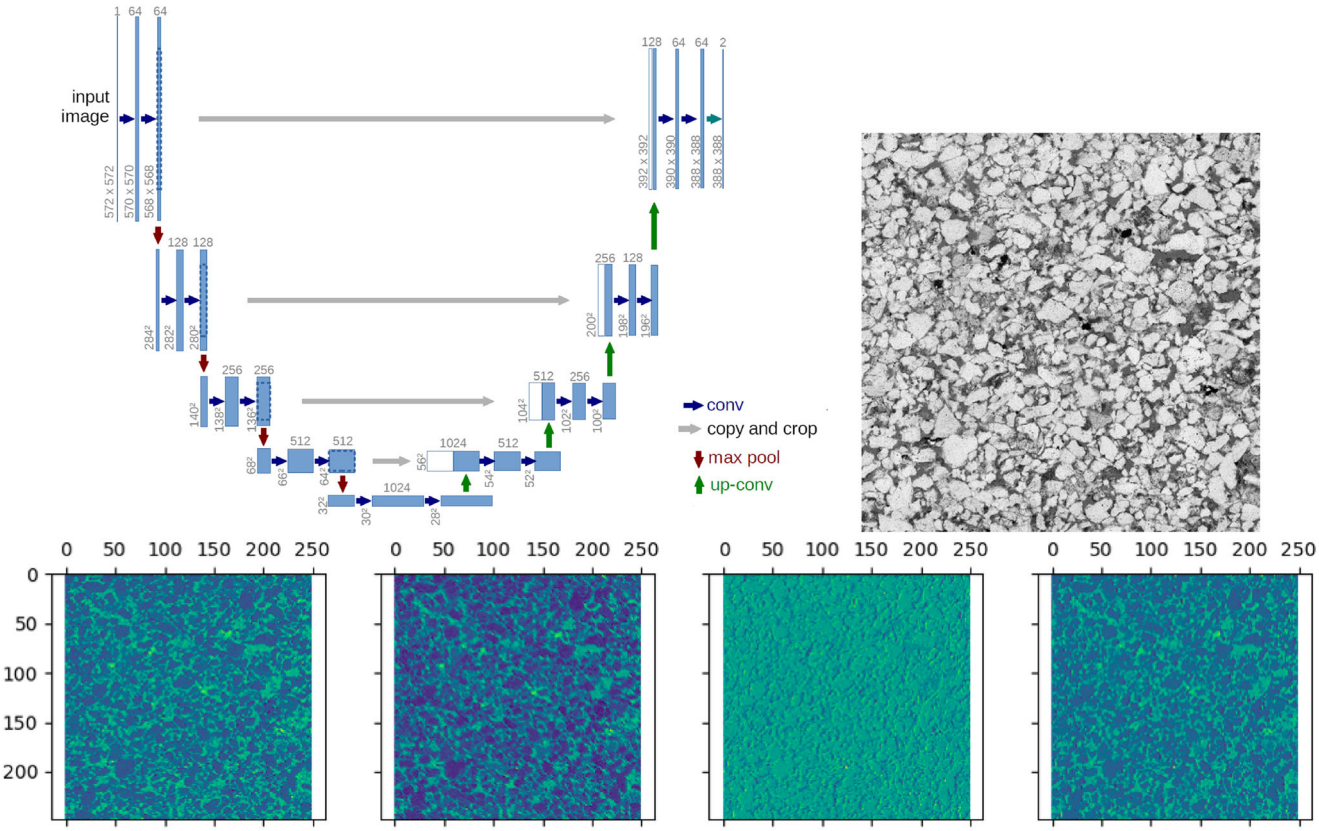

Fig. 8 (Top-left) Depiction of the U-Net architecture, pictured by [13]. (Top-right) Example of input image, sub sampled and cropped from the original image, one channel. (Bottom) Features maps (four filters) after applying one convolutional layer including activation function and max pooling

Table 1 Examples of predictions for unseen samples from reservoir 7 core plug 4. The average error is 19.09%, but if only predictions for permeability above 10 mD are considered, the average error drops to 7.73%

Sample	Ground truth (mD)	Prediction (mD)	Error (%)
1	27.58	26.33	4.53
2	74.86	71.96	3.87
3	7.88	10.33	31.09
4	63.04	59.57	5.50
5	3.94	6.46	63.96
6	15.74	18.72	18.93
7	47.28	50.02	5.80

- In terms of computing implementation, python plus Keras library ([14]) with Google's TensorFlow-backend ([15]) was used, also, NVIDIA cuDNN library ([16]) is required given that most of heavy computing was carried out with NVIDIA V100 GPGPUs.
- Data normalization is applied in order to ensure that weights of all features are equal in their representation. Rectified linear unit (ReLU) activation function is used and the loss function is mean-squared error (MSE). The network optimizer is stochastic gradient descent (SGD) with Nesterov momentum ([17, 18]).

In terms of computing resources, each node of our cluster is equipped with two NVIDIA Tesla V100 GPGPU cards. The technical specification of each CPU-GPGPU node is as follows: HP Proliant XL250a Gen9, 24 Cores-Intel(R) Haswell–Xeon(R) CPU E5-2680 v3 @ 2.50 GHz, 256 GB DRAM. It is important to note that each Tesla V100 has 5120 CUDA cores and 16 GB of total HBM2 RAM at 900 GB/s of internal bandwidth. The memory capacity is a key due to the size of the images to be used as input to the DL architectures; in fact, it is the single limiting factor when considering cropping from the original high-resolution images, and it sets the bounds to the image size and sub-sampling of the mentioned images.

3.2 Performance metrics

The results will compare using the following metrics. The performance metrics are R^2 score (coefficient of determination) and mean-squared error (MSE) for the training and testing set overall. R^2 score measures the total variation of the outcomes provided by the model, it is interpreted as the goodness of the model fitting, the values can be negative, and the optimal value is 1. In terms of error, it is expected that MSE is close to 0 as possible. For individual predictions, we will use simple relative error, since ground truth is known.

3.3 Case 1: Results for sub-samples from single thin section

In this sub-section, we introduce the results of using DL to directly predict permeability on *one* image from the

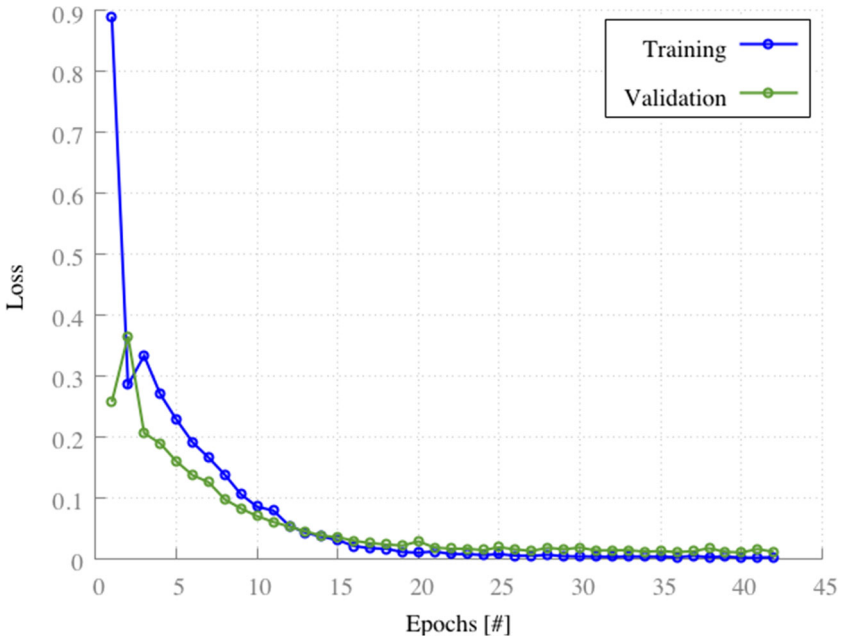

Fig. 9 The training and validation process is stopped after 42 epochs, both curves monotonically descend, this for the validation curve implies that the network is not over-fitting

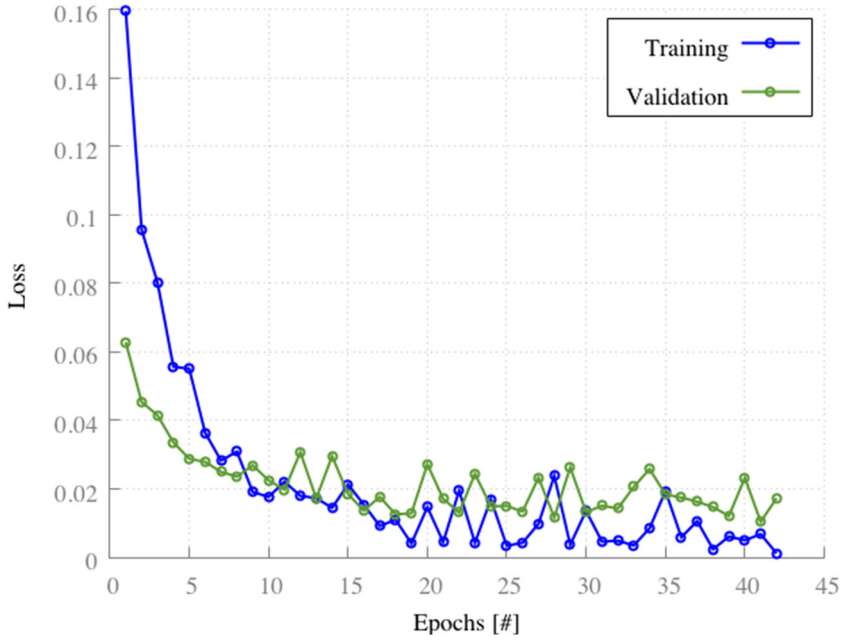

Fig. 10 The training and validation process is stopped after 42 epochs. Training curve monotonically decreased, but validation curve stop descending which might indicate some level of over-fitting

reservoir 7 core plug 4. The training and testing datasets are generated from a hi-res image of size 33020×9434; then, this image is sub-sampled and cropped to patches that can fit in GPGPU main memory. The working data set is composed of a couple of hundred lower resolution 640×480 images. The dataset is split into training and testing (unseen during training) sub-sets, 90% and 10% of the dataset, respectively. The permeability ranges from 3.94 to 74.86 mD for this dataset. Both extremes values are hold out from the training process and predicted as can be seen in Table 1 rows 2 and 5. Hyperparameter optimization ([19]) is carried out in order to set the best possible parameters to train the DL architecture (described in the implementation section). The converge of the network with the best parameters obtained by hyperparameter optimization depicted in Fig. 9.

The overall performance metrics (as described in the preceding section) are as follows:
Testing: R^2 0.9582 (1.0 max), MSE 0.0108 (0.0 max)
The performance results are encouraging in particular in terms of prediction confidence (R^2), but there is room for improvement error-wise.

In terms of specific examples of permeability predictions, the following table shows predictions for some sub-sampled unseen images:

Generalization is a concern when all data is narrow sampled, which is the case when all data is sub-sampled from the same thin section, particularly if that sample lithology is highly homogeneous, similar to what was presented in [2]. Also, it is known that small permeability values are difficult to measure, the error range can be as large as the measurements themselves. As can be seen in Table 1, in particular for samples 3 and 5, similar situation affects the DL-based predictions, but for DL the main issue is the distribution of the input data and lack of sufficient labels for small permeability values. Although the predictions accuracy is encouraging, we needed to test the concept with a larger collection of samples, where the distribution of the input data label pairs cover a large and

Table 2 Predictions for samples from the training set. The average error is 5.56%. The error is consistent across the examples, which is reflected in the R^2 metric

Sample	Ground truth (mD)	Prediction (mD)	Error (%)
Reservoir 1, core 91	479.00	448.34	6.43
Reservoir 2, core 85	929.00	956.45	2.95
Reservoir 3, core 62	160.00	185.05	15.63
Reservoir 4, core 1	238.00	238.63	0.26
Reservoir 2, 67	833.00	854.06	2.53

Table 3 Predictions for samples from the unseen set (testing). The average error is 11.69%. As expected the error here is larger than for the training dataset, still the predictions are useful and well inside the measurement uncertainties

Sample	Ground truth (mD)	Prediction (mD)	Error
Reservoir 5, core 2	271.00	239.25	12.81
Reservoir 6, core 1	456.00	443.19	2.81
Reservoir 2, core 79	707.78	708.32	0.08
Reservoir 2, core 40	667.00	574.78	13.83
Reservoir 1, core 71	339.00	437.16	28.96

even range of values, assuming that this compensates for measurement uncertainty, and at the same time, mitigates the generalization concern.

3.4 Case 2: Multiple reservoirs thin section results

To address previous concerns, the permeability (labels) in this section ranges from 50 to 1100 mD, and a careful selection of the samples helps to avoid label imbalance problem. The new data set is composed of samples from 11 different reservoirs and multiple core plugs from them. The sample distribution is as follows: reservoir 2 (70 samples), reservoir 7 (1 sample), reservoir 1 (13 samples), reservoir 3 (12 samples), reservoir 4 (8 samples), reservoir 6 (11 samples), reservoir 11 (2 samples), reservoir 5 (9 samples), reservoir 8 (6 samples), reservoir 9 (8 samples), and reservoir 10 (3 samples). The treatment of these samples is similar to the previous section, but this time the size of the input images is larger (1000×1000), therefore less cropped images where used, this given that images from all the different reservoirs have an average size of 40000×40000, which overwhelms the memory capacity of current GPGPUs. In terms of data splitting, we again used 90% of data for training (validation) and only 10% for testing.

The overall performance metrics (as described in the preceding section) are as follows:
Training: R^2 0.9958 (1.0 max), MSE 0.0003 (0.0 max)
Testing: R^2 0.7967 (1.0 max), MSE 0.0106 (0.0 max) The training performance results are excellent both in prediction confidence and error, but for testing there is room for improvement in both categories (Fig. 10).

The following tables show predictions for training (Table 2) and unseen images (Table 3):

Results cover a wide range of permeability with promising performance. Results can be improved, in particular adding more training data, this giving the fact that the DL implementation added regularization, dropout, and other techniques in order to control over-fitting.

In terms of computing, the training process for this data set takes only minutes in a couple of NVIDIA V100 GPUs, and the predictions are computed in a matter of seconds. This is the great advantage of this method with respect to laboratory measurements and simulation, once a DL model is trained, prediction time is basically negligible. The main requirement for this method is access to accurate labeled data, the source of that can be legacy laboratory measurements and/or highly accurate simulations. Any bias that the data carries will be passed by training to the DL model.

4 Conclusions

The current means of computing rock properties, such as permeability, are time consuming and computationally costly. Either laboratory measurements are required on core samples or micro-CT imaging of these samples and computer simulations with computational fluid dynamics techniques on these images are needed. An alternative method is developed based on the deep learning (DL) approach, which takes advantage of an existing library of thin-section images. We show that the resulting workflow exhibits a great promise. A set of legacy proprietary data with labels is the starting point for the DL-driven direct prediction of permeability. We show the feasibility of this method using a real dataset and a state-of-the-art machine learning algorithm taking advantage of GPUs. The resulting deep neural network is trained and tested with a dataset composed of samples from 11 different reservoirs and 135 different cores. The average error of the predictions is only 11.69% for permeabilities that range from 50 to 1100 mD. The predictions are produced almost instantaneously which speeds-up all related field-development workflows that depend on permeability estimation. It is important to mention that the DL architecture can potentially predict any property given the presence of a proper label database, such as compressibility, etc.

The DL-based property estimation technology is being tested with a larger and more demanding dataset and incorporated into the production workflow for digital rock. Also, new workflows that intertwine simulation and DL-based approaches are in development targeting even more ambitious objectives combining physics modeling and machine learning.

Acknowledgments We thank Shell International Exploration and Production Inc. for allowing the publication of this material.

References

1. Saxena, N., Hows, A., Hofmann, R., Alpak, F.O., Freeman, J., Hunter, S., Appel, M.: Imaging and computational considerations for image computed permeability: operating envelope of digital rock physics. Adv. Water Resour. **116**, 127–144 (2018)
2. Sudakov, O., Burnaev, E., Koroteev, D.: Driving Digital Rock Towards Machine Learning: Predicting Permeability with Gradient Boosting and Deep Neural Networks. ArXiv e-prints (2018)
3. Sochi, T.: Pore-Scale Modeling of Non-Newtonian Flow in Porous Media. PhD thesis, 2010 (2010)
4. Srisutthiyakorn, N.: Deep-learning methods for predicting permeability from 2D/3D binary-segmented images. pp. 3042–3046 (2016)
5. Saxena, N., Mavko, G., Hofmann, R., Srisutthiyakorn, N.: Estimating permeability from thin sections without reconstruction: digital rock study of 3d properties from 2d images. Computers and Geosciences **102**, 79–99 (2017)
6. LeCun, Y., Bengio, Y., Hinton, G.: Deep learning. Nature **436**(05), 521 (2015)
7. Goodfellow, I., Bengio, Y., Courville, A.: Deep Learning. MIT press (2016)
8. Adler, A., Boublil, D., Zibulevsky, M.: Block-based compressed sensing of images via deep learning. In: 2017 IEEE 19Th International Workshop on Multimedia Signal Processing (MMSP), pp. 1–6 (2017)
9. Wang, G.: A perspective on deep imaging. IEEE Access **4**, 8914–8924 (2016)
10. Würfl, T., Hoffmann, M., Christlein, V., Breininger, K., Huang, Y., Unberath, M., Maier, A.K.: Deep learning computed tomography: learning projection-domain weights from image domain in limited angle problems. IEEE Trans. Med. Imaging, pp. 1–1 (2018)
11. Araya-Polo, M., Jennings, J., Adler, A., Dahlke, T.: Deep-learning tomography. Lead. Edge **37**(1), 58–66 (2018)
12. Hornik, K., Stinchcombe, M., White, H.: Multilayer feedforward networks are universal approximators. Neural Netw. **2**(5), 359–366 (1989)
13. Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation. In: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (eds.) Medical Image Computing and Computer-Assisted Intervention –MICCAI 2015, pp. 234–241. Springer International Publishing, Berlin (2015)
14. Chollet, F., et al.: Keras. https://keras.io (2015)
15. Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., Ghemawat, S., Irving, G., Isard, M., Kudlur, M., Levenberg, J., Monga, R., Moore, S., Murray, D.G., Steiner, B., Tucker, P., Vasudevan, V., Warden, P., Wicke, M., Yu, Y., Zheng, X.: Tensorflow: a system for large-scale machine learning. In: Proceedings of the 12Th USENIX Conference on Operating Systems Design and Implementation, OSDI'16, pp. 265–283. USENIX Association, Berkeley (2016)
16. Chetlur, S., Woolley, C., Vandermersch, P., Cohen, J., Tran, J., Catanzaro, B., Shelhamer, E.: CuDNN: Efficient Primitives for Deep Learning. ArXiv e-prints (2014)
17. Ruder, S.: An Overview of Gradient Descent Optimization Algorithms. ArXiv e-prints (2016)
18. Sutskever, I., Martens, J., Dahl, G., Hinton, G.: On the importance of initialization and momentum in deep learning. In: Proceedings of the 30th International Conference on International Conference on Machine Learning - Volume 28, ICML'13, pp. III–1139–III–1147. JMLR.org (2013)
19. Bergstra, J., Bardenet, R., Bengio, Y., Kégl, B.: Algorithms for hyper-parameter optimization. In: Proceedings of the 24Th International Conference on Neural Information Processing Systems, NIPS'11, pp. 2546–2554. Curran Associates Inc, USA (2011)