XLVI CILAMCE
VITÓRIA - ES - BRAZIL - 2025
XLVI IBERO-LATIN AMERICAN CONGRESS ON COMPUTATIONAL METHODS IN ENGENEERING

Applications of self-supervised deep learning in micro-tomography of pre-salt carbonate rocks

Carlos E. M. dos Anjos[1], Alexandre G. Evsukoff[1], Rodrigo Surmas[2], Kamaljit Singh[3]

[1]*COPPE, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil*
cmenezes@lamce.coppe.ufrj.br
[2]*CENPES, Petróleo Brasileiro S.A, Petrobras, Rio de Janeiro, Brazil*
[2]*School of Energy, Geoscience, Infrastructure and Society, Heriot Watt University, Edinburgh, United Kingdom*

Abstract. Reservoir characterization relies on accurately capturing the complex pore-scale textures of carbonate rocks; however, annotated micro-CT data remain scarce and costly. Conventional digital rock workflows rely on CT and micro-CT imaging, plus labor-intensive image analysis, to derive petrophysical properties such as porosity and permeability, tasks complicated by the heterogeneity of pre-salt carbonates. Deep learning, leveraging multi-layered neural networks, excels in extracting patterns from data without extensive preprocessing, enabling automation of rock image analysis and property estimation. A major challenge lies in maximizing unannotated data usage while learning from limited annotated samples. Self-supervised frameworks address this by pre-training models on pretext tasks before fine-tuning for downstream applications like permeability estimation. This approach reduces reliance on annotated data, leveraging abundant unannotated samples effectively. This work proposes the use of self-supervised deep learning models for different predictive tasks: image recovery and absolute permeability estimation. The permeability values for each sample were obtained through laboratory experiments. The image recovery task was chosen since rock micro-CT images contain observable visual patterns, making it a feasible task. The results aim to streamline digital rock physics, improving the speed and reliability of reservoir rock characterization from complex carbonate samples.

Keywords: Deep Learning, Digital Rock, Micro-CT, Self-Supervised Learning, Artificial Intelligence

1 Introduction

In recent years, the rise of high-performance clusters capable of storing and processing massive datasets has been notable across many sectors. In the oil and gas industry, data are collected at nearly every stage of exploration and production—seismic surveys during early prospecting, down-hole sensors while drilling, micro-CT scans of cores and sidewall samples, among others. These diverse datasets inform critical operational decisions, yet their volume presents a considerable challenge to efficient processing and analysis. As a result, artificial-intelligence techniques have become an increasingly active research frontier, with major operators (e.g., Petrobras, Shell) investing heavily in AI-enabled tools [1].

Reservoir characterization, particularly for heterogeneous carbonate systems, exemplifies these challenges. The workflow spans geological, physical, chemical, and mathematical analyses to clarify rock composition, layering, and petrophysical behavior [2]. Of the many properties considered, porosity, permeability, and lithology are decisive for assessing reservoir quality [3]. Accurate estimation depends on high-resolution CT and micro-CT imaging, supplemented by laboratory tests and numerical flow simulations, as well as accurately determining of the texture characteristics that compose it [4, 5].

Digital Rock Physics (DRP) has advanced our ability to simulate pore-scale flow, yet still relies on precise segmentation and the inclusion of key geological parameters [6]. Manual annotation is labor-intensive and prone to subjectivity, while automated segmentation struggles with complex pore systems [7]. Deep-learning models alleviate some of this burden by learning hierarchical features directly from images [8, 9]. In the petroleum industry, these techniques have already been applied in several studies [3, 10, 11]. Nevertheless, traditional supervised training demands large labelled sets—resources rarely available in petroleum datasets.

Self-supervised learning offers a promising alternative. Frameworks such as SimCLR [12] and DINO [13] exploit vast quantities of unlabelled images, pre-training on proxy tasks before fine-tuning with limited labels [14]. Leveraging this paradigm, the present work applies self-supervised and supervised models to carbon-

ate micro-CT data for two objectives: (i) image retrieval to group samples exhibiting similar textures and (ii) absolute-permeability prediction.

2 Materials and Methods

The standard procedure employed for rock characterization, specifically for lithological classification and permeability/porosity measurement can be seen in Figure 1; they are time-consuming. The lithological classification is a biased process, since it depends on the interpreter's subjective judgment. In Figure 1, the proposed artificial intelligence workflow is presented, which can speed up the process and provide specialists with an early prediction to inform decisions earlier, while waiting for the experimental results to be ready.

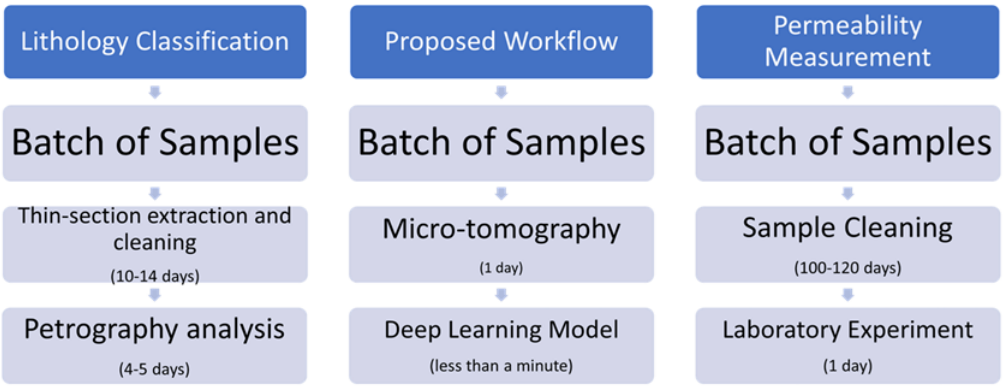


Figure 1. Characterization workflows and proposed artificial intelligence workflow.

Micro-CT is a technique incorporated into the reservoir characterization workflow, producing images that closely correspond to a series of thin sections, enabling a better analysis of the volume as a whole when compared to just one thin section. More details on micro-CT reconstruction and image acquisition of earth materials can be found in Hanna and Ketcham [15] and Cnudde and Boone [16]. This work proposes a deep-learning approach for automatic rock characterization using micro-CT images. The authors proposed a similar workflow in the past where deep convolutional models were used for lithology classification and permeability estimation in a supervised manner [10, 11].The novelty here is the use of self-supervised models, transfer learning, larger models, and the available dataset.

2.1 Dataset

The dataset available for training and testing the models consisted of 5,201 micro-CT images of rock samples with only about 3,000 samples having permeability values available. Each image was cropped to select the region of interest, as shown in Figure 2, thereby removing the void area around the sample. Twenty equally spaced slices were taken from each plug sample.

The pre-processing steps applied to each sample consisted of resizing each image to 256×256 pixels and scaling the image according to the slice's minimum and maximum values. Since the models applied here are pre-trained, we also need to rescale to the model's expected scale (using predefined average and standard deviation values). The image slices used did not undergo any filtering or processing to remove high-density artifacts.

2.2 Model Topologies

For supervised permeability prediction we evaluated three off-the-shelf backbones: DenseNet-161 [17], RegNet [18], and Vision Transformer (ViT) [19]. DenseNet-161 is a 161-layer CNN whose dense connections promote gradient flow and feature reuse, achieving strong accuracy. RegNet families are generated by a simple four-parameter rule that yields width- and depth-balanced CNNs offering favorable accuracy-per-FLOP across a wide size range. ViT dispenses with convolutions altogether, treating an image as a sequence of patch tokens processed by a transformer encoder, thereby obtaining global receptive fields from the outset. All three architectures

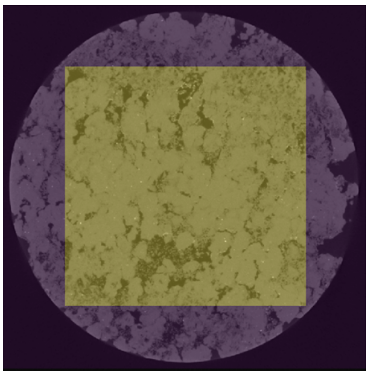


Figure 2. Cropped micro-CT slice, region of interest in yellow.

are readily available in PyTorch with ImageNet-pre-trained weights, which we use as the initialization for every experiment.

2.3 Training methods

This work compares two training paradigms for carbonate-rock micro-CT analysis. In the supervised setting, models learn permeability directly from labels via cross-entropy minimization and SGD [8]. In the self-supervised setting, labels are withheld during pre-training, allowing the network to first build generic visual representations. We adopt SimCLR, which pulls together two augmented views of the same image while pushing apart other images through an NT-Xent contrastive loss [12], and DINO, which uses a teacher-student self-distillation game with EMA weight updates to learn view-invariant features without collapse [13]. Both methods have matched or nearly surpassed supervised ImageNet baselines while requiring significantly fewer annotated samples.

After pre-training, the encoders are fine-tuned for absolute permeability regression. For image retrieval, we plug the pre-trained encoders into the SuperGlobal framework [20], which relies solely on compact global descriptors and a fast re-ranking step, thus avoiding memory and latency overheads of local-feature matching. Collectively, these pipelines enable faster, less biased rock characterization while conserving precious labelled data.

2.4 Validation statistics and cross-validation

There are many ways to evaluate machine learning algorithms [21]. Selecting the appropriate evaluation functions depends on the task executed by the model and the characteristics of the problem. In this work, RMSE and R^2 were used as evaluation metrics for the permeability estimation task, and mAP was used for the image retrieval task.

Cross-validation is a common method for prediction error estimation and is widely used in machine learning problems [21]. It consists of splitting the dataset into k subsets of equal or nearly equal sizes, where each subset is known as a fold and retains the same label distribution as the dataset. In our supervised case, the dataset was divided into five folds, with three folds used for training, one for validation, and one for testing. For the self-supervised models, all images were used during pretraining, as they do not utilize the downstream task labels. The self-supervised frameworks create their own labels based on the input, so this prevents annotation data leakage.

3 Results

Four different experiments were conducted in this work to estimate permeability, varying the topology and training paradigm. The results on the test set for each experiment are shown in Table 1. Notably, all models' results fall within each other's standard deviation when evaluating RMSE and R^2. All models achieved an R^2 score of over 0.65, and the standard deviation was lower than 0.08. Since the supervised RegNet achieved the lowest RMSE and the highest R^2 score, the predictions on the test set for this model are shown in Figure 3. Most predictions fall within an error margin of one order of magnitude, and most errors occur on low permeability. Further studies will take place to understand better the reason this is happening.

Topology	**Training**	**RMSE**	$\mathbf{R^2}$
DenseNet	Supervised	0.741 ± 0.066	0.683 ± 0.059
RegNet	**Supervised**	$\mathbf{0.720 \pm 0.064}$	$\mathbf{0.699 \pm 0.062}$
ViT	Supervised	0.732 ± 0.086	0.690 ± 0.074
DenseNet	SimCLR	0.767 ± 0.081	0.660 ± 0.073

Table 1. Performance comparison of different model topologies and training paradigms.

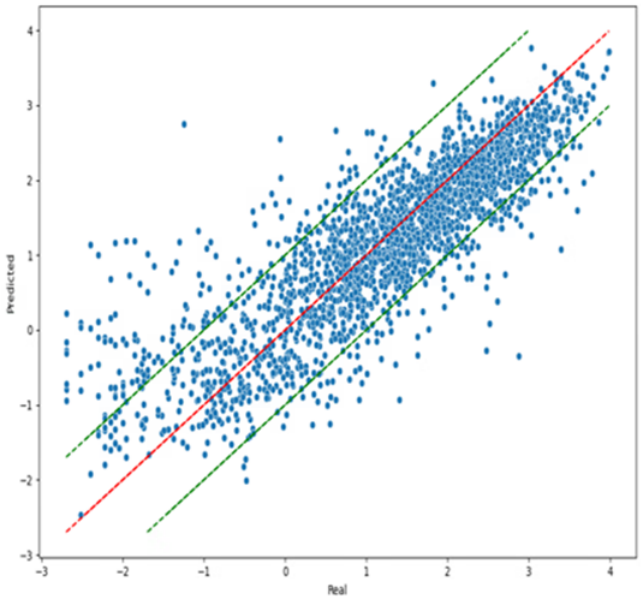


Figure 3. Permeability test set predictions of the RegNet model.

Since the SuperGlobal framework for image retrieval allows us to use almost any model for the task, four different experiments were conducted in this work varying the pretraining method. The first model is an ImageNet-pretrained DenseNet, the second a DINO Framework pretrained using micro-CT data, the third a SimCLR Framework pretrained using micro-CT data, and the last model is the supervised permeability-trained model. The results for each experiment are shown in Table 2. Notably, the models pre-trained on the micro-CT data achieved better results, indicating that training on that data improved performance on the image retrieval task. Some of the predictions can be seen in Figure 4, where the red-marked slice represents the slice used as input for image retrieval, and the rest are the predictions. The green-marked slices are from the same sample as the red-marked slice.

Model	**mAP1**	**mAP2**
ImageNet	0.59	0.54
DINO	0.74	0.65
SimCLR	**0.80**	**0.76**
Permeability	0.71	0.67

Table 2. Image-retrieval performance (mean Average Precision) of different pre-training strategies.

4 Conclusions

In this paper, we compare supervised and self-supervised approaches for permeability estimation and image retrieval on carbonate-rock micro-CT images. The findings are preliminary but satisfactory. While supervised

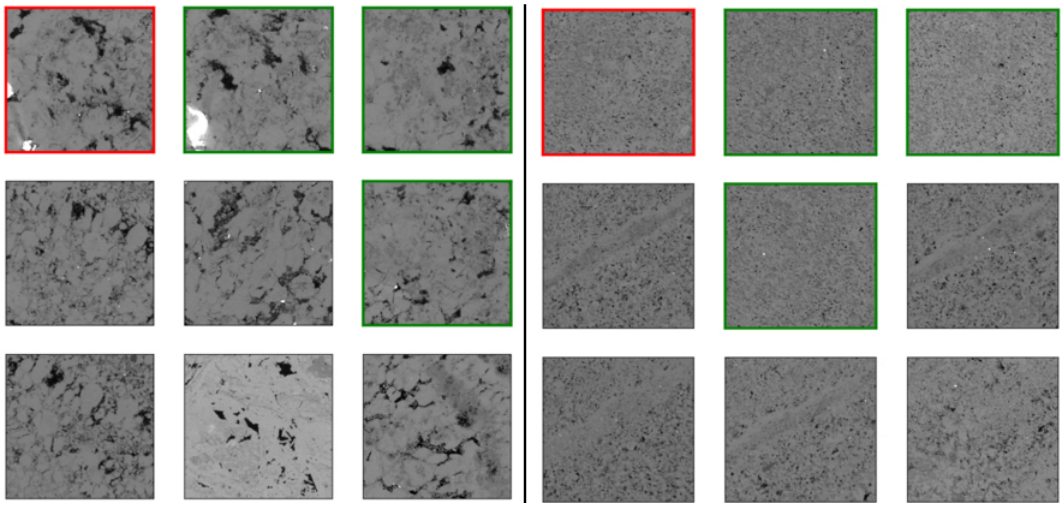


Figure 4. Image retrieval results, the reference slice is marked in red and green marked slices are from the same sample.

models achieved the best scores, the self-supervised counterparts lagged only slightly behind, indicating that label-efficient training can deliver competitive performance. Future work involves training the remaining models, conducting further analysis of the obtained results, exploring 3D approaches and additional petrophysical properties.

Acknowledgements. The authors would like to thank Petrobras for providing the data and financial support. The authors also thank the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) for the scholarships for students and researchers.

Authorship statement. The authors hereby confirm that they are the sole liable persons responsible for the authorship of this work, and that all material that has been herein included as part of the present paper is either the property (and authorship) of the authors, or has the permission of the owners to be included here.

References

[1] S. Choubey and G. P. Karmakar. Artificial intelligence techniques and their application in oil and gas industry. *Artif. Intell. Rev.*, vol. 54, n. 5, pp. 3665–3683, 2021.
[2] L. Buryakovsky, G. V. Chilingar, S. Shin, and H. H. Rieke. *Fundamentals of the Petrophysics of Oil and Gas Reservoirs*. John Wiley & Sons, 2012.
[3] M. B. Valentín, C. R. Bom, J. M. Coelho, M. D. Correia, M. P. de Albuquerque, M. P. de Albuquerque, and E. L. Faria. A deep residual convolutional neural network for automatic lithological facies identification in brazilian pre-salt oilfield wellbore image logs. *Journal of Petroleum Science and Engineering*, vol. 179, pp. 474–503, 2019.
[4] C. H. Arns, M. A. Knackstedt, W. V. Pinczewski, and E. J. Garboczi. Computation of linear elastic properties from microtomographic images: Methodology and agreement between theory and experiment. *Geophysics*, vol. 67, n. 5, pp. 1396–1405, 2002.
[5] E. Flügel. *Microfacies of Carbonate Rocks: Analysis, Interpretation and Application*. Springer, 2004.
[6] *Multi-Scale Rock Analysis for Improved Characterization of Complex Carbonates*, SPE Reservoir Characterisation and Simulation Conference and Exhibition, 2015.
[7] S. Hunter, R. Hofmann, and I. Espejo. Integrating grain-scale geology in digital rock physics. *The Leading Edge*, vol. 37, n. 6, pp. 428–434, 2018.
[8] I. Goodfellow, Y. Bengio, and A. Courville. *Deep Learning*. MIT Press, 2016.
[9] T. Zhou, S. Ruan, and S. Canu. A review: Deep learning for medical image segmentation using multi-modality fusion. *Array*, vol. 3–4, pp. 100004, 2019.
[10] C. E. M. dos Anjos, M. R. V. Avila, A. G. P. Vasconcelos, A. M. P. Neta, L. C. Medeiros, A. G. Evsukoff, R. Surmas, and L. Landau. Deep learning for lithological classification of carbonate rock micro-ct images. *Computational Geosciences*, vol. 25, n. 3, pp. 971–983, 2021.

[11] C. E. M. dos Anjos, T. F. de Matos, M. R. V. Avila, de J. Castro Vargas Fernandes, R. Surmas, and A. G. Evsukoff. Permeability estimation on raw micro-ct of carbonate rock samples using deep learning. *Geoenergy Science and Engineering*, vol. 222, pp. 211335, 2023.
[12] T. Chen, S. Kornblith, M. Norouzi, and G. Hinton. A simple framework for contrastive learning of visual representations. In *International conference on machine learning*, pp. 1597–1607. PMLR, 2020a.
[13] M. Caron, H. Touvron, I. Misra, H. Jégou, J. Mairal, P. Bojanowski, and A. Joulin. Emerging properties in self-supervised vision transformers. *CoRR*, vol. abs/2104.14294, 2021.
[14] T. Chen, S. Kornblith, K. Swersky, M. Norouzi, and G. E. Hinton. Big self-supervised models are strong semi-supervised learners. *CoRR*, vol. abs/2006.10029, 2020b.
[15] R. D. Hanna and R. A. Ketcham. X-ray computed tomography of planetary materials: A primer and review of recent studies. *Chemie der Erde-Geochemistry*, vol. 77, n. 4, pp. 547–572, 2017.
[16] V. Cnudde and M. N. Boone. High-resolution x-ray computed tomography in geosciences: A review of the current technology and applications. *Earth-Science Reviews*, vol. 123, pp. 1–17, 2013.
[17] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger. Densely connected convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 4700–4708, 2017.
[18] I. Radosavovic, R. P. Kosaraju, R. B. Girshick, K. He, and P. Dollár. Designing network design spaces. *CoRR*, vol. abs/2003.13678, 2020.
[19] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, J. Uszkoreit, and N. Houlsby. An image is worth 16x16 words: Transformers for image recognition at scale. *CoRR*, vol. abs/2010.11929, 2020.
[20] S. Shao, K. Chen, A. Karpur, Q. Cui, A. Araujo, and B. Cao. Global features are all you need for image retrieval and reranking, 2023.
[21] J. Friedman, T. Hastie, and R. Tibshirani. *The elements of statistical learning*. Springer series in statistics New York, 2001.