www.nature.com/scientificdata

scientific data

OPEN

DATA DESCRIPTOR

Multi-resolution X-ray micro-computed tomography images of carbonate rocks from brazilian pre-salt

Alyne Duarte Vidal[1] ✉, Aurea Pereira Martins Neta[1], Júlio de Castro Vargas Fernandes[1], Lizianne Carvalho Medeiros[1], Carlos Eduardo Menezes dos Anjos[1], Rodrigo Surmas[2] & Alexandre Gonçalves Evsukoff[1]

The recent surge in artificial intelligence (AI) advancements has been driven by the availability of open datasets for model development and evaluation. However, in the field of earth sciences, particularly in digital rock physics applications, open data remains scarce. To bridge this gap, we introduce a dataset comprising 16 rock samples from the Brazilian pre-salt region, available in both low resolution (48 μm - 64 μm) and high resolution (6 μm - 8 μm). The dataset also includes their respective segmented images into pore and matrix. Furthermore, porosity and permeability values obtained from laboratory measurements are provided for all samples. This dataset serves as a valuable resource for developing and benchmarking AI-based superresolution/segmentation models. Additionally, it can be utilized to develop models for predicting porosity and permeability directly from μ-CT images.

Background & Summary

Characterizing reservoirs is a significant challenge and critical step in various applications, including carbon capture and storage as well as oil extraction in the oil and gas industry. This process typically involves a multidisciplinary approach, combining physical, geological, and mathematical modeling techniques to obtain accurate representations of rock layer structures, compositions, and physical properties[1]. However, when dealing with carbonate rocks from the Brazilian pre-salt region, the complexity of this task increases due to their inherently heterogeneous nature. Therefore, acquiring comprehensive knowledge of their petrophysical and structural properties is essential for optimizing operational efficiency and maximizing resource utilization rates in these applications[2].

Digital rock physics is increasingly being utilized for the precise analysis and characterization of physical properties such as porosity and permeability[3–5]. One key technology in this domain is X-ray computed tomography (CT), which has been adapted for geological investigations, particularly in petroleum geology[6–8]. This non-invasive technique excels at generating high-resolution, three-dimensional images of a sample's internal structure without causing any damage to the sample itself. Notably, this digital approach significantly accelerates the characterization process compared to traditional laboratory-based methods, which require time-consuming rock cleaning procedures that typically span 100-120 days before analysis can commence[9,10]. However, while numerical simulations enable rapid estimation of petrophysical properties, they often rely on high-resolution images, which are costly and time-consuming to acquire through microtomography (μ-CT). To address this limitation, this descriptor is designed to provide both high and low-resolution images, accompanied by their corresponding segmentations. In this context, the release of a comprehensive dataset comprising microtomography images of pre-salt carbonate rocks is of paramount importance to the digital rock physics community. Its value lies particularly in the fields of artificial intelligence (AI) and petrophysical property characterization, where it enables comprehensive analysis at different scales. The availability of multi-resolution images presents a significant opportunity for the development and validation of superresolution algorithms[11–13]. This rapidly

[1]COPPE - Federal University of Rio de Janeiro, mailbox 68506, Rio de Janeiro, 21941-972, Rio de Janeiro, Brazil. [2]CENPES - Petrobras, Av. Horácio Macedo, 950, Rio de Janeiro, 21941-915, Rio de Janeiro, Brazil. ✉e-mail: alyne.vidal@coppe.ufrj.br

evolving area of research in image analysis aims to reconstruct high-resolution images from low-resolution ones. Successful implementation of these algorithms can significantly enhance the quality and accuracy of analyses, leading to more precise characterizations and improved predictive models. Furthermore, this dataset facilitates the development of segmentation algorithms for distinguishing between pore and matrix structures at both high and low resolutions. The inclusion of laboratory-measured permeability values also enables the development of predictive models that can enhance our understanding of rock properties and improve the accuracy of simulations in various geological and engineering applications.

In digital rock applications, it is common to acquire images of the same rock sample-such as plugs, sidewall samples, or subsamples of a rock matrix-at multiple resolutions. These images, whether low-resolution, high-resolution, or zoomed, depict the same underlying rock matrix but vary in the level of detail and quality of information they provide. Zoomed images typically serve as an intermediate step between larger samples and more detailed subsamples. Having images at different scales offers vast opportunities for petrophysical, geological, and AI applications. In the field of petrophysics, this dataset can be used for porosity and permeability analyses, characterizations of rock connectivity and heterogeneity, and simulations of advanced properties like acoustics, wettability, and relative permeability.In geology, high-resolution images are particularly useful for investigating diagenetic processes that impact rocks over time, such as cementation and mineral dissolution. This analysis can provide valuable insights into the diagenetic history of a formation and its effects on petrophysical properties. Additionally, these images enable precise identification of specific minerals and rock components, allowing three-dimensional visualizations of textures and internal structures. AI further enhances the utility of this dataset through various applications. Image classification algorithms can automatically segment rock samples into pore and matrix files, while object detection techniques identify features like pores, grains, or fractures within samples. Semantic segmentation models partition images into distinct regions based on their properties, enabling quantitative assessments of material distributions. Feature extraction algorithms extract meaningful attributes from images, which can be used for conducting predictive modeling on petrophysical properties. Moreover, AI facilitates anomaly detection to identify unusual patterns within images, indicating potential geological features or artefacts. Transfer learning leverages pretrained models to extract relevant features from images, reducing the need for extensive labelled data. Generative modelling techniques can augment the dataset by generating synthetic images with characteristics similar to those obtained from real samples, expanding the training data for achieving improved model performance.

The dataset presented in this descriptor was compiled from 16 samples of carbonate rocks sourced from the Brazilian pre-salt region, a critical and challenging exploratory environment for the oil and gas industry. The μ-CT acquisition processes were performed using the X-ray emission μ-CT VTomex M model manufactured by Waygate Technology. The resulting images were subsequently processed, filtered, registered, and cropped into equal volumes for both high and low-resolution datasets. Furthermore, these images were segmented into pore and matrix structures. A detailed description of the workflow employed to obtain and process the μ-CT data, as well as their segmentations, is provided in the Methods section. Additionally, the Data Records section provides a comprehensive list of all images included in the dataset, along with their main characteristics, while the Usage Notes section offers guidance on how to access and utilize these datasets using Python.

Methods

In this section, we provide a detailed description of the methodology used in this study, focusing on three fundamental aspects: descriptions of the database, the image acquisition process using a μ-CT, and the workflow employed in image processing.

Data Description. The dataset presented comprises 16 carbonate rock samples from the Brazilian Pre-Salt. The selected samples fall into two distinct categories: sidewall core samples with a diameter of 2.54 cm and plug samples with a diameter of 3.81 cm. Each sample was used to generate μ-CT images at two resolutions, facilitating the development of super-resolution methods. Specifically, the sidewall core samples yielded high and low-resolution images at 6 μm and 48 μm, respectively, while the plug samples produced images at 8 μm and 64 μm, respectively. This resolution scheme was chosen to maintain a consistent eight-fold difference between low and high resolutions, thereby ensuring consistent analysis and comparison across the dataset.

The dataset comprises both grayscale (16-bit) and segmented images. The grayscale images underwent processing, including artifact reduction through filtering, registration, and cropping, ultimately yielding 32 image files. Meanwhile, the segmented images were obtained using Otsu[14,15] and Watershed[16,17] methods. Each original image has a corresponding segmented version, divided into pore and matrix components, thereby adding 32 more files to the dataset and bringing the total to 64 files.

The integration of both grayscale and segmented images significantly enhances the versatility and applicability of this dataset. The grayscale images are particularly well-suited for the development and validation of super-resolution algorithms, whereas the segmented images can be directly utilized for the evaluation of segmentation algorithms or for implementing super-resolution methods that bypass the need for grayscale data. Moreover, we also provide experimentally determined permeability and porosity values, which were obtained through laboratory analysis. These values can be leveraged to develop or, given the limited size of this dataset, evaluate predictive models for these characteristics in complex heterogeneous rocks, such as carbonates.

Image Data Acquisition. The μ-CT setup consists of an X-ray source and detector, with the rock sample positioned between them. During scanning, the sample rotates 360 degrees around a fixed axis, enabling detailed imaging from multiple angles. This rotation process captures a series of radiographic (2D) images at various angles, which represent the attenuation of X-rays passing through the sample. The grayscale values in these images reflect the varying absorption of X-rays by different materials and structures within the rock.

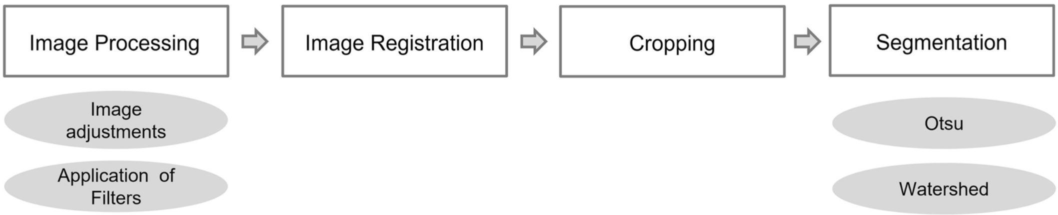

Fig. 1 Image processing workflow.

The VTomex M model was used for all μ-CT acquisitions. This equipment features two X-ray tubes (180 kV and 300 kV) and a flat screen detector, offering faster and more enhanced image quality with better resolution than competing devices. In this case, only the 300-kV tube was utilized since the rock sample needed to remain stationary for subsequent imaging purposes.

After acquiring the 2D radiographic images, reconstruction software employs algorithms like filtered back projection and iterative reconstruction to transform the 2D images into a coherent three-dimensional volume. This reconstructed 3D volume represents the spatial distribution of the X-ray attenuation coefficient within the rock sample, effectively mapping its internal structures in detail. The resulting 3D volume provides a detailed and accurate representation of the internal features of rocks, enabling further analyses and simulations of their petrophysical properties, such as porosity and permeability.

μ-CT scans were performed at both low and high resolutions for all rock samples considered in this descriptor with energy settings of 150 KeV and 140 KeV, respectively. Low-resolution scans employed a current setting of 250 μA with an average factor of 3 to improve the signal-to-noise ratio, while the skip factor was set to 1, ensuring no data points were skipped during scanning. A 0.15-mm Cu filter was used to reduce beam hardening effects and noise by absorbing lower-energy X-rays. In contrast, high-resolution settings utilized a current of 140 μA with an average factor of 6 to further improve image quality by reducing random fluctuations in the measurements, and maintained the same jump parameter as low-resolution scans. The same type of 0.15-mm Cu filter was used for both cases to maintain consistency in image quality.

Image Processing Workflow. After acquiring the images, a post-processing workflow was applied, as illustrated in Fig. 1. This comprehensive process consisted of four primary steps: image processing, image registration, cropping, and segmentation. The first step, image processing, aimed to enhance the quality of the acquired images by reducing noise and artefacts. This ensured that the images were clear and suitable for further analysis. In the second step, image registration was performed to align the low-resolution images with their high-resolution counterparts. This step is crucial because there is an inherent trade-off between resolution and the sample size that can be captured; higher resolutions provide more detail but limit the sample area. Essentially, the high-resolution image serves as a detailed "zoom-in" on a specific part of the sample initially captured in low resolution, and, thus, image registration ensures that the same physical structure is aligned in both high and low-resolution images. Following registration, the images were cropped in the third step to generate images representing the same physical structure and maintain a consistent enlargement factor, chosen as 8x, between the low and high-resolution images. The final step was segmentation, where the images were divided into regions of interest (rock and pore).

Image Processing. After reconstructing the images into a 3D volume from the 2D projections, a quality control check was performed using the reconstruction software provided with the μ-CT equipment (Fig. 2). The goal of this step was to ensure that all 2D projections, from the first to the last, were processed correctly. This was achieved by selecting the entire area of interest of each image (red rectangle in Fig. 2(a)) and verifying its acquisition. Once the 3D digital volume was processed, a verification procedure was carried out to inspect the reconstructed volume (Fig. 2(b),(c)). The first slice (Fig. 2(b)) showed the beginning of the sample, while the last slice (Fig. 2(c)) displayed the end. This step ensured that all slices of the sample were preserved and processed correctly.

Following the initial quality control procedures, the analysis of rock images proceeds with the application of filters, a fundamental step in enhancing image clarity for detailed examination. Industry-standard software, typically incorporates several filtering tools designed to refine the visualization of rock structures. These filters play a crucial role in smoothing the rock matrix, thereby facilitating the differentiation and segmentation of various gray shades that correspond to different mineral compounds within the rock. In addition to enhancing matrix visibility, filters are instrumental in achieving a clearer distinction between the rock matrix and its pores, a key factor in accurate rock characterization. In this study, two specific filters were employed: the Median Filter and the Gaussian Filter, both applied using the Pergeos software. The Median Filter, used on low-resolution images, utilizes morphological operators to assign pixel values based on the median of their surrounding neighborhood. This approach is effective in mitigating non-Gaussian noise and minimizing the impact of small-scale artifacts, which are common in lower quality captures. In Pergeos, the adjustable parameter for the Median filter is called number of interaction, set to 2 in this study. Conversely, for high-resolution images, the Gaussian filter was used. This filter enhances contrast and smooths edges, facilitating the precise segmentation of pores in the rock. In Pergeos, the available parameters for the Gaussian filter include the kernel type, which was set to separable, and the standard deviation on the (x) and (y) axes which were defined as 2.

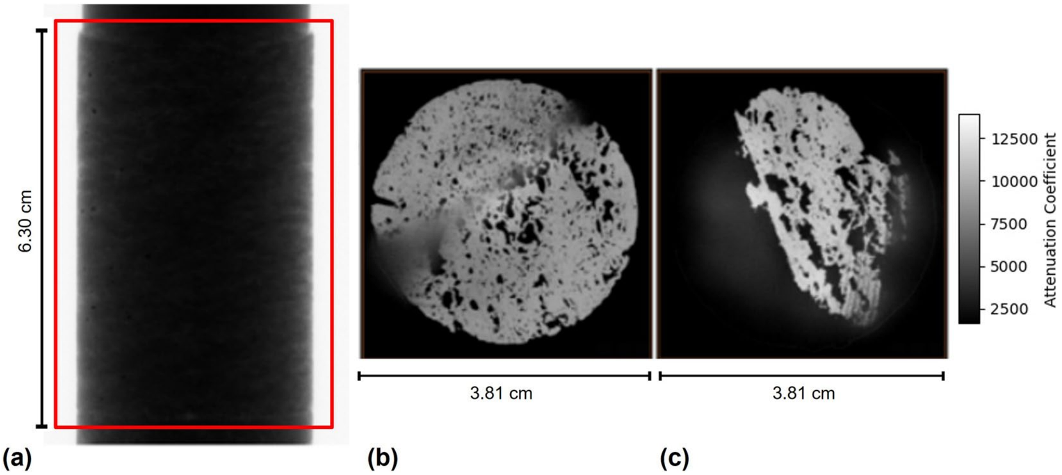

Fig. 2 Visual Quality Control. (**a**) 2D Projections Image (**b**) The first slice of the 3D volume. (**c**) The last slice of the 3D volume.

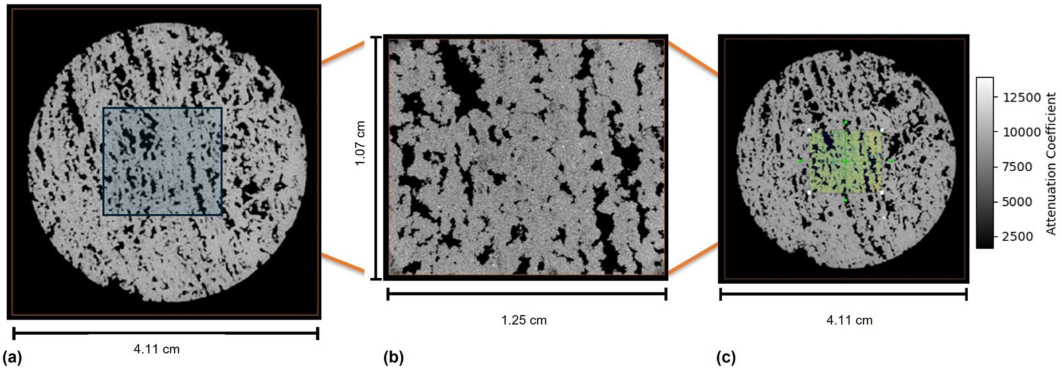

Fig. 3 Registration process for Sample P01: (**a**) Low-Resolution Image (64 μm); (**b**) High-Resolution Image (8 μm) and (**c**) Registered Image of (**a**) and (**b**).

In the subsequent step, adjustments are made to align the low and high-resolution images. The objective of this procedure is to ensure that the high-resolution image is precisely centered over the low-resolution image, enhancing the consistency and accuracy of the analysis. This alignment is critical as it sets the stage for the next phase of the process: image registration. Proper alignment ensures that the registration process is feasible.

Image Registration. In μ-CT, image registration is a crucial process that aligns images with varying resolutions into a unified voxel spatial coordinate system. This requires a transformation that precisely maps each point in one image to another, ensuring spatial coherence across different image scales. In this project, we consider zoomed images to be high-resolution images. Pre-adjustment processes facilitate the effectiveness of image registration by ensuring that the spatial positioning of the highest-resolution 3D image aligns closely with the reference image, which usually has a lower resolution. Figure 3 illustrates the alignment process, showing the overlay of a low-resolution image in Fig. 3(a) onto the corresponding zoomed image in Fig. 3(b), with successful registration displayed in Fig. 3(c). This technique confirms that all images accurately represent the same area of the sample at different scales and resolutions.

Cropping. At this stage, the high and low-resolution images are precisely aligned, ensuring accurate spatial correspondence between them. This alignment is crucial for cropping both sets of images to focus on identical regions of interest at different resolutions. The cropping process is primarily guided by the high-resolution image, as depicted in Fig. 4. Figure 4(a) shows a sample at a resolution of 64 μm. In Fig. 4(b), we see the cropping step guided by the details visible in the high-resolution image, resulting in Fig. 4(c), which is a cropped version from the same area as the high-resolution image. This method guarantees that the data extracted at varying resolutions are consistent and precise, enhancing the reliability of comparative analyses and interpretations. However, it is important to note that the resulting cubes do not necessarily have the same voxel count since their resolution differs. The image registration was verified to ensure that the high and low-resolution images were precisely aligned, corresponding voxel to voxel in the same position.

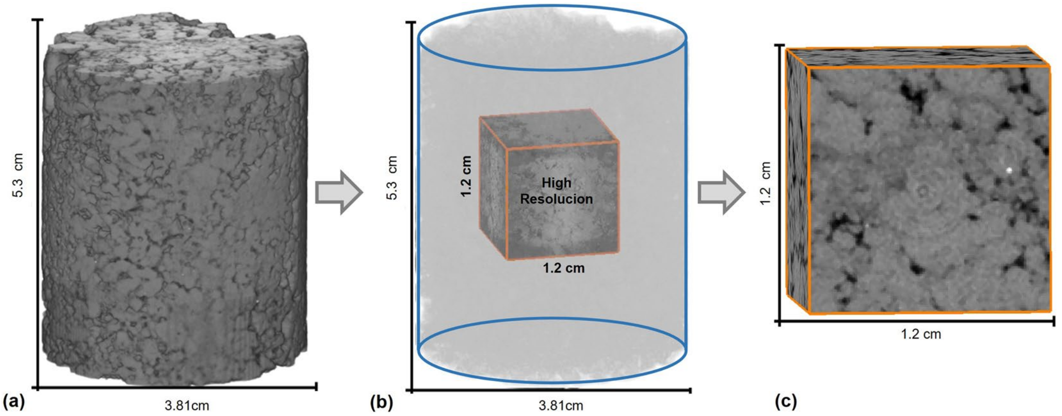

Fig. 4 Example of steps in the low-resolution cropping stage. (**a**) Low-resolution image, (**b**) cropping based on the high-resolution image, and (**c**) the final low-resolution image.

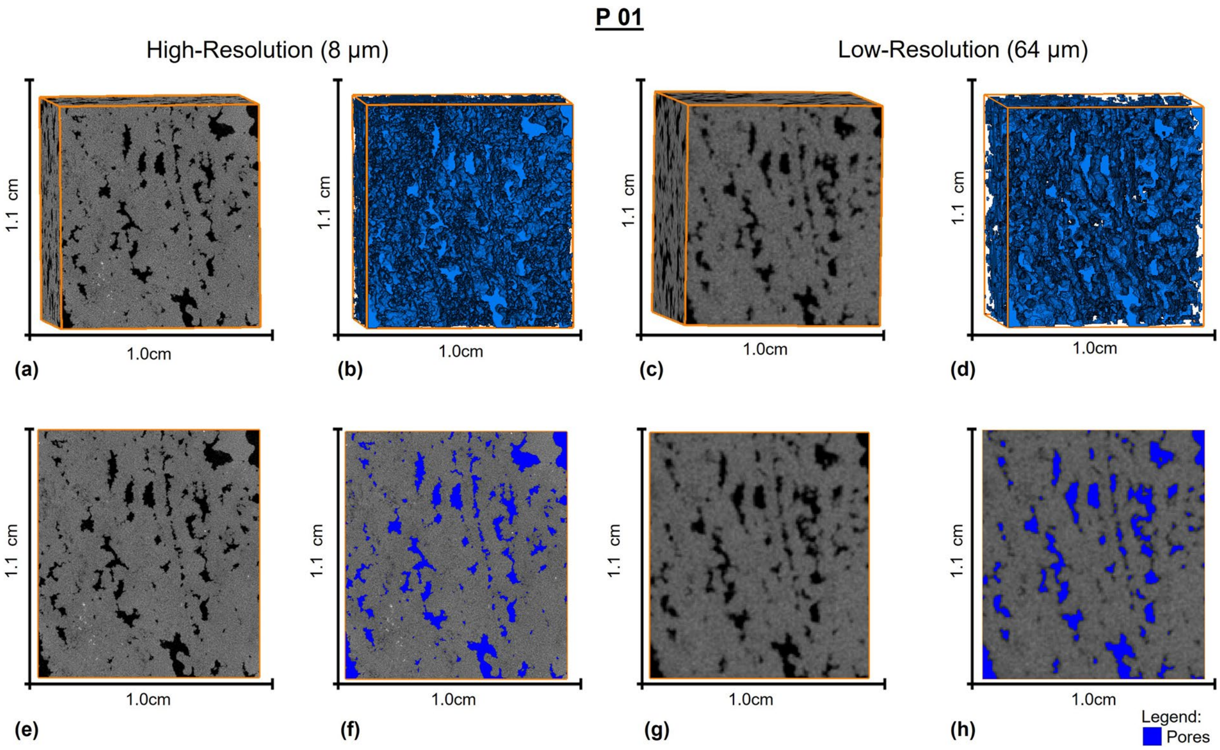

Fig. 5 High and low resolution images for the P01 sample. In (**a**) and (**c**) we have the images before segmentation for the high and low resolution respectively, in (**b**) and (**d**) we have their respective segmentations. In (**e**) and (**g**) we have 2D images before segmentation. In (**f**) and (**g**), we have 2D images showing the pores highlighted in blue.

Image Segmentation. After being cropped into cubes, both high-resolution and low-resolution images were then subjected to segmentation using the watershed algorithm. Initially, segmentation was performed using the Otsu method, which established ideal thresholds to distinguish between pores and matrix. These thresholds were carefully verified for accuracy before advancing. With the selected thresholds in place, the next step involved applying them to the watershed algorithm. The watershed segmentation method, which is a robust technique in image processing, has been employed in various research domains. The method interprets an image as a topographic landscape, where grayscale levels represent elevational heights-darker regions signify valleys, and lighter regions denote peaks. Segmentation is achieved by identifying the dividing lines between valleys, known as watershed lines, which partition the image into distinct regions or segments. This allows for the precise separation of objects and structures of interest[16,17]. This method proves particularly effective in scenarios with object overlaps or gradual variations in grayscale intensity, leading to more accurate segmentations and minimizing the risk of errors[18]. The resulting segmented images clearly delineate the pores and matrix, providing a detailed analysis of the sample's microstructure and enabling a more precise characterization of its

	Colormap		Threshold value	
Code	Low Res.	High Res.	Low Res.	High Res.
P01	5011–19022	8574–16720	7100	10300
P02	2756–9734	19094–22829	4200	20000
P03	3405–13721	20916–26352	5075	21900
P04	3760–13497	16536–20831	5500	17294
P05	3562–9667	14277–18713	5200	15250
P06	5367–16918	13391–17147	7710	14150
SW01	4071–13463	19100–27229	5646	20003
SW02	2996–10291	17055–21334	4300	17927
SW03	3282–8992	21682–25259	4030	22050
SW04	2529–10692	11837–18230	4100	13268
SW05	3083–11436	15461–21112	4300	16000
SW06	7167–12956	9971–13185	8410	10450
SW07	2469–9131	19669–22968	3851	20400
SW08	2303–11483	11204–17767	3800	12410
SW09	2899–8192	16451–19923	3850	17077
SW10	3100–9470	20391–25164	3812	21151

Table 1. Parameters used for image segmentation.

	Dimensions (pixels)		Pixel Size	
Code	Low Res.	High Res.	Low Res.	High Res.
P01	150 × 168 × 186	1200 × 1340 × 1488	64 um	8 um
P02	176 × 183 × 207	1401 × 1461 × 1655	64 um	8 um
P03	144 × 158 × 201	1149 × 1259 × 1603	64 um	8 um
P04	142 × 171 × 193	1132 × 1362 × 1541	64 um	8 um
P05	154 × 158 × 189	1226 × 1261 × 1507	64 um	8 um
P06	163 × 151 × 196	1299 × 1202 × 1562	64 um	8 um
SW01	161 × 164 × 203	1281 × 1308 × 1622	48 um	6 um
SW02	169 × 150 × 198	1348 × 1199 × 1579	48 um	6 um
SW03	139 × 164 × 197	1105 × 1310 × 1572	48 um	6 um
SW04	156 × 154 × 203	1241 × 1226 × 1620	48 um	6 um
SW05	179 × 156 × 187	1427 × 1244 × 1491	48 um	6 um
SW06	162 × 150 × 197	1294 × 1194 × 1569	48 um	6 um
SW07	151 × 164 × 200	1202 × 1307 × 1599	48 um	6 um
SW08	150 × 166 × 204	1199 × 1324 × 1625	48 um	6 um
SW09	163 × 156 × 198	1298 × 1241 × 1582	48 um	6 um
SW10	155 × 143 × 184	1236 × 1140 × 1466	48 um	6 um

Table 2. Size and resolution details for each sample μ-CT.

petrophysical properties. An example of a segmented sample is shown in Fig. 5. To ensure the reproducibility of the segmentation process of the files provided, the threshold values applied in the segmentation, as well as the colormap parameters used in the PerGeos software, are listed in the Table 1.

Data Records

The dataset acquired in this study and reported in the manuscript is available on the Digital Rocks Portal at https://www.digitalrocksportal.org/projects/503[19]. The dataset comprises 16 samples, as shown in the Table 2, each comprising two sets of grayscale images: one high resolution and one low resolution, for a total of 32 images. To further enhance the dataset, segmentation was applied to distinguish between pores and matrix, resulting in an expanded dataset of 64 files, including both original grayscale images (32) and segmented images (32). Notably, 10 of these samples are sidewall cores, while the remaining 6 are plugs. In μ-CT applications, precise differentiation between various materials and structures is crucial. To achieve this level of precision, we utilized 16-bit unsigned images due to their superior numerical resolution compared to the more common 8-bit images. The use of 16-bit images offers several advantages. They provide enhanced capability to differentiate between varying levels of X-ray absorption, which is essential for identifying and characterizing pores, minerals, fractures, and other features in rock samples. Additionally, these higher-resolution images offer significant benefits in post-processing, quantitative analysis, and image manipulation. The increased bit depth allows for

Code	Laboratory Permeability (*mD*)	Laboratory Porosity (%)	Digital Porosity high (%)	Digital Porosity low (%)
P01	1610	19.9	17.1	16.2
P02	4923	13.1	11.3	10.2
P03	1518	7.7	5.4	4.5
P04	247	11.8	10.4	8.7
P05	340	10.5	7.2	4.7
P06	1950	11.2	8.9	7.9
SW01	179	13.8	12.7	6.0
SW02	642	17.7	13.4	5.5
SW03	24.6	11.7	8.2	5.5
SW04	945	18.3	12.6	12.9
SW05	50.3	10.0	10.1	6.9
SW06	352	16.2	8.2	7.0
SW07	25.1	10.0	8.9	6.3
SW08	422	18.9	12.1	7.7
SW09	14.2	14.0	10.2	5.5
SW10	7.92	12.8	11.3	7.4

Table 3. Porosity and permeability values of the rock samples analyzed in this study.

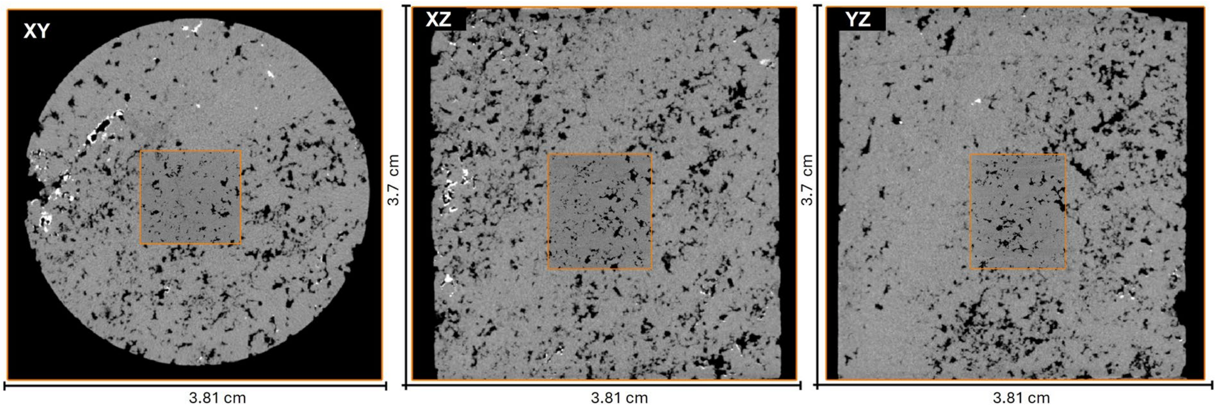

Fig. 6 The volume analyzed, highlighted by the orange rectangle at the center of the sample P05, represents a portion of the sample included in the dataset.

more substantial adjustments in contrast and brightness without compromising information, leading to more accurate and reliable data interpretation. This is particularly important when working with rock samples that often exhibit complex structures and subtle features.

Technical Validation

The technical validation process is divided into two parts: the microtomography system itself and the images it produces. To ensure the precision and accuracy of the measurements, the μ-CT VTomex M system undergoes regular calibration checks. This procedure is crucial for verifying that the images do not contain distortions and that the dimensions, density, and other parameters are measured accurately.

The Jima test, a standard evaluation method developed by the Japan Inspection Instruments Manufacturers' Association, is used to assess the spatial resolution and performance of the imaging system. In this test, the system's ability to distinguish between close details in a sample is evaluated. The Jima phantom consists of a pattern of fine lines engraved on a metal substrate with known and variable widths. This allows for the evaluation of the imaging system's ability to resolve different levels of detail.

During calibration, the Jima phantom is placed in the inspection position of the μ-CT imaging system, which then captures an image of the phantom through scanning. The resulting image is analyzed to determine the smallest line width that can be clearly resolved by the system. This result is compared with established standards to ensure that the system meets the spatial resolution requirements, thereby verifying that the image pixels accurately represent the details of the rock. Through this metrological calibration process, it was ensured that the μ-CT imaging system was capable of resolving and identifying fine details in the evaluated rock samples.

To assess the quality of our dataset, we compared the estimated digital porosities with laboratory-measured porosity values. Digital porosity was calculated by counting voxels in segmented images, which were previously

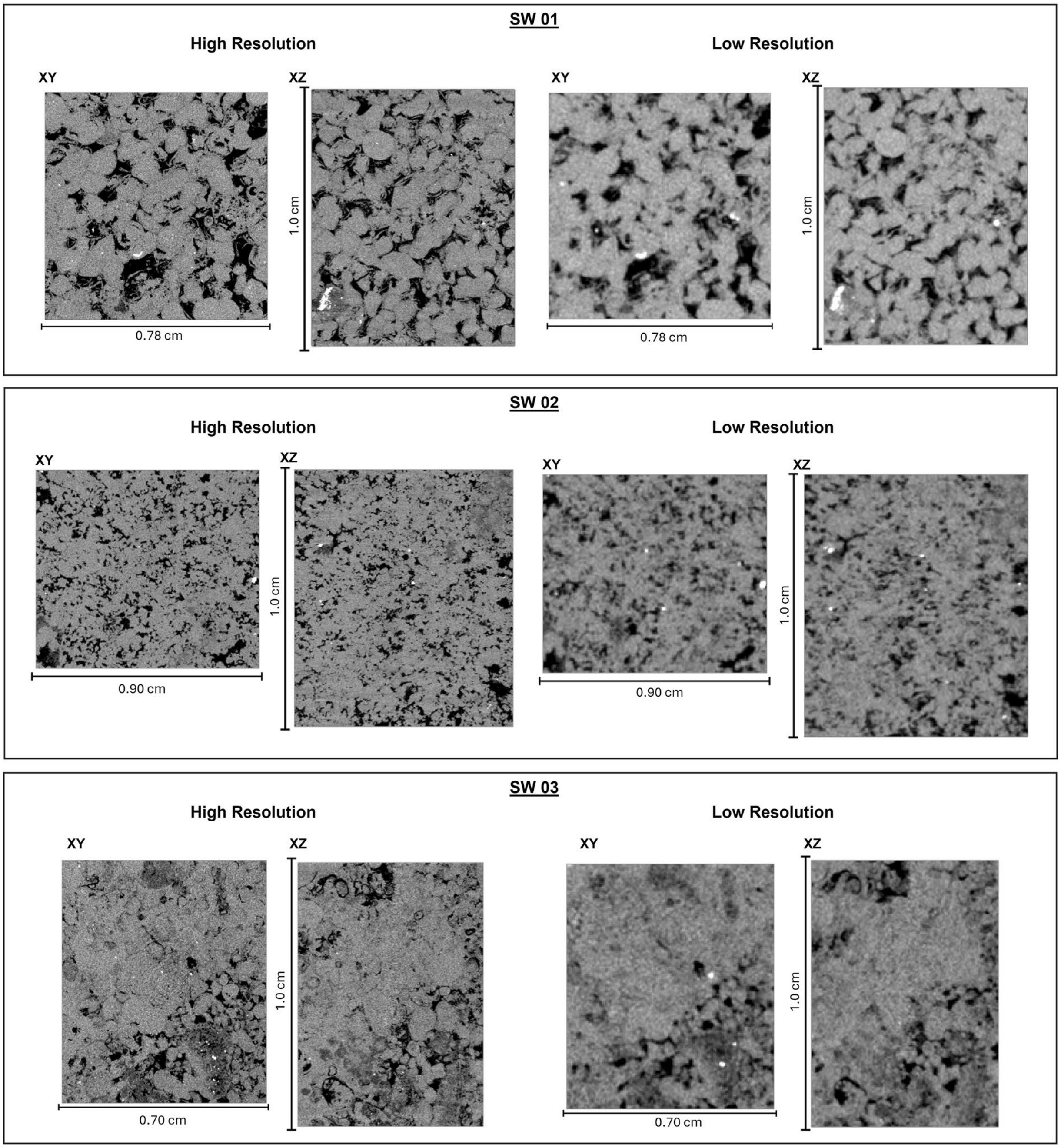

Fig. 7 High and low resolution images of samples SW01, SW02 and SW03.

divided into two phases: pore voxels (representing the pore space within the rock) and matrix voxels (comprising the solid phase of the rock). In a conventional binary image, pore voxels are assigned a value of 1, while matrix voxels are assigned a value of 0. The porosity was then calculated using the following equation:

$$\phi = \frac{N}{T} \tag{1}$$

where N represents the total number of pore voxels, and T denotes the total number of voxels in the image. This method provides a direct measure of the porosity based on the proportion of pore space observed in the segmented images.

Table 3 presents the estimated digital porosities alongside laboratory-measured porosity and permeability values. Some images within our dataset exhibit ring artifacts, which are a result of limitations inherent to the microtomography technique. Although these artifacts do not affect the segmentation process, they do impact the overall quality of the grayscale images.

As expected, the digital porosity values for all samples were lower than those obtained in the laboratory, both for low and high-resolution images. This discrepancy is attributed to limitations in image resolution and the potential unrepresentativeness of the cubes extracted from the whole image, particularly given the heterogeneity of the pre-salt carbonate samples. Figure 6 displays examples of the cubes extracted from the whole image, which were used to construct the dataset. Notably, sample P05 contains more porous regions that were not included in the central cube cut. This exclusion of porous areas results in a diminished porous volume and a lower digital porosity measurement compared to the laboratory-determined porosity reference value. This phenomenon is also evident in other samples where digital porosities are derived from high-resolution images. In contrast, the primary influencing factor for porosity results obtained from low-resolution images (48 μm-64 μm) was the limitation of image resolution itself. Given the heterogeneous nature of the samples and the diversity in pore sizes, pores smaller than the image resolution are not captured, leading to a decreased pore volume being analyzed in the low-resolution images. Figure 7 illustrates the differences between high and low-resolution images.

From these results, we can conclude that the segmentations, at least in the high resolution, are close to the real segmentation given the values obtained. However, additional work is needed to incorporate permeability simulations, this step is essential to establish a correlation between the porosity and permeability values for each sample based on μ-CT images. Including permeability will enable a more comprehensive and precise analysis of the petrophysical properties of the samples, enhancing our understanding of the behavior of the porous media. Additionally, future studies could aim to improve segmentation techniques and address the challenges posed by image resolution limitations. In summary, while the current results are promising, ongoing research is crucial to refine and extend the application of μ-CT images in assessing the petrophysical properties of heterogeneous samples.

Usage Notes

The dataset files are provided in NetCDF (.nc) format, a widely used standard for array-oriented scientific data. This format allows for easy manipulation and analysis using libraries such as xarray in Python. To work with the dataset, you can load a file named “filename.nc” using the following code snippet:

```
import xarray as xr
dataset = xr.open_dataset('filename.nc')
data = dataset['data']
```

Here, “data” refers to the specific field within the NetCDF file that contains the image data. Notably, if the filename contains the string “segmented”, it indicates that the associated images have been segmented into pores and matrices. This structured format and accompanying naming convention enable efficient data handling and analysis.

Code availability

No custom code was utilized to generate or process the data. All data processing was done using commercial softwares such as Avizo.

Received: 7 October 2024; Accepted: 28 November 2024;
Published online: 18 December 2024

References

1. Buryakovsky, L., Chilingar, G. V., Shin, S. & Rieke, H. H.*Fundamentals of the Petrophysics of Oil and Gas Reservoirs* (John Wiley & Sons, 2012). https://doi.org/10.1002/9781118472750.
2. Haro, H. A. V.*Simulação de Injeção de CO2 em Reservatórios de Petróleo para EOR e Armazenamento de Carbono*. Ph.D. thesis, PUC-Rio (2014).
3. Knackstedt, M. A. *et al*. Digital core laboratory: Properties of reservoir core derived from 3d images. In *SPE Asia Pacific Conference on Integrated Modelling for Asset Management*, 14 (Society of Petroleum Engineers, https://doi.org/10.2118/87009-MS (2004).
4. Arns, C. H. *et al*. Pore scale characterization of carbonates using x-ray microtomography. *Spe Journal* **10**, 475–484, https://doi.org/10.2118/90368-PA (2005).
5. Claes, S., Soete, J., Cnudde, V. & Swennen, R. A three-dimensional classification for mathematical pore shape description in complex carbonate reservoir rocks. *Mathematical Geosciences* **48**, 619–639, https://doi.org/10.1007/s11004-016-9636-z (2016).
6. Vinegar, H. J. X-ray ct and nmr imaging of rocks. *Journal of Petroleum Technology* **38**, 257–259, https://doi.org/10.2118/15277-PA (1986).
7. Honarpour, M. M., Cromwell, V., Hatton, D. & Satchwell, R. Reservoir rock descriptions using computed tomography (ct). *60th Annual Technical Conference and Exhibition*, https://doi.org/10.2118/14272-MS (1985).
8. Wildenschild, D. & Sheppard, A. P. X-ray imaging and analysis techniques for quantifying pore-scale structure and processes in subsurface porous medium systems. *Advances in Water Resources* **51**, 217–246, https://doi.org/10.1016/j.advwatres.2012.07.018 (2013).
9. API. Api rp 40-recommended practices for core analysis (API Washington, DC, 1998).
10. McPhee, C., Reed, J. & Zubizarreta, I.*Core analysis: a best practice guide* (Elsevier, 2015).
11. Gerchberg, R. Super-resolution through error energy reduction. *Optica Acta: International Journal of Optics* **21**, 709–720 (1974).
12. Yang, C.-Y. & Yang, M.-H. Fast direct super-resolution by simple functions. In *Proceedings of the IEEE international conference on computer vision*, 561–568 (2013).
13. de Castro Vargas Fernandes, J. *et al*. Absolute permeability estimation from microtomography rock images through deep learning super-resolution and adversarial fine tuning. *Scientific Reports* **14**, 16704, https://doi.org/10.1038/s41598-024-67367-1 (2024).
14. Feng, Y., Zhao, H., Li, X., Zhang, X. & Li, H. A multi-scale 3d otsu thresholding algorithm for medical image segmentation. *Digital Signal Processing* **60**, 186–199 (2017).
15. Liu, D. & Yu, J. Otsu method and k-means. In *2009 Ninth International conference on hybrid intelligent systems*, vol. 1, 344–349 (IEEE, 2009).
16. Vincent, L. & Soille, P. Watersheds in digital spaces: an efficient algorithm based on immersion simulations. *IEEE Transactions on Pattern Analysis & Machine Intelligence* **13**, 583–598 (1991).

17. Roerdink, J. B. & Meijster, A. The watershed transform: Definitions, algorithms and parallelization strategies. *Fundamenta informaticae* **41**, 187–228 (2000).
18. Guo, Q., Wang, Y., Yang, S. & Xiang, Z. A method of blasted rock image segmentation based on improved watershed algorithm. *Scientific Reports* **12**, 7143 (2022).
19. Vidal, A. *et al.* 16 brazilian pre-salt carbonates: multi-resolution micro-ct images. https://www.digitalrocksportal.org/projects/503 (2024).

Acknowledgements

The authors would like to thank ANP and PETREC for providing the samples, and PETROBRAS for providing financial support to this project. We would also like to thank the Conselho Nacional de Desenvolvimento Cientifico e Tecnhnolgico (CNPq) and Coordenacao de Aperfeicoamento de Pessoal de Nivel Superior (CAPES) for providing scholarships to the authors.

Author contributions

Conceptualization, A.D.V., J.C.V.F., C.E.M.A., A.G.E. and R.S.; Data Acquisition and Processing, A.D.V., A.N., J.C.V.F., L.C.M. and C.E.M.A.; Draft of the manuscript, A.D.V., A.N., J.C.V.F., L.C.M. and C.E.M.A.; All authors reviewed and approved the final manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to A.D.V.

Reprints and permissions information is available at www.nature.com/reprints.