CNN Prediction Enhancement by Post-Processing for Hydrocarbon Detection in Seismic Images

JOSÉ FABRÍCIO L. SOUZA[1], GABRIEL L. SANTANA[1], LEONARDO V. BATISTA[2], GUSTAVO P. OLIVEIRA [1], EDUARDO ROEMERS-OLIVEIRA[3], AND MOISÉS D. SANTOS[1]

[1]Petroleum Engineering Modelling Laboratory, Federal University of Paraíba, João Pessoa 58051-900, Brazil
[2]Computer Vision Laboratory, Federal University of Paraíba, João Pessoa 58051-900, Brazil
[3]Petrobras Research Center, Rio de Janeiro 21941-915, Brazil

Corresponding author: Gustavo P. Oliveira (gustavo.oliveira@ci.ufpb.br)

ABSTRACT Seismic image interpretation is indispensable for oil and gas industry. Currently, artificial intelligence has been undertaken to increase the level of confidence in exploratory activities. Detecting potentially recoverable hydrocarbon zones (leads) under the viewpoint of computer vision is an emerging problem that demands thorough examination. This paper introduces a processing workflow to recognize geologic leads in seismic images that resorts to encoder-decoder architectures of a convolutional neural network (CNN) accompanied by segmentation maps and post-processing operations. We have used seismic images collected at offshore sites of the Sergipe-Alagoas Basin (northeast of Brazil) as input. After performing a patch-based data augmentation, a total of 29600 patches were achieved. Out of these, 24000 were used for training, 5000 for validation, and 600 for testing. Each image generated for the training set was post-processed through reconstruction, thresholding – binarization and deblurring –, and outlier removal. By using the dice loss function, intersection-over-union index, and relative areal residual computed after intense cross-validation training rounds, we have shown that the accuracy of the network to detect leads was higher than 80%. Furthermore, the validation error limits were found stable within 5% - 10% in all validation rounds, thereby resulting in a fairly accurate prediction of the pre-labelled hydrocarbon spots.

INDEX TERMS Encoder-decoder, petroleum exploration, segmentation, seismic imaging.

I. INTRODUCTION

Spotting hydrocarbon reservoirs accurately is one of the leading problems faced by geologists who have the tricky responsibility to read and interpret seismic images. When seeking some evidence that spurs exploratory activities to achieve successful prospects, much effort and time are spent with the acquisition, analysis, and interpretation of geophysical and geological data [1]. In the petroleum industry, well drilling is costly and strongly dependent on human decisions. Consequently, exhaustive rounds of image interpretation are undertaken [2] to minimize potential failures that eventually are caused by subjective judgments [3]. In a few cases, this may incur in operational downsides.

Over the last decades, the traditional seismic interpretation evolved from essential qualitative analyses to consistent quantitative methods, such as post-stack amplitude analysis

(e.g. direct hydrocarbon indicators, bright-spots, dim-outs), offset-dependent amplitude analysis (AVO analysis), acoustic and elastic impedance inversion, and forward seismic modeling [4]. These methods are strongly supported by maps, such as time structure, depth, and seismic amplitude maps. In particular, self-organizing maps (SOMs) are useful for data clustering and applicable in facies mapping when information about the subsurface is scarce [5].

So far, reflection seismology is the commonplace method used to identify hydrocarbon reservoirs. In this process, the subsurface rocks are represented by differences of acoustic impedances [6]. Next, powerful algorithms are tasked to process images and recognize patterns, Besides, pattern recognition also became indispensable for the petroleum industry [7] to have fine-tuned image features.

Despite the wide class of methods available, seismic image interpretation still is sometimes undermined by the viewpoints of different interpreters. Hence, reasonable levels of precision and certainty should be always sought while

FIGURE 1. Example of architecture for a convolutional neural network aimed to classify hydrocarbon lead/no-lead regions in seismic images. Implemented in [21].

looking for economically viable hydrocarbon zones. At this point, artificial intelligence is a hand-picked resource.

Neural networks and their extension to deep neural networks have been increasing exponentially for a few years as another class of methods to deal with seismic interpretation. In particular, convolutional neural networks (CNNs) and Deep CNNs (DCNNs) have attracted much interest due to their ability for generalization [8] and feature extraction from input images by itself [9]. The latter quality is especially devoted to reducing time while searching for the most suitable features to be considered during data training.

CNNs are becoming a popular technique among people who work with computer vision for seismic imaging. Usually, most of the papers upon this subject study two major classes of procedures: patch-based image classification and image segmentation. In the former class, the input images are classified according to the seismic sample's labeling after being parsed by the network. Recent applications of CNNs involve the identification of salt bodies [10], [11], fault detection [12], [13], and seismic classification by using waveform amplitudes [14], [15].

Alternatively, the second class deals with the seismic interpretation through segmentation, which is based on encoder-decoder architectures. Encoder-decoders are unsupervised networks that operate at a self-learning pace by reducing the dimensionality of the data. This kind of architecture uses a nonlinear feature extraction method that focuses on codifying and decodifying the information to enhance their representation instead of simply assigning true/false labels to image portions [16].

Better performance of encoder-decoders over patch-based CNNs for salt body identification was verified although with high computer processing [17]. An encoder-decoder architecture was also used to overcome database training gaps for the identification of chaotic failures, where "weak labels" of the original image originated a larger image database to fulfill training requirements [18]. Similar work is reported in [19] to recover essential features for training a DCNN.

A kind of encoder-decoder architecture that received highlight over the last few years for its successful applications in biomedicine is U-net [20]. One of its main advantages is the ability to learn satisfactorily from a few images. Naturally, this positive point makes U-net architectures considerably

useful for deep learning applications, where large datasets are generally required, as with for petroleum exploration problems.

U-net is a "U-shaped" symmetric architecture formed by a contraction path and an expansion path. Along the contraction path, convolution and max pooling operations predominate. The image is reduced in dimensionality for analysis. On the other hand, the expansion path concerns to reconstruct a new image with the same dimensions as the original image.

The purpose of this paper is the identification of geological leads of hydrocarbons in petroleum reservoirs by using the U-net architecture. Here, "lead" means a geological structure whose hydrocarbon recovery potential is uncertain. In [21], the authors have presented a method focused on classification. By using multilayer perceptron networks (MLPs), CNNs and Haralick's descriptors, patches of size 20 x 20 were labeled as "lead" or "no-lead" regions as depicted in Fig. 1. However, that strategy was found limited in classifying regions globally. To put another way, a proper classification of leads was hindered due to partial comprehensions of the network while parsing the image patches. Here, we resort to U-net to improve the process of lead identification through patch-based semantic segmentation and post-processing (image reconstruction, thresholding and outlier removal). With a pixel-wise scrolling, a much more accurate prediction of the potential hydrocarbon zones is delivered. Moreover, since U-net is tailored to label whole images, its architecture becomes a suitable choice to deal with this particular problem. As it will be seen, our outcomes overperform those previously obtained with other methods.

II. METHODS
A. IMAGE ACQUISITION AND GEOLOGICAL CONTEXT
The examples of 2D seismic images used in this paper were acquired at an offshore region of the Sergipe-Alagoas (SEAL) Basin in northeast of Brazil and released by the Brazilian National Petroleum, Natural Gas and Biofuels Agency (ANP) [22]. SEAL Basin's structure is defined by an elongated asymmetric rift with extension of 350 km. Lithostratigraphically, the basement is formed by Proterozoic metamorphic rocks. Furthermore, five supersequences can be recognized: i) Paleozoic, represented by permo-carboniferous sediments; ii) Pre-Rift, made up of sandstones; iii) Rift, composed
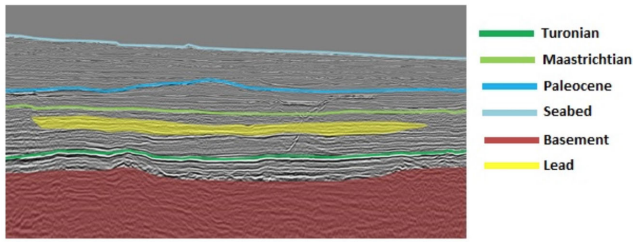
FIGURE 2. Sample of seismic image of an offshore block collected at SEAL Basin (Brazil) annotated after interpretation. Geological bounds are labeled, along with the lead region (in yellow).
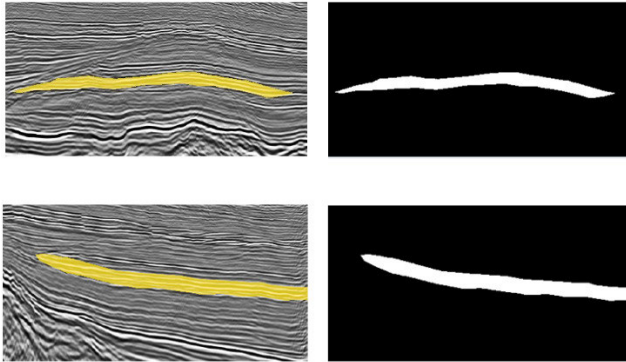

FIGURE 3. Examples of original seismic annotated images (on the left side) in contrast with versions binarized and resized (on the right side) to label each pixel as "lead" (white ones) or "no-lead" (black ones).

by shale and lacustrine sandstones; iv) Post-Rift, formed by coarse-grained siliciclastics, evaporites, carbonates, and shale; v) Drift, comprised by an interval predominantly carbonatic which is followed by a second interval mainly clastic formed by sandstones, carbonates, and shales with turbiditic sandstones. The hydrocarbon leads presented in this paper (highlighted in yellow at Figures 2 and 3) belong to Drift supersequence and correspond to turbiditic sandstones from Upper Cretaceous of the Calumbi Formation with porosity from 15 to 20% sealed by shales from the same Formation.

B. IMAGE PRE-PROCESSING

All the images used as input data were pre-processed to deliver a proper flow for training. The pre-processing stage can be organized into three major steps: cleaning, patch generation, and data augmentation, with their technicalities. In the first step, a simple binarization was applied to the images to leave them with two labels only, thus identifying a given pixel as "lead" or "no-lead". This way, the original annotations of the geological bounds were eliminated to purge all the no-interest areas of the image and maintain the hydrocarbon region as interpreted. *Lead pixels* were colored in white, whereas *no-lead pixels* were colored in black, resulting in a binary mask as depicted in Fig. 3.

Secondly, we have expanded the image bank by using a technique of patch generation based on the concept of *lead-pixel local area*. To make this clear, suppose that a window of size 3×3 is sliding over an image with stride 2. Every time that the sliding window finds a dominant "white area"
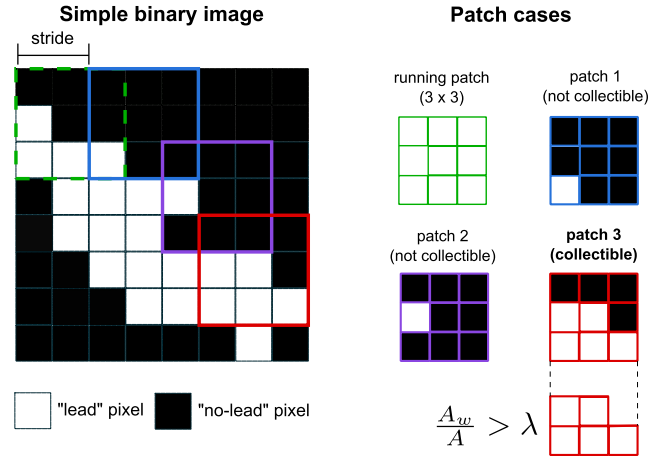

FIGURE 4. Process of patch generation from a simple binary image and a 3×3 patch model running with stride of 2. New image patches (subimages) are collectible to form the larger patch bank used for training only if the area of white pixels (lead region) overlaps the running patch area by a proportion above λ. In this example, two (blue and purple) of the three patches are non-collectible because they do not satisfy the criterion of the lead-pixel local area, but one of them is collectible (red).

of the covered area by a factor higher than λ, we collect this patch area and generate a new smaller image (subimage). The process is schematized in Fig. 4 for a simple image already binarized. Two cases of non-collectible patches and one case of collectible patch are illustrated. Here, we assume that the "white area" A_w is the portion of each window that overlaps a lead region, which is given by the sum of all the lead-pixels lying in the patch area A. Hence, a patch is collectible if $A_w/A > \lambda$, which defines the *lead-pixel local area*. In this paper, patches empirically defined with sizes of 80×80, strides of 2, and $\lambda = 0.2$ were the parameters selected. This threshold of λ, equivalent to 20% of no-lead pixels, was found after some pre-tests so allowing us to get a reasonable number of patches for posterior execution by the network.

Thirdly, to expand the image bank, data augmentation operations were applied to the patches, such as rotations and shearing limited by 15°, horizontal flipping, and scaling restricted to ±10%. Out of 29600 patches generated, 24000 were used for training, 5000 for validation, and others 600 exclusively for testing.

C. U-NET CONFIGURATION

Although several CNN architectures for segmentation are available, we have opted for U-net because of its powerful ability for generalization in low-variance domains. That is to say, U-net architectures can capture fine details of the images quickly in banks whose images are slightly different from each other. Even with a limited number of images, U-net performs effectively under training. This is the reason why U-nets are being recognized as an appropriate and innovative instrument for unsupervised seismic interpretation in petroleum-related applications [23]. The U-net base model [20] is represented in Fig. 5.

Because of the low dimensionality and scarcity of the input images, overfitting signs soon arise when the U-net
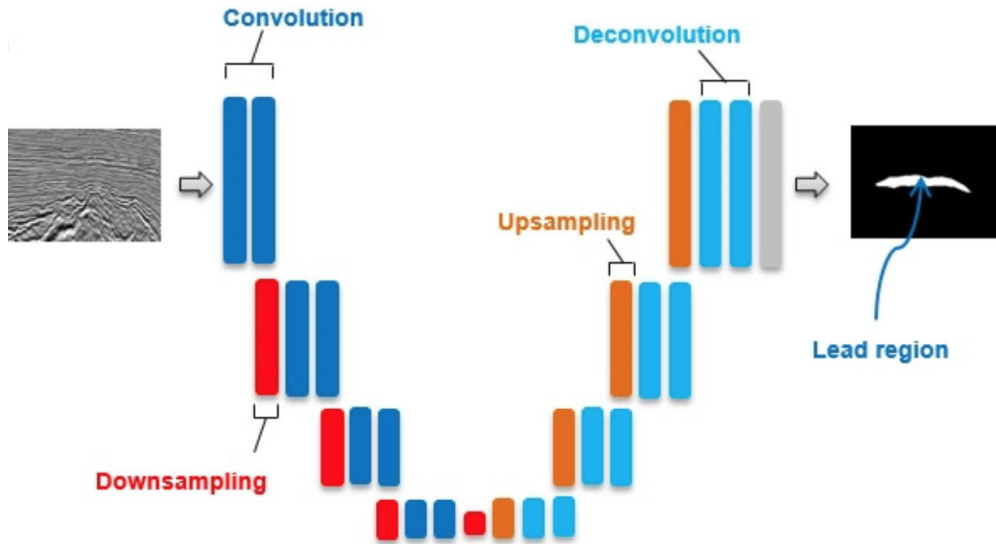
FIGURE 5. Seminal model of the U-net architecture adapted from [20] that highlights the operational walkthrough. Blue boxes correspond to multi-channel feature maps.

TABLE 1. Configuration of the U-net architecture used in this paper for training runs, which has considered a total of 1,940,817 parameters.

layer	kernel size	pool	output size
input	-	-	80 x 80 x 1
flat-conv1 + dropout = 0.3	3 x 3	-	80 x 80 x 16
down-conv1	3 x 3	2 x 2	40 x 40 x 16
flat-conv2 + dropout = 0.3	3 x 3	-	40 x 40 x 32
down-conv2	3 x 3	2 x 2	20 x 20 x 32
flat-conv3 + dropout = 0.4	3 x 3	-	20 x 20 x 64
down-conv3	3 x 3	2 x 2	10 x 10 x 64
flat-conv4 + dropout = 0.4	3 x 3	-	10 x 10 x 128
down-conv4	3 x 3	2 x 2	5 x 5 x 128
flat-conv5 + dropout = 0.5	3 x 3	-	5 x 5 x 256
up-conv1 + concatenate w/ flat-conv4	2 x 2	1/2 x 1/2	10 x 10 x 256
flat-conv6 + dropout = 0.4	3 x 3	-	10 x 10 x 128
up-conv2 + concatenate w/ flat-conv3	2 x 2	1/2 x 1/2	20 x 20 x 128
flat-conv7 + dropout = 0.4	3 x 3	-	20 x 20 x 64
up-conv3 + concatenate w/ flat-conv2	2 x 2	1/2 x 1/2	40 x 40 x 64
flat-conv8 + dropout = 0.3	3 x 3	-	40 x 40 x 32
up-conv4 + concatenate w/ flat-conv1	2 x 2	1/2 x 1/2	80 x 80 x 32
flat-conv9 + dropout = 0.3	3 x 3	-	80 x 80 x 16
output	1 x 1	-	80 x 80 x 1

works on its standard form. To surmount this difficulty, one recommends that the number of filters per layer reduces and further dropout at the convolutional layers applies [24]. In this paper, the configuration used for the U-net architecture is systematically described in Table 1.

D. ACTIVATION FUNCTIONS

Acceptable results are obtainable in deep learning applications that implement the ReLU (Rectified Linear Unit) function, or even better with variants, such as the ELU (Exponential Linear Unit) [25]. The ELU function may have a superior performance in seismic image applications when compared to the ReLU function as its usage prevents intricacies with the so-called ''dead nodes'' (network nodes whose response will always vanish) and speeds up the convergence [23]. For this reason, we have set up the activation function of all the U-net layers to be ELU and only the output layer

has stood with the logistic function. Both ReLU and ELU functions are defined in Eq. 1 and plotted just below in Fig. 6.

$$\text{ReLU}(x) = \max(0, x) \tag{1}$$

$$\text{ELU}(x) = \begin{cases} \alpha(\exp(x) - 1), & \text{if} \quad x \leq 0 \\ x, & \text{if} \quad x > 0 \end{cases} \tag{2}$$

E. QUALITY ASSESSMENT

To assess the quality level of the network outputs, we have calculated the Intersection-over-Union score (IoU) – also known as Jaccard's Similarity Index – since it is a popular metric amongst image segmentation experts [26]–[28]. The IoU score measures the overlap ratio between the predicted image and the ground truth. A perfect match would imply $IoU = 1.0$. An illustration of the IoU concept is provided in Fig. 7.

The notion of the IoU is revamped when resorting to a confusion matrix dependent of the number of predicted and true labels. By considering $p_i, i = 1, 2, \ldots, n$ a pixel of the input image and n the total number of pixels, we can define, respectively, $y_i = y(p_i)$ as the binary value of the pixel p_i predicted by the network and $\hat{y}_i = \hat{y}(p_i)$ as that of the true image. Therefore, $y_i, \hat{y}_i \in \{0, 1\}$ and the following 2×2 matrix can be written as

	$\hat{y}_i \geq \tau$	$y_i \leq \tau$
$y_i = 1$	TP	FN
$y_i = 0$	FP	TN

where *TP*, *FN*, *FP* and *TN* stand for true positive, false negative, false positive and true negative, in this order. The parameter $\tau \in \mathbb{R}$ is a control threshold. This way, taking into account the number of pixels # for each class, the IoU reads as

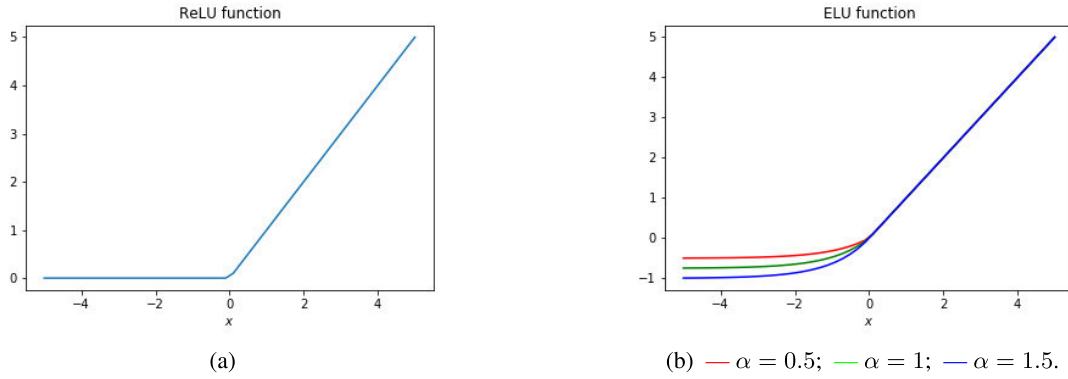$$IoU = \frac{\#TP}{\#TP + \#FN + \#FP}. \tag{3}$$

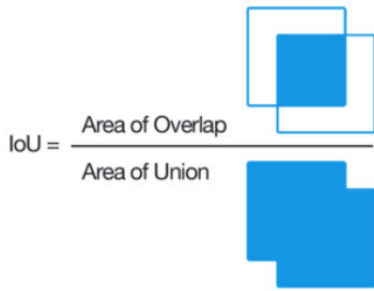**FIGURE 6.** Usual activation functions: (a) ReLU; (b) ELU.

FIGURE 7. Sketch of the geometric idea behind the Intersection-over-Union (IoU) score, or Jaccard's Similarity Index). The more the predicted image overlaps the true image, the more the IoU score approaches the maximum admissible value of 1.0.

Therewith, two behaviors are expected. On one hand, if $\#FN \to 0.0$ and $\#FP \to 0.0$, then $IoU \to 1.0$ and the level of prediction of the network enhances. On the other hand, if $\#FN \gg \#TP$ or $\#FP \gg \#TP$, then $IoU \to 0.0$ and the level of prediction of the networks worsens. Furthermore, the dice loss – or F_1-score – was used as loss function during the training [29], [30]. Likewise IoU, the dice loss function is defined as:

$$\text{dice} = \frac{2\#TP}{2\#TP + \#FN + \#FP}. \qquad (4)$$

F. TRAINING INFORMATION
The code was run for a period of 250 epochs by using Adam's optimizer [31] starting with an initial learning rate of 10^{-4}. At the end of each new epoch, the IoU was computed to measure the network's performance over the validation dataset and if it persisted on a plateau state (i.e. no improvements on IoU) for 25 epochs, the learning rate was halved. After verifying if the network's performance at the most updated epoch had been found better than the best performance among all those at the previous epochs, the current epoch was stored as the latest network.

G. POST-PROCESSING
The U-net handles patches and delivers segmentation maps over these input patches. This training process, however, requires further intervention to recover the whole image and polish the identification of the lead regions. Likewise the pre-processing stage, the post-processing stage is subdivided into minor steps, namely, image reconstruction, thresholding, and outlier removal. They are explained in detail below.

1) IMAGE RECONSTRUCTION
Every patch parsed by the U-net during the training flow will be, later, a small piece of the reconstructed image. The image reconstruction follows an incremental assembling of adjoining patches based on strides that can be described by recursive unions. By considering $\hat{P}$ the running window and σ an abstract operator that stores a new patch and shifts the running window by a stride of s scrolling over the image,[1] it is possible to recombine each piece of a certain image I sequentially, so that its reconstruction is given by

$$
\begin{aligned}
I_0 &= \sigma(I, \hat{P}) \\
I_1 &= I_1 \cup \sigma(I, \hat{P} + s) \\
I_2 &= I_2 \cup \sigma(I, \hat{P} + 2s) \\
&\vdots \quad \vdots \quad \vdots \\
I_j &= I_{j-1} \cup \sigma(I, \hat{P} + js)
\end{aligned}
$$

where each $I_1, I_2, \ldots, I_j$ is an image piece obtained cumulatively through sequential strides and additions, as opposed to the first patch I_0, for which $s = 0$. This way, Eq.(5) results in the whole image, fully reconstructed, when $k = j$.

$$I_k = I_{k-1} \cup \sigma(I, \hat{P} + ks), \quad k = 0, 1, 2, \ldots, j. \qquad (5)$$

Figure 8 illustrates the reconstruction process.

2) THRESHOLDING
Due to the successive intersections implied by the reconstruction, much overlap of patches is experienced, mainly at the central portions of the image where the running patch has a higher frequency of visitation. This overlap triggers a blurring effect (see image I_j in Fig.8) whose consequence is the rise of false positives. To restrain the occurrence of false positives,

---

[1]Rigorously, the stride shifts the patch from left to right and from top to down. At this point, without losing generality, one assumes that this two-dimensional displacement is implicit on s.
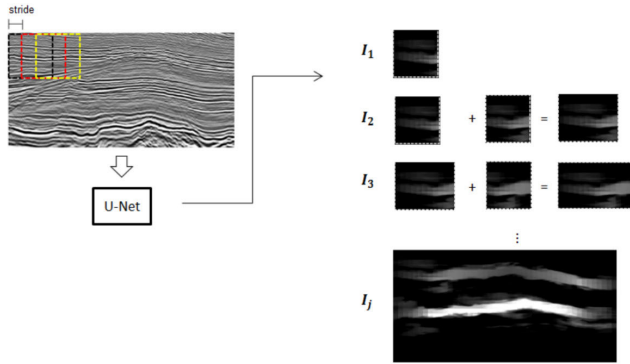
FIGURE 8. Image reconstruction process: the original image is parsed by the U-net in the form of smaller 80 × 80 patches. Once segmented, the patches are joined by a succession of compositions based on a 10-stride scrolling that recovers the whole image through incremental additions.
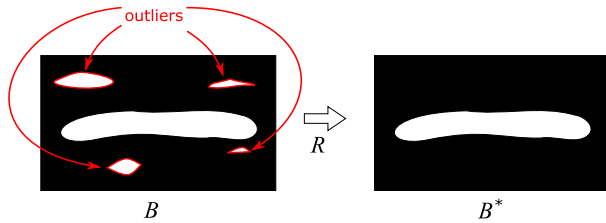

FIGURE 9. Outlier removal process in a segmented image after thresholding. The smaller marked pixel areas are eliminated, whereas the largest one remains as the unique "lead" region. The image on the left (right) is $\mathcal{B}$ ($\mathcal{B}^*$).
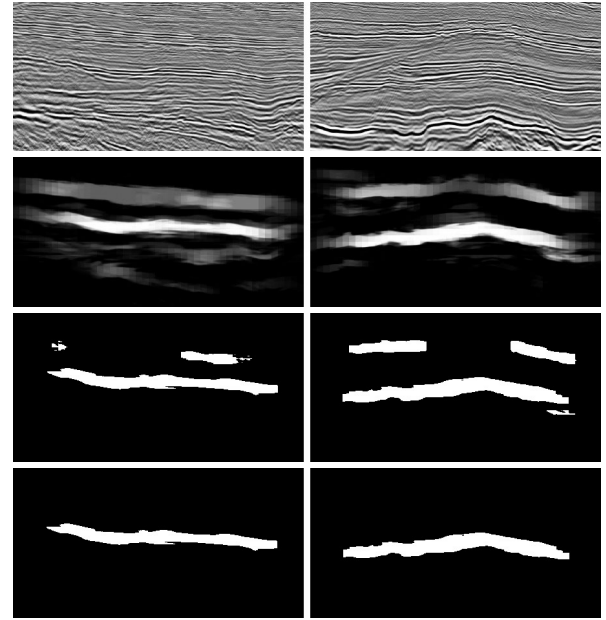

FIGURE 10. Post-processing operations for two testing images (labelled as 4, at the left, and 8, at the right) parsed by the U-net: raw seismic images for reference (1st. row); images after reconstruction (2nd. row); images after thresholding (binarization and deblurring) (3rd. row); images after outlier removal (4th. row).

we have applied a thresholding technique that has the twofold mission of binarizing and deblurring the output image.

It should be underlined at this point that the U-net segmentation delivers a scalar field that associates a normalized value to each pixel of the reconstructed image, namely the *confidence index* (*CI*). As the *CI* is the variable subjected to the blurring effect, the proposed thresholding will map the field *CI* onto a new binary mask so deblurring the output image. Here, the thresholding is implemented by choosing the threshold $\tau_0 = 0.5$ because it is an average value understood as a reasonable starting point.

Suppose that $\mathcal{B}^0$ is the binary representation of one of the original annotated image and $\mathcal{G}$ is the reconstructed image. As we have said above, $\mathcal{G}$ is, in fact, a grayscale blurred image. To make its binary version, we define a thresholding function $\mathcal{L}_k$ over $\mathcal{G}$, given by

$$\mathcal{L}_k(p) = \begin{cases} 0, & \text{if } CI(p) < \tau_k \\ 1, & \text{if } CI(p) \geq \tau_k, \end{cases} \quad (6)$$

for each pixel p of the image $\mathcal{G}$ and a threshold value τ_k determined incrementally, so that $\mathcal{L}_k : \mathcal{G} \to \mathcal{B}_k$.

3) OUTLIER REMOVAL
The third step of post-processing concerns an elimination of outliers that possibly remain after applying the thresholding

function. In this context, outlier is a borrowed term of statistics that is equivalent to a disjoint set of pixels that survive to the elimination due to remaining "isolated islands" where the confidence index is found above the cutoff point τ_k. As such, they cause false impressions of lead portions. In the current study, we have firstly used an algorithm for identification of connected components available in the Python module `scikit-image` to separate the regions. Then, we have performed the outlier removal automatically after a simplified segmentation. Given the bedding and local continuity features expected of lead regions observed in the images of our bank, we have assumed that only the component with the largest pixel area should be kept.[2] All the remnant components were eliminated. Although the largest area may not be the best criterion to choose, we have verified that the number of outliers was found much smaller in relation to the largest component, as with their areas. More formally, the removal operation can be represented by a function, let us say $\mathcal{R}$, that takes a binary thresholded image $\mathcal{B}$ and gives back a cleaner image B^*. When the computation of the *IoU* value is performed over $\mathcal{B}$, it is affected by influence of the outliers. After removing them, a more accurate estimation of the *IoU* is obtained with B_*. Figure 9 depicts the outlier removal for a sample image.

III. RESULTS
A. POST-PROCESSING SHOWCASE
In this section, we show examples of application of the post-processing steps over two testing images since the

---

[2] The simplified segmentation yields a unique integer label for each disjoint set. The lead area is the set whose label is the most frequent one (the statistical mode).
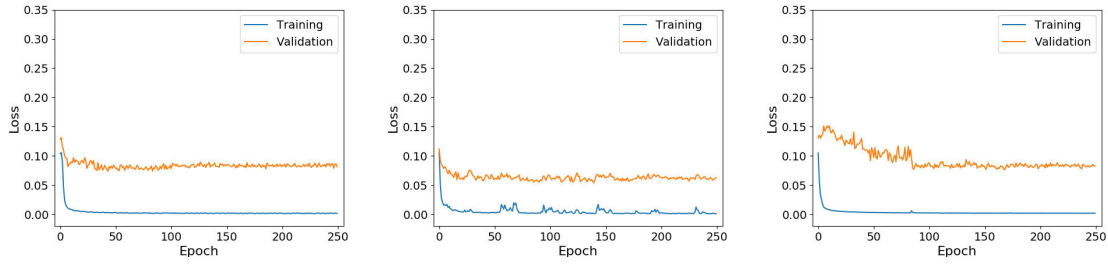
FIGURE 11. Normalized dice loss profiles for training and validation computed after 250 epochs per round. In all cases, the validation loss stabilizes at values within 5% - 10%.

TABLE 2. Accuracy and error of validation obtained per training round.

round	accuracy	error
1	0.8425	0.0739
2	0.8848	0.0538
3	0.8370	0.0771

pre-processing is based on usual cleaning and resizing operations. During the post-processing, as explained along the paper, after the training step, the images undergo operations of reconstruction, thresholding (binarization and deblurring), and outlier removal. A few results for these two images (labelled as 4 and 8, from left to right) are illustrated in Figure 10. Each row represents the sequential post-processing operations, except the first one, which brings out the raw seismic image for the reader's reference.

B. CROSS-VALIDATION

To ensure that the U-net performs successfully and is able to generalize its predictions, we have carried out three rounds of cross-validation, a process that consists of repeating the training steps with exchanging the training, testing, and validation sets each other. This way, the occurrence of a biased dataset is overturned. Table 2 outlines the accuracy and error of validation obtained during each training round. Additionally, we have plotted in Fig. 11 the normalized dice loss obtained after carrying out the code for each training round. By taking 250 epochs, the computational time spent to run all epochs in a round was kept on the average of approximately 4 hours. As seen, the validation loss stabilizes at values within 5% - 10%, thus showing that the U-net results are fairly accurate for the intended application.

C. IoU ENHANCEMENT

The outlier removal operation has caused a significant enhancement of the *IoU* ratio. Since it represents a pixel-wise certitude percentage of the region of interest, as it grows, more pixels of lesser interest are unconsidered. Table 3 compares the *IoU* values for three test images obtained before the removal operation (with outliers) and after the removal (with no outliers). For all cases, the average of hits of the U-net was above 60%, except for a slight reduction experienced by Image 8.
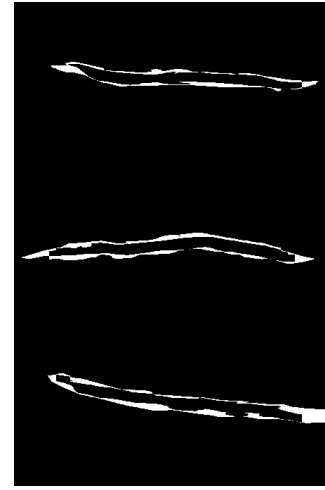

FIGURE 12. Segmentation residuals (pixel-wise white areas) taken when comparing the ground-truth image and the ultimate image after being parsed by the U-net. The thin fringe area around the lead portion shows an evidence that the network's prediction is consistent with the original interpretation.

TABLE 3. Enhancement of *IoU* values computed before the removal operation (with outliers) and after the removal (with no outliers).

test image	IoU with outliers	IoU without outliers
4	0.5517	0.6633
7	0.4279	0.6279
8	0.3974	0.5947

D. RELATIVE AREAL RESIDUALS

Another perspective that brings enrichment to the method is exhibited by segmentation residuals. The areal residuals are pixel-wise differences existing between the white area occupied by the ground-truth image and the post-processed one. Quantitatively, we can estimate the relative areal residual for each testing round previously discussed in percentage terms as simply as

$$RAR = (1 - IoU^*) \times 100\%, \tag{7}$$

where IoU^* is computed after the outlier removal operation (third column of Table 2). Therefore, $RAR = \{33.67\%, 37.21\%, 40.53\%\}$, for the rounds 4, 7 and 8, respectively. A visual representation of the differential lead areas are depicted in Fig. 12. As seen, the respective errors are consistently concentrated around the region of interest as
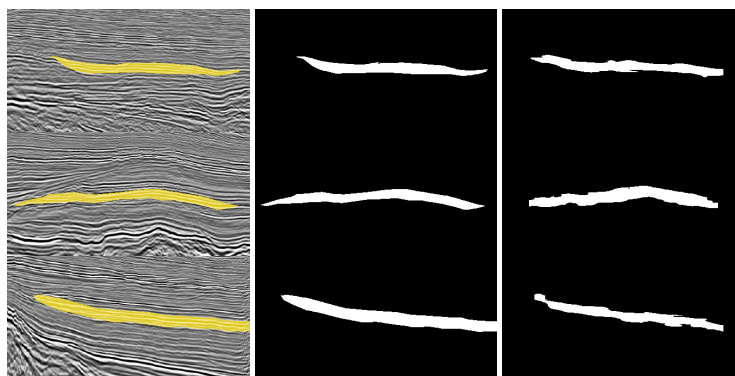
FIGURE 13. Synoptic image frame of the workflow implemented in this paper highlighting three distinct images of the SEAL Basin: original annotated images (1st. column), binary versions of the original images (2nd. column), and post-processed versions of the images parsed by the U-net (3rd. column). The last white pixel area strongly agrees with the region previously interpreted (in yellow).

expected, thereby ensuring a reliable interpretation given the tolerance margin.

E. SYNOPTIC IMAGE FRAME

To summarize the main capabilities of segmentation of the U-net, we have organized a final synoptic image frame that provides a qualitative overview of the methodology. Figure 13 is a synopsis of the implemented workflow from the original annotated image (1st. column), going through its binary version (2nd. column) until getting the ultimate post-processed version (3rd. column) for three images.

IV. CONCLUSION

We have proposed a workflow for identification of potential hydrocarbon-bearing accumulations recognized as geological leads in seismic images. The image segmentation process was performed by a convolutional neural network equipped with encoder-decoder, namely the U-net, with varying architectures.

The advantage of working with segmentation instead of classification is due to the higher accuracy of the former technique in relation to the latter. For segmentation, the U-net is particularly interesting because of its ability for generalization with only a few training images.

In this paper, we have shown that reasonable accuracy can be achieved when some pre- and post-processing operations are performed. Although some limitations of the method still prevail due to the iterative workflow, such as the reconstruction overlaps, the general path to achieve an ultimate prediction is straightforward and compensatory. Future improvements should require a larger image bank, suppress overblurring effects, and possibly eliminate the need to deal with remnant outliers. Moreover, a series of seismic attributes must be brought into the network's crux for better prediction.

REFERENCES

[1] A. Roy, A. S. Romero-Peláez, T. J. Kwiatkowski, and K. J. Marfurt, "Generative topographic mapping for seismic facies estimation of a carbonate wash, veracruz basin, southern mexico," *Interpretation*, vol. 2, no. 1, pp. SA31–SA47, Feb. 2014, doi: 10.1190/int-2013-0077.1.

[2] C. Song, Z. Liu, Y. Wang, X. Li, and G. Hu, "Multi-waveform classification for seismic facies analysis," *Comput. Geosci.*, vol. 101, pp. 1–9, Apr. 2017, doi: 10.1016/j.cageo.2016.12.014.

[3] T. Zhao, V. Jayaram, A. Roy, and K. J. Marfurt, "A comparison of classification techniques for seismic facies recognition," *Interpretation*, vol. 3, no. 4, pp. SAE29–SAE58, Nov. 2015, doi: 10.1190/int-2015-0044.1.

[4] P. Avseth, T. Mukerji, and G. Mavko, *Quantitative Seismic Interpretation: Applying Rock Physics Tools to Reduce Interpretation Risk*. Cambridge, U.K.: Cambridge Univ. Press, 2005, doi: 10.2113/gscanmin.43.5.1801.

[5] M. M. Saggaf, M. N. Toksöz, and M. I. Marhoon, "Seismic facies classification and identification by competitive neural networks," *Geophysics*, vol. 68, no. 6, pp. 1984–1999, Nov. 2003, doi: 10.1190/1.1635052.

[6] M. C. de Matos, P. L. Osorio, and P. R. Johann, "Unsupervised seismic facies analysis using wavelet transform and self-organizing maps," *Geophysics*, vol. 72, no. 1, pp. P9–P21, Jan. 2007, doi: 10.1190/1.2392789.

[7] H.-K. Du, J.-X. Cao, Y.-J. Xue, and X.-J. Wang, "Seismic facies analysis based on self-organizing map and empirical mode decomposition," *J. Appl. Geophys.*, vol. 112, pp. 52–61, Jan. 2015, doi: 10.1016/j.jappgeo.2014.11.007.

[8] Z. Wang, H. Di, M. A. Shafiq, Y. Alaudah, and G. AlRegib, "Successful leveraging of image processing and machine learning in seismic structural interpretation: A review," *Lead. Edge*, vol. 37, no. 6, pp. 451–461, Jun. 2018, doi: 10.1190/tle37060451.1.

[9] I. Goodfellow, Y. Bengio, A. Courville, and Y. Bengio, *Deep Learning*, vol. 1. Cambridge, MA, USA: MIT Press, 2016.

[10] A. U. Waldeland, A. C. Jensen, L.-J. Gelius, and A. H. S. Solberg, "Convolutional neural networks for automated seismic interpretation," *Lead. Edge*, vol. 37, no. 7, pp. 529–537, Jul. 2018, doi: 10.1190/tle37070529.1.

[11] H. Di, Z. Wang, and G. AlRegib, "Why using CNN for seismic interpretation? An investigation," in *Proc. SEG Tech. Program Expanded Abstr.*, Aug. 2018, pp. 2216–2220, doi: 10.1190/segam2018-2997155.1.

[12] X. Wu, Y. Shi, S. Fomel, and L. Liang, "Convolutional neural networks for fault interpretation in seismic images," in *Proc. SEG Tech. Program Expanded Abstr.*, Aug. 2018, pp. 1946–1950, doi: 10.1190/segam2018-2995341.1.

[13] A. Cunha, A. Pochet, H. Lopes, and M. Gattass, "Seismic fault detection in real data using transfer learning from a convolutional neural network pretrained with synthetic seismic data," *Comput. Geosci.*, vol. 135, Feb. 2020, Art. no. 104344, doi: 10.1016/j.cageo.2019.104344.

[14] M. Araya-Polo, T. Dahlke, C. Frogner, C. Zhang, T. Poggio, and D. Hohl, "Automated fault detection without seismic processing," *Lead. Edge*, vol. 36, no. 3, pp. 208–214, Mar. 2017, doi: 10.1190/tle36030208.1.

[15] A. Pochet, P. H. B. Diniz, H. Lopes, and M. Gattass, "Seismic fault detection using convolutional neural networks trained on synthetic poststacked amplitude maps," *IEEE Geosci. Remote Sens. Lett.*, vol. 16, no. 3, pp. 352–356, Mar. 2019, doi: 10.1109/lgrs.2018.2875836.

[16] L. Deng and D. Yu, "Deep learning: Methods and applications," *Found. Trends Signal Process.*, vol. 7, nos. 3–4, pp. 197–387, Jun. 2014, doi: 10.1561/2000000039.

[17] T. Zhao, "Seismic facies classification using different deep convolutional neural networks," in *Proc. SEG Tech. Program Expanded Abstr.*, 2018, pp. 2046–2050, doi: 10.1190/segam2018-2997085.1.

[18] H. Di, Z. Wang, and G. AlRegib, "Deep convolutional neural networks for seismic salt-body delineation," in *Proc. AAPG Annu. Conv. Exhib.*, 2018, pp. 1–7. [Online]. Available: http://www.searchanddiscovery.com/pdfz/documents/2018/70360di/ndx_di.pdf.html

[19] Y. Shi, X. Wu, and S. Fomel, "Automatic salt-body classification using deep-convolutional neural network," in *Proc. SEG Tech. Program Expanded Abstr.*, 2018, pp. 1971–1975, doi: 10.1190/segam2018-2997304.1.

[20] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Proc. Int. Conf. Med. image Comput. Comput.-Assist. Intervent.* Tulsa, OK, USA: Springer, 2015, pp. 234–241, doi: 10.1007/978-3-319-24574-4_28.

[21] J. F. L. Souza, M. D. Santos, R. M. Magalháes, E. M. Neto, G. P. Oliveira, and W. L. Roque, "Automatic classification of hydrocarbon-leads-in seismic images through artificial and convolutional neural networks," *Comput. Geosci.*, vol. 132, pp. 23–32, Nov. 2019, doi: 10.1016/j.cageo.2019.07.002.

[22] B. Haeser, "National agency of petroleum, natural gas and biofuels, super-intendence of definition of blocks: Geologic summary and block bidding," Brazilian Nat. Agency Petroleum, Natural Gas Biofuels, Rio de Janeiro, Brazil, Tech. Rep., 2015.

[23] Y. Zeng, K. Jiang, and J. Chen, "Automatic seismic salt interpretation with deep convolutional neural networks," in *Proc. 3rd Int. Conf. Inf. Syst. Data Mining*, 2019, pp. 16–20, doi: 10.1145/3325917.3325926.

[24] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: A simple way to prevent neural networks from overfitting," *J. Mach. Learn. Res.*, vol. 15, no. 1, pp. 1929–1958, 2014.

[25] D. Pedamonti, "Comparison of non-linear activation functions for deep neural networks on MNIST classification task," 2018, *arXiv:1804.02763*. [Online]. Available: http://arxiv.org/abs/1804.02763

[26] L. P. Coelho, A. Shariff, and R. F. Murphy, "Nuclear segmentation in microscope cell images: A hand-segmented dataset and comparison of algorithms," in *Proc. IEEE Int. Symp. Biomed. Imag.*, Dec. 2009, pp. 518–521, doi: 10.1109/ISBI.2009.5193098.

[27] C. V. Narayana, E. Sreenivasa Reddy, and M. Seetharama Prasad, "Automatic image segmentation using ultra fuzziness," *Int. J. Comput. Appl.*, vol. 49, no. 12, pp. 6–13, Jul. 2012, doi: 10.5120/7677-0977.

[28] V. Iglovikov and A. Shvets, "Ternausnet: U-net with vgg11 encoder pre-trained on imagenet for image segmentation," 2018, *arXiv:1801.05746*. [Online]. Available: https://arxiv.org/abs/1801.05746

[29] R. Lguensat, M. Sun, R. Fablet, P. Tandeo, E. Mason, and G. Chen, "Eddynet: A deep neural network for pixel-wise classification of oceanic eddies," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Jun. 2018, pp. 1764–1767, doi: 10.1109/IGARSS.2018.8518411.

[30] L. Fidon, W. Li, L. C. Garcia-Peraza-Herrera, J. Ekanayake, N. Kitchen, S. Ourselin, and T. Vercauteren, "Generalised wasserstein dice score for imbalanced multi-class segmentation using holistic convolutional networks," in *Proc. Int. MICCAI Brainlesion Workshop*. Cham, Switzerland: Springer, 2017, pp. 64–76, doi: 10.1007/978-3-319-75238-9_6.

[31] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, *arXiv:1412.6980*. [Online]. Available: http://arxiv.org/abs/1412.6980

JOSÉ FABRÍCIO L. SOUZA received the B.Sc. and M.Sc. degrees in mathematics and the D.Sc. degree in mechanical engineering from the Federal University of Paraíba, João Pessoa, Brazil, in 1998, 2000, and 2019, respectively. He is currently a Professor with the Exact Sciences Department, Federal University of Paraíba. He is also a Collaborator with the Petroleum Engineering Modelling Laboratory. His research interests include statistics, probability, digital inclusion, and artificial intelligence.

GABRIEL L. SANTANA is currently pursuing the degree in computer science with the Federal University of Paraíba. He is also an Assistant Researcher with the Petroleum Engineering Modelling Laboratory. His research interests include mobile deployment, computer graphics, and artificial intelligence.

LEONARDO V. BATISTA received the bachelor's degree in electrical engineering from the Pontifical Catholic University of Rio de Janeiro and the Ph.D. degree in electrical engineering from the Federal University of Campina Grande, in 2002. He has been with the Federal University of Paraíba, Brazil, since 1996, where he is currently a Full Professor with the Department of Informatics. He has coordinated several government-funded research projects. He has published approximately 15 journal articles and 80 complete conference papers. He has also been an invited speaker and a program committee member of several scientific conferences. He has contributed to two books in his areas of interest and co-edited another. His research interests include digital signal and image and video processing and analysis.

GUSTAVO P. OLIVEIRA received the Lic. degree in mathematics from Federal Fluminense University, Rio de Janeiro, Brazil, in 2009, the M.Sc. degree in metallurgical and materials engineering from the Federal University of Rio de Janeiro-COPPE, in 2011, and the D.Sc. degree in mechanical engineering from the State University of Rio de Janeiro, in 2015.

He is currently an Adjunct Professor with the Federal University of Paraíba, Brazil. He is also a member with the Scientific Computing Department, a Permanent Member with the Graduate Program in mechanical engineering, and a Co-Leader with the Petroleum Engineering Modelling Laboratory. His research interests include numerical methods, fluid mechanics, scientific computing, reservoir modeling, and data science. He received the Postdoctoral Fellowship (Score Ten) from FAPERJ, Brazil, and PNPD/CNPQ, Brazil, Grantee in Computational and Mathematical Modeling.

EDUARDO ROEMERS-OLIVEIRA received the B.Sc. degree in geology from the Federal University of Paraná, Brazil, in 2008, and the M.Sc. degree in regional geology from the State University of São Paulo, Rio Claro, Brazil, in 2014. He is currently a Researcher with the Petrobras Research Center (CENPES). His research interests include reservoir modeling, carbonate rocks, sedimentology, and stratigraphy.

MOISÉS D. SANTOS received the B.Sc. degree in mathematics from the Federal University of Paraíba, Brazil, in 2002, the M.Sc. degree in mathematics from the Federal University of Campina Grande, Brazil, in 2005, and the D.Sc. degree in petroleum science and engineering from the Federal University of Rio Grande do Norte, in 2010.

He is currently an Associate Professor with the Federal University of Paraíba. He is also a member with the Scientific Computing Department, a Permanent Member with the Graduate Program in mechanical engineering, and a Co-Leader with the Petroleum Engineering Modelling Laboratory. His research interests include thermal oil recovery methods, reservoir characterization, seismic processing, and artificial intelligence.

• • •