Seismic Facies Segmentation Using Convolutional Neural Networks

1st Jonh Lemos
GAIA
Universidade Federal da Bahia
Salvador, Brazil
jonh.brian@ufba.br

2nd Lorena Santos
GAIA
Universidade Federal da Bahia
Salvador, Brazil
lorena.santos@ufba.br

3rd Alexsandro Cerqueira
GAIA
Universidade Federal da Bahia
Salvador, Brazil
alexsandrocerqueira@ufba.br

***Abstract*—Seismic facies segmentation is a crucial task for hydrocarbon exploration and reservoir characterization. However, due to the high amount of data involved in geophysics, manual seismic labeling is very time-consuming and sometimes prone to errors due to its subjective nature. Artificial intelligence algorithms, especially deep learning techniques, have had a significant impact on seismic data analysis. Nevertheless, many techniques require large amounts of labeled data, often becoming unfeasible in the context of seismic segmentation. This study proposes a workflow that uses only 10% of the data for training, which is generally the amount labeled in real-world analysis, to predict the remaining sections. We employed a U-Net architecture to train two models on the Parihaka 3D seismic survey, one for the inline orientation and another for the crossline orientation. The data processing included normalization, patching, and the merging of the two prediction cubes based on the maximum probability of their classes. The proposed method achieved an accuracy of 95% and a Mean IoU of 87% on the merged cube. These results demonstrate that the U-Net architecture was able to extract key features from seismic data and provide accurate labels. This approach represents a practical and efficient tool to accelerate the seismic labeling process while generating reliable results.**

***Index Terms*—Seismic Facies Segmentation, Seismic Data, Convolution Neural Networks, U-Net**

I. Introduction

Seismic data interpretation is a critical step in hydrocarbon exploration, particularly for reservoir characterization and stratigraphic analysis. One key task in this process is seismic facies classification, which involves segmenting seismic volumes into different classes based on signal characteristics such as amplitude, frequency, reflector continuity, and geometry. Traditionally, this task is performed manually by geoscientists who examine seismic sections, identify patterns, and assign labels to regions based on geological knowledge. However, the increasing size and complexity of seismic datasets make manual interpretation a time-consuming and often subjective process [1].

In recent years, Convolutional Neural Networks (CNNs) have shown remarkable success in image segmentation tasks and are now being widely applied to seismic data interpretation. CNNs can automatically extract hierarchical features from seismic images, enabling more efficient and consistent identification of geological structures as faults [2] and salt structures [3]. This approach allows interpreters to label a small portion of the data and use the trained network to generalize segmentation across larger areas, reducing interpretation time and improving reliability [4], [5].

This study investigates the use of CNNs for seismic facies segmentation, focusing on improving the objectivity and reproducibility of the interpretation workflow. We apply the methodology to the Parihaka 3D seismic dataset (Figure 1), acquired in the Taranaki Basin offshore New Zealand. The dataset includes 590 inlines and 782 crosslines, with annotated labels that serve as ground truth for training and evaluating the model. The goal is to demonstrate the potential of deep learning to automate and enhance seismic facies analysis in a real-world context.

II. Geophysical Context

A. Seismic Method

The seismic method is the most well-known geophysical technique regarding investment, operational cost, and the number of professionals involved, particularly in the oil and gas industry. Its popularity is attributed to its high resolution, deep penetration, versatility, and ability to produce accurate subsurface images [6], [7].

The seismic method acquisition involves the controlled generation of elastic waves using controlled energy sources such as vibroseis trucks or explosives and recording returning wavefields with arrays of geophones or hydrophones [8]. These waves propagate through the Earth and are partially reflected, refracted, or diffracted when encountering boundaries between layers with different elastic properties. After that, data processing steps include trace editing, deconvolution, velocity analysis, stacking, and migration to produce an interpretable image of subsurface reflectors [7].

The velocity of seismic waves is a critical parameter in seismic imaging, as it governs wave propagation and travel times. It is influenced primarily by the elastic properties of the rocks and varies according to lithology, mineral composition, porosity, fluid saturation, and compaction [9]. Although velocity typically increases with depth due to compaction and diagenesis, lateral variations also occur due to changes in facies and structural complexity. These variations are fundamental for interpreting stratigraphic patterns and geological features.

By analyzing attributes such as travel time, amplitude, frequency content, and phase variations, seismic data allow for the identification of complex subsurface features, including faults, folds, unconformities, and salt domes. These interpretations provide essential information on structural frameworks and sedimentary environments, which form the basis for the classification of facies and reservoir modeling [10].

B. Seismic Facies Interpretation

Seismic facies interpretation plays a fundamental role in seismic analysis workflows, as it allows geoscientists to extract geological meaning from patterns observed in seismic data. Seismic facies are defined as units which are different from adjacent units in its seismic characteristics, such as reflection amplitude, dominant reflection frequency, polarity, interval velocity, continuity and geometry. Seismic facies analysis uses different seismic parameters to get more than structural information [11].

Seismic facies analysis uses different seismic parameters to get more than structural information [11], integrating both qualitative and quantitative approaches to differentiate depositional environments and lithological variations based on seismic attributes. According to [7], variations in amplitude, continuity, and configuration of reflections provide critical clues about stratigraphic and structural settings. Moreover, other analysis as amplitude variation with offset (AVO), frequency decomposition, and seismic inversion contribute to a richer interpretation framework, revealing subtle geological features often missed in conventional structural interpretation [11].

With the aid of Convolutional Neural Networks, the interpretation of seismic facies becomes more efficient and consistent. CNNs are capable of learning and highlighting subtle variations in seismic texture that may be overlooked in manual analysis. Once the segmentation is complete, interpreters can focus on validating and contextualizing the results, linking them to geological models and guiding exploration decisions.

C. Dataset

The dataset used in the present work was the Parihaka 3D seismic cube provided by New Zealand Crown Minerals. Consisting of 590 inlines, 782 crosslines, and 1006 time samples recorded at a $3\,\mathrm{ms}$ interval, this survey is located in the Northwest of the Taranaki Basin, offshore New Zealand (Figure 1).

The seismic facies labeled in the aforementioned dataset are classified into six categories based on their geological characteristics, including reflectors, continuity, and amplitude.

According to [12], the classes present the following geological features:

1) **Basement:** Shows a low signal-to-noise echo amplitudes, with few internal reflectors.
2) **Slope Mudstone A:** Identified by low-amplitude continuous/semi-continuous internal reflectors.
3) **Mass Transport Deposit:** Displays a mix of chaotic facies and low amplitude parallel reflectors.
4) **Slope Mudstone B:** Exhibits high-amplitude parallel reflectors.
5) **Slope Valley:** Characterized by high-amplitude incised channels and valleys.
6) **Submarine Canyon System:** Internal fill is low-amplitude mix of parallel inclined surfaces and chaotic disrupted reflectors.

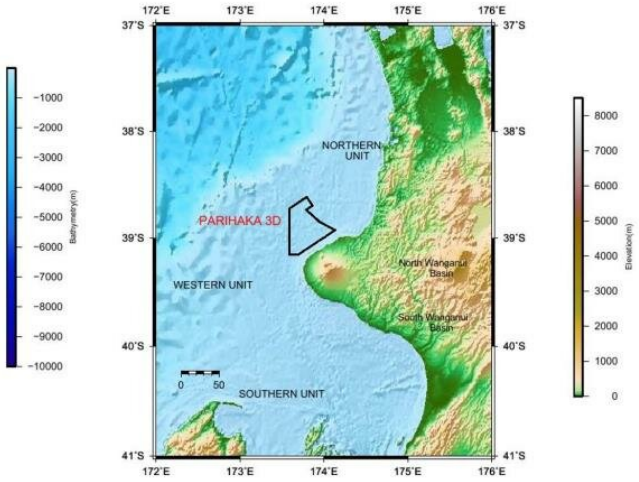


Fig. 1: Geological map of Taranaki basin with the Parihaka 3D seismic block marked by black polygon (Source: [13]).

III. Methodology

A. Convolutional Neural Networks

Convolutional Neural Networks are a specific case of neural networks, inspired by the visual neural cortex structure [14], [15], and are very effective to deal with grid-like data such as 2D images [16], [17]. These networks differ from the Multi-Layer Perceptron by applying the convolution operator to extract relevant patterns from the input [18]. To achieve this, different filters are convolved with the input to extract multiple feature maps. In a classification task, a combination of multiple feature maps is used to predict the correct label.

B. U-Net Architecture

In this research, a U-Net architecture was employed [19]. This technique is based on an encoder-decoder structure. Initially, the network extracts key characteristics through a sequence of successive contracting layers, where each step transforms the input into a lower-dimensional representation. Subsequently, the U-Net upsamples these features and concatenates them with the corresponding higher-resolution feature maps via skip connections. This step provides a spatial context to the upsampled features during the decoding process. The effective mapping of different levels of information makes this method well-suited and efficient for image segmentation tasks.

The current research scope is to perform seismic facies segmentation, assigning each pixel to a specific class. A seismic section can be represented as a 2D image, where

the reflector geometry, amplitude, and direction constitute the patterns to be extracted. For this task, a neural network must be able to not only extract patterns but also link these patterns to their spatial region. The U-Net architecture is effective here due to its skip connections, which map the extracted patterns to the spatial region in the image. Therefore, U-Net represents a robust technique for capturing the key features of each seismic facies and enabling the correct classification.

The U-Net applied in this work, illustrated in Figure 2, is simpler than the version developed by Ronneberger [19]. Our version has four levels of contraction and expansion, starting with 32 filters, then increasing to 64, 128, and reaching 256 filters at the bottleneck. The kernel size is 3x3 for convolutions and 2x2 for max-pooling, used for downsampling. For the activation function, a Leaky ReLU [20] was used because the negative values of the seismic traces contain relevant information.

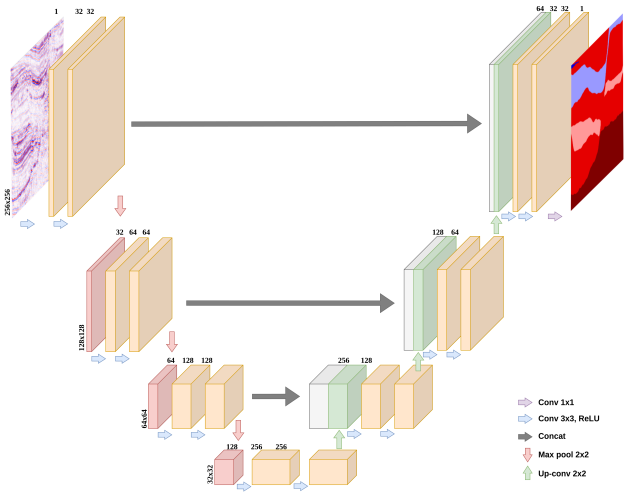


Fig. 2: The U-Net architecture employed for seismic facies segmentation, featuring four levels of contraction and expansion.

C. Experimental Workflow

To take advantage of the 3D seismic information, the workflow methodology processed the data along the inline and crossline orientations. The initial step was normalizing the amplitude data. Subsequently, both the normalized amplitude cube and its corresponding labels were split into training and testing datasets. The patches were extracted from these datasets to serve as input for the neural network. Thereafter, two neural networks were trained with the same settings for each orientation, and their respective prediction cubes were generated. In the last step, these two cubes were merged to create the final segmentation result. The workflow is illustrated in Figure 3.

1) Data Preprocessing: Neural Networks tend to be sensitive to the scale of input values, which can affect convergence and stability. Thus, the seismic data was normalized, using standardization, but only using the standard deviation, assuming that the seismic data is already centered around zero [7].

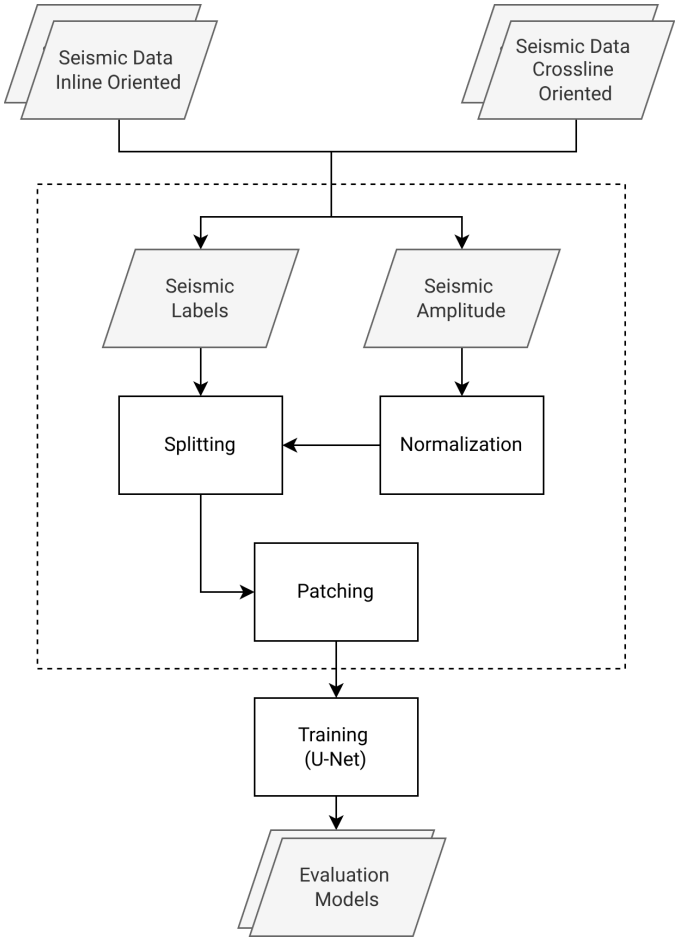


Fig. 3: Workflow diagram illustrating the main steps: normalization, splitting, patching, U-Net training, and evaluation.

2) Dataset Splitting: The seismic labeling step is typically a highly time-consuming and manual process. Usually, seismic interpreters analyze a part of the data in a process called picking. This method involves selecting a seismic section, labeling it, and then skipping several others before conducting a new analysis. The data in between is often interpolated. To emulate the scenario of a few labeled data points, we propose using only 10% of the seismic sections for training and the remaining 90% for testing. This method reproduces the interpreting process and is more computationally efficient.

3) Patching: After splitting the data, a step of patching is applied to extract a smaller subsection of the input. The original image is 400x500 pixels, whereas each patch is 256x256 pixels. The objective is to increase the amount of data and make the filters more precise by processing a smaller chunk of data each time. Also, the computational cost decreased because each training iteration involves processing a smaller input, which reduces memory requirements [21].

4) Model Training: The neural network was implemented using the Keras library [22] and trained for 150 epochs using the Adam optimizer [23]. Due to the high computational cost, a batch size of 1 was adopted. The loss function used was

Categorical Crossentropy.

5) Cubes Merging: Considering the 'picking' workflow for seismic interpretation, geophysical interpreters typically analyze data along both inline and crossline directions when evaluating an area. Therefore, we trained two separate U-Net models: one for processing inline-oriented sections and another for crossline-oriented sections. After generating prediction volumes from both models, these were merged to produce a single output cube. To achieve this, the cubes were first transposed to a common orientation and shape. Then, the last dimension of each model's output, which corresponds to class probabilities, was concatenated. The final class, among all possibilities for each voxel, was determined by selecting the maximum probability found across the corresponding predictions from the two models. This merging strategy enables the extraction of the most confident prediction from the network, thereby enhancing overall accuracy.

D. Evaluation Metrics

1) Mean Intersection over Union (Mean IoU): The performance metric employed in work was the Mean Intersection over Union (Mean IoU), also known as the Jaccard index (Equation 1). This metric displays the overlap between predicted and ground-truth segmentation for each class. For each class, the intersection (true positive) and union (all pixels in the ground truth and predicted area) are calculated [24]. The IoU for each class is computed as the ratio of the intersection and the union. A value of 1 indicates perfect agreement, while 0 represents no overlap. The Mean IoU is the average of the IoU from all classes (Equation 2).

The dataset is imbalanced, with few samples of classes *Submarine Canyon System* and *Basement*. Consequently, metrics such as overall accuracy are unsuitable for evaluating model performance, as they are insensitive to the performance of the minority classes. In contrast, the Mean IoU provides a better approach to evaluate by weighting classes equally, thus offering a more informative assessment for the geological application considered in this study.

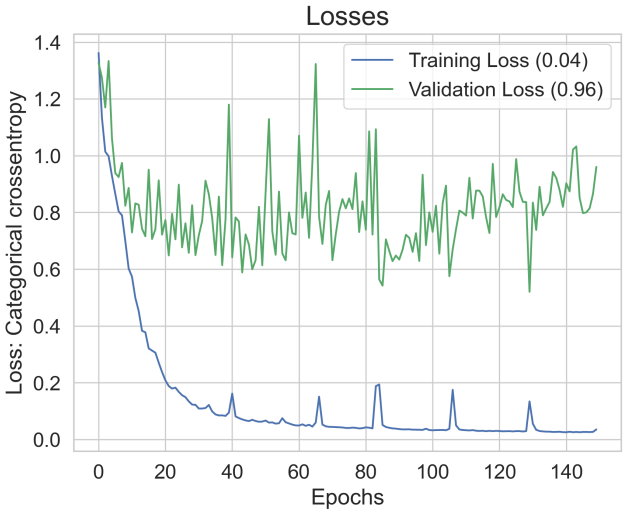
$$\mathbf{IoU}_i = \frac{\mathrm{TP}_i}{\mathrm{TP}_i + \mathrm{FP}_i + \mathrm{FN}_i} = \frac{\text{Intersection}}{\text{Union}} \tag{1}$$

$$\textbf{Mean IoU}_i = \frac{1}{n}\sum_{i=0}^{n} \mathrm{IoU}_i \tag{2}$$

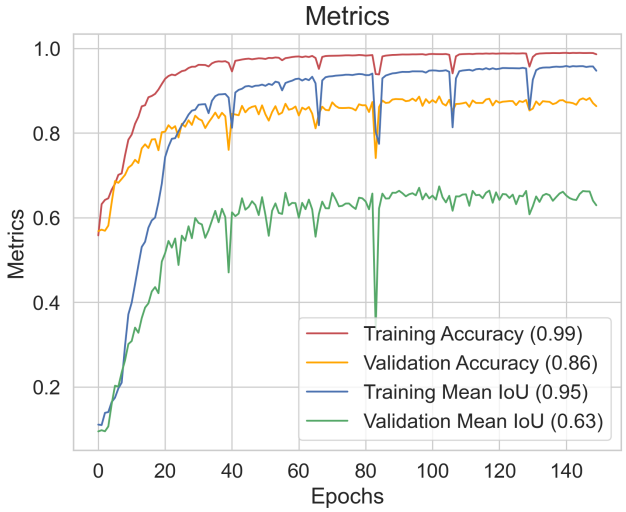
IV. Results

A. Inline Training

According to the approach of splitting the test and training sets based on the picking principle, only 10% of the data is used for training the model. Out of this portion, 20% is used for validation. Because of that, it is natural to see that sometimes the model does not adapt well to the validation data, leading to spikes in the loss graph and a lack of convergence, as observed in Figure 4a. This happens due to the low statistical representation and variety of scenarios in the training set. Still, the results are acceptable since the structures are, to some extent, similar. This is reflected in the values observed in the Mean IoU metric, shown in Figure 4b.



(a) Loss (Categorical Crossentropy) vs. Epochs.

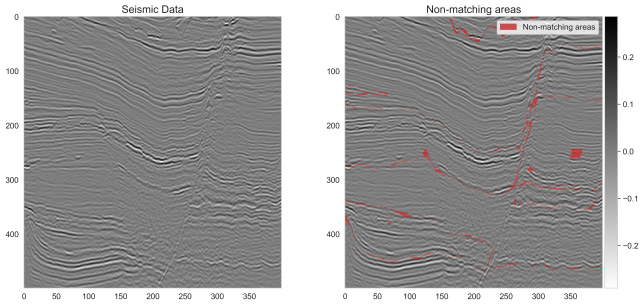


(b) Metrics (Accuracy and Mean IoU) vs. Epochs.

Fig. 4: Training history for the U-Net model on inline-oriented data: (a) Loss curves and (b) Performance metrics.

Some fragments of the *Submarine Canyon* class can be seen in the *Slope Mudstone B* area, and *Slope Mudstone B* is also identified in a region labeled as *Slope Mudstone A* (Figure 5d). An important point to highlight is that areas with very distinct characteristics, such as the *Basement*, were not displaced. Additionally, the connection within the *Slope Mudstone B* region represents a fault, meaning a fracture in the area. From a geological perspective, this region should not have a clearly defined class label since the surrounding content is the same.

This is reflected in the model's results. In the non-matching area (Figure 5b), the model misrepresents the thickness of that area, which is actually consistent with the geological structure.



(a) Normalized Seismic Section (b) Non-matching Areas

(c) Ground Truth Labels (d) Predicted Labels

Fig. 5: Example of seismic segmentation results inline-oriented. Comparison between (a) input seismic, (b) highlighting the non-matching regions, (c) true labels and (d) model prediction.

B. Cubes Merging

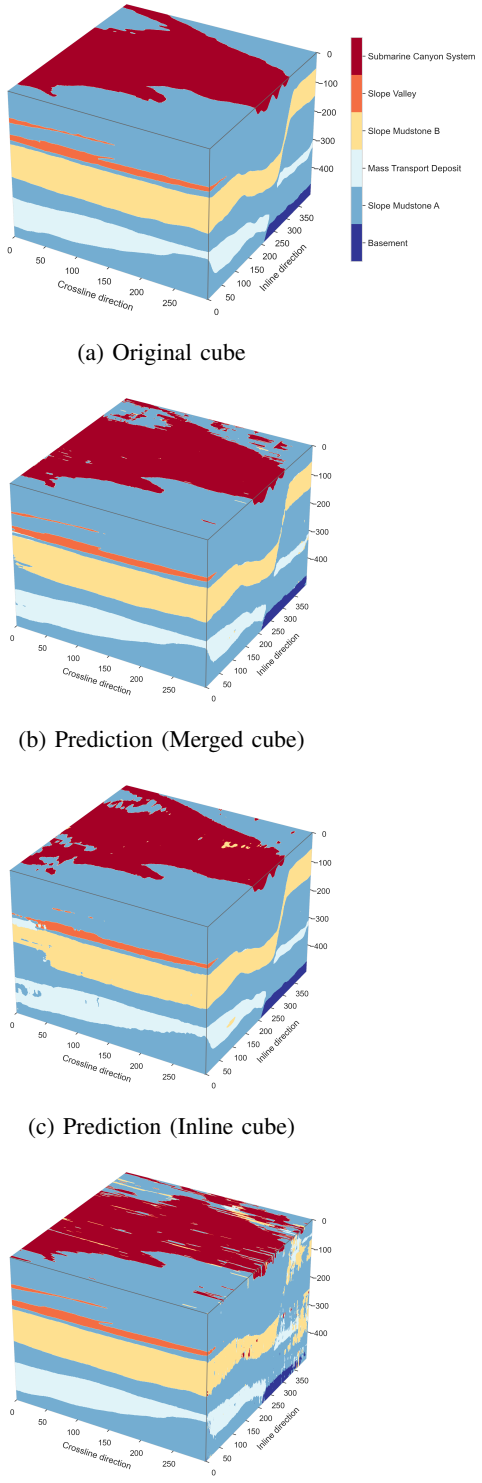
Table 1 presents the comparative metrics between the 3D ground truth cube and those generated by the model. The accuracy difference between the cube trained along the inline direction and the merged cube is relatively small, only 0.2%. However, this difference becomes more relevant when analyzing the mean IoU, which was increased by 2%.

TABLE I: Accuracy and Mean IoU for each cube

Cube	Accuracy	Mean IoU
Crossline	0.916	0.796
Inline	0.954	0.852
Merged	0.956	0.870

The dataset used is very imbalanced. Therefore, this improvement in mean IoU suggests that merging the cubes enabled the model to identify certain underrepresented classes more effectively. These typically correspond to special events, indicating that the proposed approach is effective.

It is also important to highlight that the metrics revealed a slight difference in performance depending on the training direction. Specifically, training along the inline direction consistently yielded better results when extrapolated to the full volume compared to training along the crossline direction. This difference is likely related to the fact that inlines generally present higher spatial resolution and better continuity of seismic reflectors than crosslines, largely due to acquisition geometry, where inlines correspond to the primary shot direction.



(a) Original cube

(b) Prediction (Merged cube)

(c) Prediction (Inline cube)

(d) Prediction (Crossline cube)

Fig. 6: Comparison of results: (a) The ground truth cube, Prediction from the U-Net trained on (c) inline orientation, (d) on crossline orientation, and (b) the final merged result.

As a result, models trained along crosslines tend to generalize less effectively to the full volume, particularly struggling with complex or subtle geological features. The performance gain observed in the merged cube demonstrates that combining information from different directions helps compensate for these limitations, leading to improved generalization and better identification of underrepresented classes.

V. Final Remarks

The application of the Deep Learning model has been successfully utilized in seismic interpretation, yielding notable results. Since seismic data can be treated as images, Convolutional Neural Networks have also been employed with remarkable success. Consequently, convolutional models such as U-Net represent a powerful tool for seismic facies segmentation.

This work demonstrates that the proposed train-test split approach, which emulates the picking process by utilizing a limited amount of labeled data — a common scenario in real-world analysis — proves to be effective in identifying the key characteristics of seismic data.

Future work could use other patching strategies, such as rectangular patches that cover the full height of the seismic section. This approach might provide an in-depth context to the neural network. Furthermore, the use of different loss functions, such as Dice Loss or Categorical Focal Loss, could be employed as a way to address the challenge of imbalanced classes.

In conclusion, the proposed methodology demonstrated its effectiveness in segmenting seismic data, showing an alternative to interpolation techniques between labeled sections. This methodology demonstrates the potential for efficient application in a broader geological settings context.

Acknowledgment

The authors thank the Instituto Nacional de Ciencia e Tecnologia de Geofisica do Petroleo (INCT-GP) for supporting this research. Alexsandro Cerqueira would like to thank the National Council for Scientific and Technological Development (CNPQ) for financing the project nº 409718/2022-0. We would also like to acknowledge the support of the Postgraduate Program in Geophysics (PPGEOF) of the Federal University of Bahia (UFBA).

References

[1] A. U. Waldeland, A. C. Jensen, L.-J. Gelius, and A. H. S. Solberg, "Convolutional neural networks for automated seismic interpretation," *The Leading Edge*, vol. 37, no. 7, pp. 529–537, 2018.

[2] X. Wu, L. Liang, Y. Shi, and S. Fomel, "Faultseg3d: using synthetic data sets to train an end-to-end convolutional neural network for 3d seismic fault segmentation," *Geophysics*, vol. 84, no. 3, pp. IM35–IM45, 2019.

[3] Y. Shi, X. Wu, and S. Fomel, "Saltseg: automatic 3d salt segmentation using a deep convolutional neural network," *Interpretation*, vol. 7, no. 3, pp. SE113–SE122, 2019.

[4] H. Di, D. Gao, R. Zhang, and J. Zhang, "A machine-learning workflow for seismic facies classification and interpretation," *Interpretation*, vol. 6, no. 3, pp. SE147–SE158, 2018.

[5] M. A. Islam, "Using deep learning based methods to classify salt bodies in seismic images," *Journal of Applied Geophysics*, vol. 178, p. 104054, 2020.

[6] W. M. Telford, L. P. Geldart, and R. E. Sheriff, *Applied geophysics*. Cambridge University Press, 1990.

[7] O. Yilmaz, *Seismic data analysis: Processing, inversion, and interpretation of seismic data*, 1st ed. Society of Exploration Geophysicists, 2001.

[8] P. Kearey, M. Brooks, and I. Hill, *An introduction to geophysical exploration*, 3rd ed. John Wiley & Sons, 2013.

[9] A. Schuck and G. Lange, *Environmental Geology*. Springer, 2007.

[10] S. Chopra and K. J. Marfurt, *Seismic attributes for prospect identification and reservoir characterization*. Society of Exploration Geophysicists, 2007.

[11] M. M. Roksandíc, "Seismic facies analysis concepts *," *Geophysical Prospecting*, vol. 26, no. 2, pp. 383–398, 1978.

[12] S. Stamm, "2020 SEG Annual Meeting Machine Learning Interpretation Workshop," Basecamp, Sep. 2020, last updated: Sep 3, 2020. [Online]. Available: https://public.3.basecamp.com/p/JyT276MM7krjYrMoLqLQ6xST

[13] P. Chinmoy Kumar, "Application of geometric attributes for interpreting faults from seismic data: An example from taranaki basin, new zealand," in *SEG Technical Program Expanded Abstracts 2016*. Society of Exploration Geophysicists, Sep. 2016.

[14] D. H. Hubel and T. N. Wiesel, "Receptive fields and functional architecture of monkey striate cortex," *J. Physiol.*, vol. 195, no. 1, pp. 215–243, Mar. 1968.

[15] K. Fukushima, "Neocognitron: a self organizing neural network model for a mechanism of pattern recognition unaffected by shift in position," *Biol. Cybern.*, vol. 36, no. 4, pp. 193–202, 1980.

[16] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.

[17] A. Krizhevsky, I. Sutskever, and G. Hinton, "Imagenet classification with deep convolutional neural networks," *Neural Information Processing Systems*, vol. 25, 01 2012.

[18] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016, http://www.deeplearningbook.org.

[19] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," 2015. [Online]. Available: https://arxiv.org/abs/1505.04597

[20] A. L. Maas, "Rectifier nonlinearities improve neural network acoustic models," 2013. [Online]. Available: https://api.semanticscholar.org/CorpusID:16489696

[21] L. Hou, D. Samaras, T. M. Kurc, Y. Gao, J. E. Davis, and J. H. Saltz, "Patch-based convolutional neural network for whole slide tissue image classification," 2016. [Online]. Available: https://arxiv.org/abs/1504.07947

[22] F. Chollet *et al.*, "Keras," https://keras.io, 2015.

[23] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2017. [Online]. Available: https://arxiv.org/abs/1412.6980

[24] M. Everingham, L. Van Gool, C. K. I. Williams, J. Winn, and A. Zisserman, "The pascal visual object classes (VOC) challenge," *Int. J. Comput. Vis.*, vol. 88, no. 2, pp. 303–338, Jun. 2010.