



Seismic data pattern recognition based on analogous reservoirs via Deep Transfer Learning

Pedro Silvany (Petrobras), Marcos Machado (Petrobras), Théo de Tarzo (Petrobras), Leonardo Oliveira (Petrobras), Marcilio Mattos (SISMO)

Copyright 2021, SBGF - Sociedade Brasileira de Geofísica

This paper was prepared for presentation during the 17th International Congress of the Brazilian Geophysical Society held in Rio de Janeiro, Brazil, 16-19 August 2021.

Contents of this paper were reviewed by the Technical Committee of the 17th International Congress of the Brazilian Geophysical Society and do not necessarily represent any position of the SBGF, its officers or members. Electronic reproduction or storage of any part of this paper for commercial purposes without the written consent of the Brazilian Geophysical Society is prohibited.

Abstract

The shortage of static and dynamic data in areas newly acquired or in the early stages of production and production development makes extremely difficult to understand the spatial distribution of the different scales of geological heterogeneities. This often impacts in obtaining realistic flow scenarios. To address the data scarcity for reservoir modelling and characterization, a commonly used approach is the identification and study of analogous reservoirs. This can bring many different data which permit create important quantitative criteria to modelling a production zone. Due to the large volume and types of information, the task of recognizing pattern in these data mountain can be overwhelming. This is particularly true if pre-stack seismic data are considered. To tackle this problem, different transfer learning strategies based on artificial intelligence can be used. Recently, it has been proposed to combine deep autoencoders with clustering algorithms to extract seismic facies from pre-stack seismic data. In this paper, we extend the methodology to consider the possibility of using two different input data: the analogous and the target. These two data can be used in the training of a deep convolutional autoencoder or in the clustering algorithm. The proposed transfer learning strategies are applied to a Brazilian siliciclastic reservoir, showing an improvement in the recognition of the depositional architectural elements and understanding of rock and fluid quality based on other analogous reservoirs.

Introduction

The quick and at low cost construction of reliable geological models of an oil field is a challenge for newly acquired areas or areas in the initial phase of production and development of production – DP&P. This is one of the most important steps for understanding the hydrodynamic behavior of the reservoir and one of the main sources of uncertainties. However, the shortage of static and dynamic data in most of these areas makes it extremely difficult to understand the spatial distribution of the different scales of geological heterogeneities and, consequently, influence in obtaining realistic flow scenarios.

To address the data scarcity for reservoir modelling and characterization, a commonly used approach is the

identification and study of analogous reservoirs. The understanding of the occurrence of geological heterogeneities, from the study of analogous areas, and the patterns variation in reservoir scale, can bring important quantitative criteria for the parameterization and spatial distribution of the main geological features within a production zone. Generally, this approach enhances the understand of how certain properties (eg. porosity, permeabilities, mechanical properties, etc.) are related, integrating different data scales, to improve the assessment of probabilistic uncertainty scenarios.

Petrobras S.A. has been acquiring and producing literally a mountain of data in the last 70 years of Exploration and Production campaigns - E&P sector. There are data from different scales, supports and from different physical measures, such as log data (gamma rays, neutron density, sonic dipole, resistivities, acoustic and resistive images), test data (RFTs, mini-TIs, DSTs, others.), seismic data (4D seismic, VSP, others.), cores, analogues outcrops 3D models, geological models, simulation models, others. Digital solutions and automated frameworks both integrated to a large database of reservoirs and outcrops analogous to the modeled reservoirs is one of the main technological challenges in the conception of new projects. This will promote break down the silos and greater agility in build and condition new models.

Due to the large volume and types of information, the task of recognizing pattern in these data mountain can be overwhelming. To tackle this problem, different transfer learning strategies based on artificial intelligence are being proposed. Transfer learning is a neural network approach in which knowledge obtained with a problem is transferred to a different but related problem (Yosinski, 2014). In deep learning the technique can be used to learn the features of a huge amount of data avoiding overfitting. In the first step, the model is initialized and trained with one set of images. Then, the network with pre-trained filters continues to be trained with a second set of images, for which it is desired to recognize the main patterns. The network layers have different levels of specialization. The key point is the generalization ability of the convolutional network. But as found by (Yosinski, 2014), initializing with transferred features can improve generalization performance even after substantial fine-tuning on a new task.

Reservoir transfer learning studies must consider that methodologies to build a reservoir model integrate several information sources. The information from the wells, although more precise, is punctual and does not reflect the behavior of the field as a whole. In this way, the spatial identification of architectural and structural elements is dependent on the analysis and good representation of the

spatial distribution of the main characteristics of the seismic signal.

In the stacked seismic data, the signal amplitude is a mean of the contributions of the amplitudes obtained for different source-receiver offsets. This vector information does not allow to extract all the richness of details that exists in the data, different from the matrix of traces of the pre-stacked seismic data, which carries the information of the reflection coefficients as a function of the angles they were illuminated. Quantitative seismic interpretation based on pre-stack seismic data, such as amplitude variation with offset (AVO) analysis, allowing the extraction of geological features with greater accuracy and resolution, as lithology and fluid in the porous media (Simm and Bacon, 2014).

(Qian et al., 2018) proposed to combine deep autoencoders with clustering algorithms to extract seismic facies. Deep Learning appears as a non-linear statistical approach for patterns recognition of the elastic information within the common-depth-point (CDP) gathers. The Deep Convolutional Autoencoder (DCAE) is used to learn efficient data encodings, reducing the data dimensionality. The latent features generated by the DCAE encoder are used as input to the K-means algorithm.

Recently, (Silvany et al., 2020, Silvany et al., 2021) showed how to apply the (Qian et al., 2018) methodology to the multi-azimuthal pre-stack data and frequency-gathers extracted from the post-stack data, using different input channels to accommodate each azimuth.

Here, we extend the methodology to consider the possibility of using transfer learning. In this case, the algorithm receives two different input data: the target and the analogous data. We propose two strategies to deal with them. In the first one the analogous data is only used to pre-train the DCAE network. In the second strategy, the analogous data is used on both the DCAE training and the K-means algorithm.

In the next section we present the methodology to extract seismic facies using DCAE and K-means and the two proposed transfer learning strategies. Also, we use Student's t-distribution to estimate the probability for each facies (Silvany et al., 2021). After, we discuss the results of the methodology of learning seismic features in analogs reservoirs and predict different seismic facies with fine-tuned features or frozen features approaches. The results are interpreted based on the lithology and fluid characteristics of analogues fields in a real siliciclastic play in Sergipe-Alagoas Basin, eastern Brazilian margin.

Method

The input data are the region of interest (for example, the window around a horizon), the pre-stack CDP gathers, and the frequency-gathers. For each point of the region of interest, one time-offset and one time-frequency panel are extracted from the input gathers. The panels are parametrized by the number of samples, the number of offsets and the number of frequencies to be considered. Also, as we are able to consider multi-azimuthal data, we can have one time-offset panel and one time-frequency panel for each azimuth (Silvany et al., 2021). The goal is the identification of a certain number k of facies present in the data.

In the context of the present work, the space \mathbf{X} has the dimension of all possible gray images. However, the actual gathers images used are only a small subset of \mathbf{X} . Work with a high dimension space involves problems known as "curse of dimensionality" (Bellman, 1961). So, assuming that the real gathers images form a manifold embedded in \mathbf{X} , we will first transform the data with a nonlinear mapping $f_{\theta} : \mathbf{X} \rightarrow \mathbf{Z}$, where θ are learnable parameters and \mathbf{Z} is the latent feature space. The dimensionality of \mathbf{Z} is smaller than the \mathbf{X} . The set of transformed points $\{\mathbf{z}_i\}$ will be the input to the clustering algorithm. Thus, the method has two steps: (1) the f_{θ} building and (2) the clustering method application.

To parametrize f_{θ} , it will be used a Deep Convolutional Autoencoder (DCAE). An autoencoder (LeCun, 1987) is a type of artificial neural network used to learn efficient data encodings. DCAE is composed by two subnets: the encoder, which generates the code (latent feature space) for each input \mathbf{X}_i , and a decoder, which receives the code and makes the reconstruction $\hat{\mathbf{X}}_i$, as similar as possible to the original input. The training of the DCAE is done minimizing the reconstruction error. As it does not depend on labelled data, it is an unsupervised learning method.

The encoder is composed by a sequence of convolutional and max pooling layers. A convolutional layer (LeCun, 1989) has a set of filters, all of them with the same size. As the filter size is less than the input size, the weights that define the filter are applied over the input as a convolution. This permits to identify patterns in the input image in a way that is invariant with translation. The max pooling layer down samples the input. Typically, it runs a 2x2 mask over the input, taking the maximum value and shifting with stride of 2 along both directions, reducing the output size. This implementation of an encoder can be seen as a pyramid filter extracting features with different levels of abstraction (or scale).

The decoder, on the other hand, implements an inverted pyramid: it is composed by a sequence of upsampling and convolution layers. Upsampling layer typically doubles the size of the image, assigning to the output pixel the nearest pixel of the input.

DCAE is implemented using the denoising autoencoders (Vincent et al, 2008) strategy: it learns to approximate the original input by training on the input vectors with noises. The DCAE is designed to reconstruct the original data from the corrupted version of the original images, the process of which forces the hidden layer to discover more robust features and prevents overfitting noises. Figure 1 shows schematically the structure of DCAE.

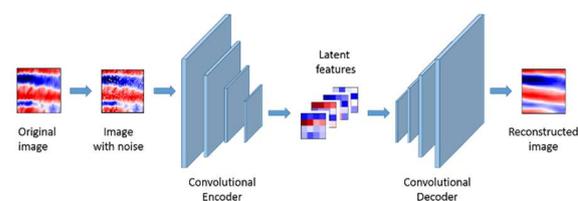


Figure 1 - The structure of DCAE.

Once trained the DCAE network (the first step of the method), we use the Encoder to make the transformation

$f_{\theta} : \mathbf{X} \rightarrow \mathbf{Z}$ for all time-offsets and time-frequency panels. Observe that the input for the Encoder are the same data used in the DCAE training, but now without noise. The code set is submitted to a clustering algorithm, obtaining a cluster identification (a number) for each input element.

In order to evaluate the probability of each point to belong to each cluster, we adopt the Student t-distribution, following the classical work of (van der Maaten & Hinton, 2008). In equation 6, q_{ij} is the probability of the point i to belong to the cluster j (z_i is the code for the i -th point and μ_j is the centroid of the j -th cluster).

$$q_{ij} = \frac{(1 + \|z_i - \mu_j\|^2)^{-1}}{\sum_j (1 + \|z_i - \mu_j\|^2)^{-1}} \quad (6)$$

Figure 2 describes visually this second step.

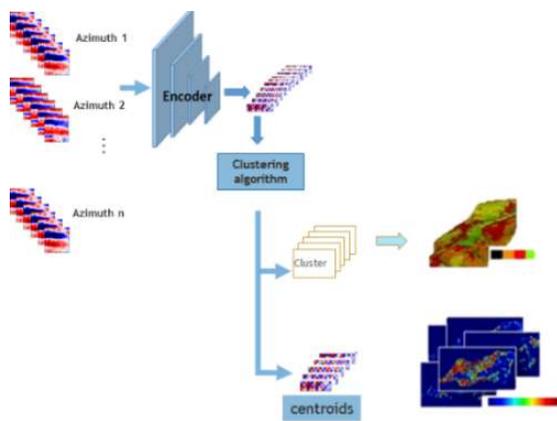


Figure 2 - The second step of the method.

In this paper, we extend the methodology to consider the possibility of using transfer learning. In this case, the algorithm receives two different input data: the target and the analogous data. Here we propose two strategies.

In the first one the analogous data is used to pre-train the DCAE; then, the weights and bias obtained are used for initializing the network which is, now, trained with the target data. The clustering algorithm is applied to the target feature vectors (extracted from the DCAE encoder). With this initialization, we hope to induce the encoder to retain some features learned from the analogous data (with this purpose, in the final training, we use a learning rate value lower than the that used for pre-training). We call this as fine-tuning strategy.

In the second strategy, the DCAE is trained with the analogous data. The obtained encoder is then used to extract two sets of feature vectors: one from the analogous data and one from the target data. The K-means is computed over the analogous feature vectors. To obtain the clusters for the target data its feature vectors are classified according to the distance from the centroids obtained previously. Observe that, in this case, there is a direct correspondence between the target facies and the analogous facies. However, with this autoencoder, there is

a risk of deteriorating the target panels encoding and decoding. We call this strategy as frozen strategy.

Siliciclastic Reservoir and Analogues Case Study

The geological province is located in the deep-water region of the Sergipe-Alagoas basin, northeast continental Brazilian margin. It is classified as a divergent margin tectonic environment. The area has hydrocarbon reserves related to siliciclastic turbidite deposits from the Calumbi Formation, that corresponds to a predominantly pelitic sequence, with turbidites embedded in siltstones, shales and marls. The conceptual depositional model of these reservoirs is based on changes in the depositional energy gradient in the physiographic paleoenvironment of slope. The major trapping mechanism involved in the region is stratigraphic, with the top and lateral seals provided by overlying shale. Characterizing these seismic facies is not a straightforward process, involving uncertainties that shall impact, for instance, in non-optimal well locations of the proposed drainage plan. The Figure 3(A), (B) and (C) illustrates a VPVS anomaly map, seismic amplitude dip section and VPVS dip section crossing the wells 1 and 2 of an analogue reservoir.

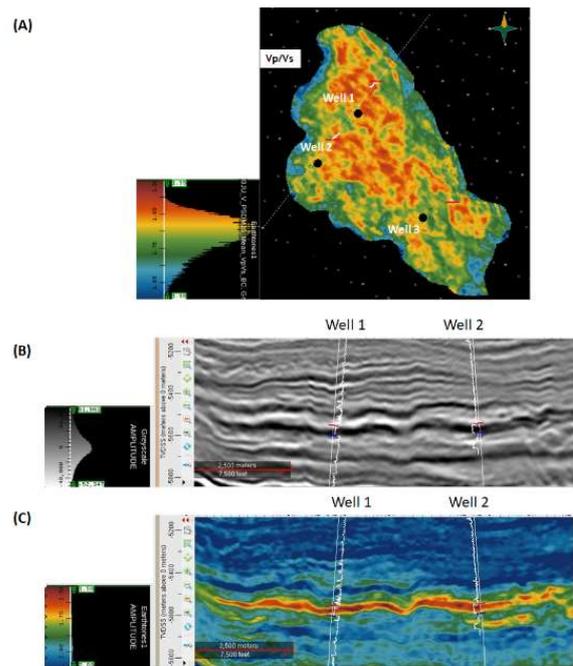


Figure 3 (a) VPVS anomaly map of an analogue reservoir; (b) seismic amplitude dip section crossing the wells 1 and 2; (c) VPVS attribute dip section crossing the wells 1 and 2.

Petrophysics and geology analysis in the existing wells have shown that these reservoirs are mostly composed of conglomerate, sandstone, shaly sand and laminate facies, based on core and well-log information. The turbidite reservoir has low to medium porosity saturated by oil. Although this petrophysical property suffers reduction due to diagenetic processes, pore-filling clay deposition, and

reservoir compaction. The compressional-to-shear velocity ratio (V_p/V_s) generally increases with the clay content in a quartz-rich turbidities sedimentary rock. So, the elastic property can be used as an indicator of reservoir quality (Figure 4). Due to these effects, we can observe AVO answer in the commons-depth-points gathers on the clastic reservoir with dominance of class III anomalies.

The project database is composed by a Broadband narrow-azimuth seismic survey with sail-line head of N119o. The seismic processing is a PSDM *Kirchhoff*. We use the pre-stack CDP gathers N119o to extract the codes which are submitted to a clustering algorithm.

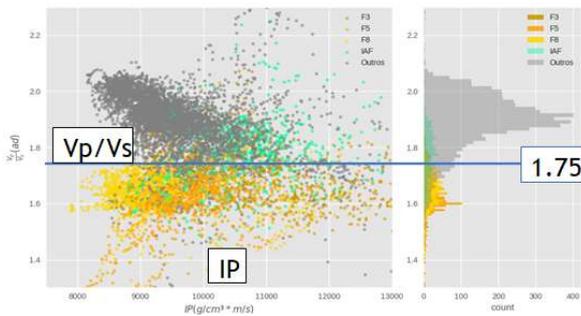


Figure 4 Crossplot of IP vs. $VPVS$ colored by depositional facies.

The reference horizons to select the data window were the top of the target and analogous reservoirs. The chosen interval comprised 100 meters below it. The DCAE was trained with images from the analogue reservoirs, each one with 20 samples in depth and 60 offsets. The feature space generated was a vector space with 32 components (with a dimensionality reduction from 1200 to 32). In the first strategy (fine-tuning), the weights and bias obtained are used for initializing the network which is, now, trained with the target data. The clustering algorithm is applied to the target feature vectors (extracted from the DCAE encoder). In the second strategy (frozen), the encoder obtained from the analogous is used to extract two sets of feature vectors: one from the analogous data and one from the target data. The K-means is computed over the target data, its feature vectors are classified according to the distance from the previously obtained centroids.

Figure 5 shows how well works the DCAE in the different situations to which it is submitted. Figure 5(B) shows ten analogous panels encoded and decoded by a DCAE trained with analogous data. Figure 5(C) shows ten target panels encoded and decoded by a DCAE trained with target data. Figure 5(D) shows ten target panels encoded and decoded by a fine-tuning DCAE. Figure 5(E) shows ten target panels encoded and decoded by a frozen DCAE. Observe that the reconstructed images are a smoothed version of the input images. The main features are preserved in all the situations, even in the frozen DCAE.

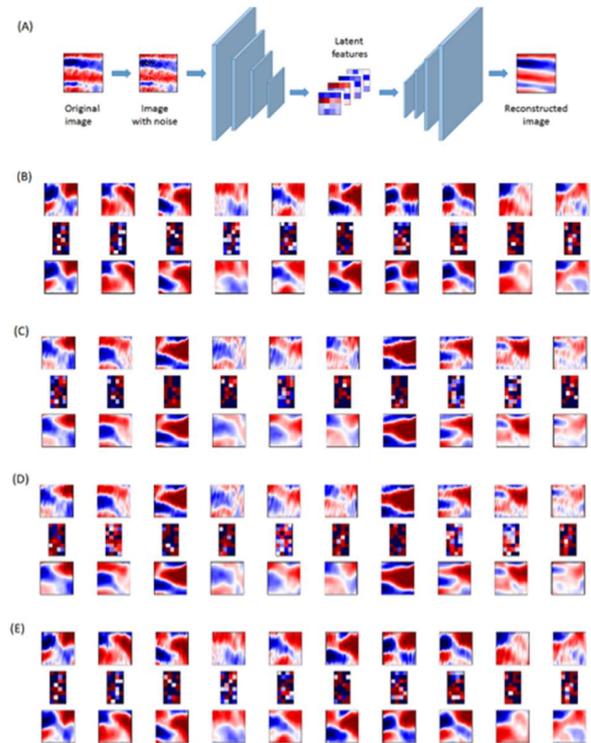


Figure 5 (A) The structure of DCAE; (B) ten analogous panels encoded and decoded by a DCAE trained with analogous data; (C) ten target panels encoded and decoded by a DCAE trained with target data; (D) ten target panels encoded and decoded by a fine-tuning DCAE; (E) ten target panels encoded and decoded by a frozen DCAE.

The Figure 6 (A) and (B) show the obtained seismic facies with DCAE applied in a single azimuth pre-stack CDP gathers (N119o), and the maximum probabilities maps associated for a specific analogue reservoir, respectively. The deposition of the analogue field occurred by turbidite flows with NW-SE direction. These flows filled minibasins in the slope environment, building systems of lobes and channels amalgamated. In the early stages of deposition, the region was filled by lower energy turbidite flows, with deposition in local depressions that compose thin sandstone intervals. In the final stages of deposition, with the same depositional direction, occurred turbidite flows of higher energy that compose thicker deposits (~80 meters). The lower energy lithofacies are described as medium to fine sandstone and shaly sand. These are associated with the red color seismic facies area. The higher energy lithofacies are described as conglomerate lithofacies and are associated with the light and dark green seismic facies areas. The interval crossed by Well1 shows these characteristics, where conglomerates lithofacies were described. The E and W edges of the reservoir are described as erosive features of late depositional events, associated with the orange seismic facies areas.

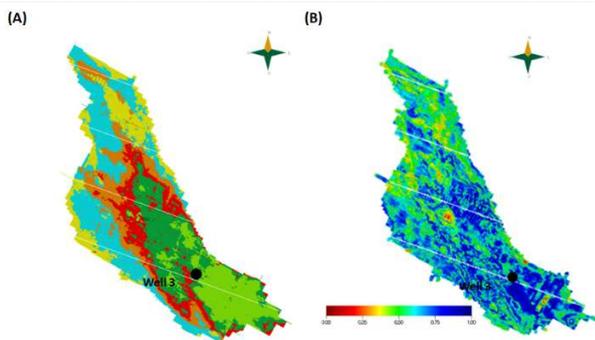


Figure 6 (a) Six facies map obtained by DCAE+K-Means in a single pre-stack CDP gathers (N119o) for a specific analogue reservoir; (b) Maximum probabilities map associated with the seismic facies.

The Figure 7 (A) shows the structural map and logs from Well 3. The Figure 7 (B) and (C) show the six facies and probability maps obtained by DCAE+K-means using only the target data. The Figure 7 (D) and (E) present the six facies and probability maps obtained using fine-tuning strategy. Finally, the Figure 7 (F) and (G) show the six facies and probability maps using frozen strategy.

The six colors (target facies) in Figure 7 (F) have a direct correspondence to the facies identified in Figure 6 for the analogous data. Naturally, due to the fact that both DCAE and K-means have been trained only with the analogous data, the probabilities shown in Figure 7 (G) have lower values than the ones shown in the other probability maps (Figure 7 (C) and (E)). But observe that, while the yellow and cyan facies occur with low probabilities, the dark green, the light green, orange and the red facies have higher probability values.

If we compare the fine-tuning results (Figure 7 (D)) with the frozen ones (Figure 7 (F)), we observe that the dark green, red and orange facies are almost equal in the two figures. On the other side, those three facies have a slightly different occurrence in Figure 7 (B) (obtained without transfer learning). This is more evident for the dark green facies. Those similarities between fine-tuning and frozen results and their differences for the results obtained with no transfer learning points to the effectiveness of the two methods.

The deposition of the target reservoir also occurred by turbidite flows. The interval composes thick deposits in the central area (approximately ~ 70 to 80 meters). The mainly lithofacies described are conglomeratic sandstones at the reservoir base with fining up pattern to the top. In the early stages of deposition, the higher energy flows were controlled by a high morphological structure in the E-SE area of the field. The resulting sedimentary intervals are described in Wells 2 and 3 as conglomeratic lithofacies. In the late stage of deposition, occurred channeled deposits with less sedimentary inflow. There is predominance of medium to fine sandstone lithofacies, as described in Well 1.

One possible interpretation for these maps (Figure 7 (B), (D) and (F)) associates light green and dark green seismic

facies with fill of a mini basins in a high energy phase of the depositional system, similar to the interpretation of analogous reservoirs. The seismic facie in red is related with channeled geometries corresponding to the environments of high-depositional slope. Subsequently, there is a late erosive event that isolates the area of Well 1 from the main portion of the field. This event is associated with the seismic facie in orange, similar the interpretation of the previous analogues reservoirs.

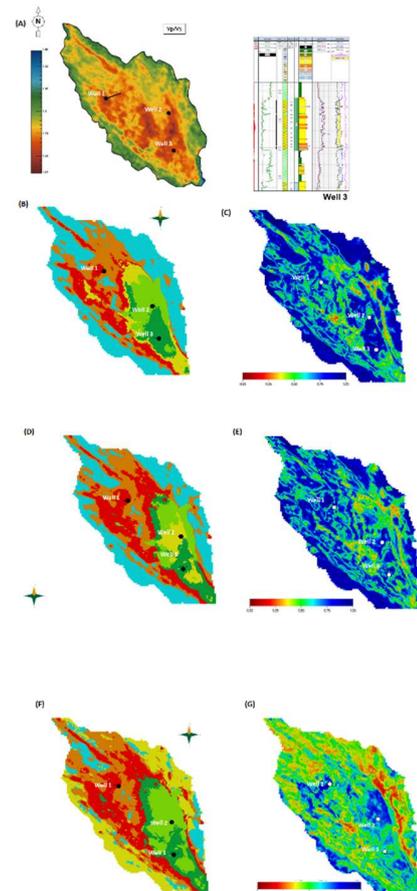


Figure 7 (A) Structural map into Calumbi Formation and logs from Well 3; (B) Six facies map obtained by DCAE+K-Means using only the target data; (C) Probability map associated with the first seismic facies map; (D) Six facies map obtained by DCAE+K-Means using fine-tuning strategy; (E) Probability map associated with the second seismic facies map; (F) Six facies map obtained by DCAE+K-Means using frozen strategy and (G) Probability map associated with the third seismic facies map.

The classification of seismic facies in the target reservoir obtained by transfer learning enhanced the understanding of architectural elements of deep-water deposits and associated lithofacies (based on analogous areas tested by wells). Both strategies are consistent and address the data scarcity for the reservoir modelling and characterization. The probability maps obtained with

different strategies can also support the uncertainty analysis of the seismic facies recognition for the geological modelling.

Conclusions

The DCAE allows extracting features from pre-stack seismic data with a high level of abstraction and in a non-linear way. Using those features as input to a clustering method it can be obtained seismic facies map using the richness of the pre-stack data. When compared with the classical approaches, the results show higher resolution in the recognition of the architectural elements of deep-water deposits and greater accuracy in the identification of zones with different depositional facies. Transfer learning enhanced the understanding of the tectonic-sedimentary patterns in the target reservoir based on analogous areas tested by wells. The strategies brought information of rock and fluid quality from other mature stages fields. The probability maps obtained with different strategies can also support the uncertainty analysis of the seismic facies recognition for the geological modelling.

Acknowledgments

We thank Petrobras for permission to publish this work.

References

- Aggarwal, Charu & Hinneburg, Alexander & Keim, Daniel. (2002). On the Surprising Behavior of Distance Metric in High-Dimensional Space. Database theory, ICDT 200, 8th International Conference, London, UK, January 4 - 6, 2001.
- Bellman, R. Adaptive Control Processes: A Guided Tour. Princeton University Press, Princeton, New Jersey, 1961.
- Cunha Filho, C.A., Ruthner, M.P., Gontijo, R.C., Statistical Seismic Facies Estimation from Pseudo Impedance Data, EAGE, 2013.
- Dumay, J., and F. Fournier, 1988, Multivariate statistical analyses applied to seismic facies recognition: Geophysics, 53, 1151–1159.
- Fournier, F., and J. F. Derain, 1995, A statistical methodology for deriving reservoir properties from seismic data: Geophysics, 60, 1437–1450.
- Johann, P.R.S., Castro D.D. and Barroso A.S., 2001, Reservoir geophysics: Seismic pattern recognition applied to ultra-deepwater oilfield in Campos basin, offshore Brazil: SPE LatinAmerica and Caribbean Petroleum Engineering Conference;
- Kohonen, T., 2001, Self-organizing maps: Springer.
- LeCun, Y. (1987). Modèles connexionistes de l'apprentissage. Ph.D. thesis, Université de Paris VI.
- LeCun, Y. (1989). Generalization and network design strategies. Technical Report CRG-TR-89-4, University of Toronto.
- Lloyd, Stuart P. "Least squares quantization in PCM." Information Theory, IEEE Transactions on 28.2 (1982)
- Marfurt, K. J., and R. L. Kirlin, 2001, Narrow-band spectral analysis and thin-bed tuning: Geophysics, 66, 1274–1283.
- Matos, M.C., Osório, P.L.M. and Johann, P.R.S. [2007] Unsupervised seismic facies analysis using wavelet transform and self-organizing maps. Geophysics, 72, 9-21.
- Silvany, P., Machado, M., Tarzo, T., Pre-stack seismic facies prediction via deep convolutional autoencoders: an application to a turbidite reservoir. Sixteenth International Congress of the Brazilian Geophysical Society (2019).
- Silvany, P., Machado, M., Matos, M., Paes, M., Joint multi-azimuthal pre-stack and time-frequency attributes seismic facies prediction via Deep learning: an application to a Brazilian Pre-salt reservoir. SEG Annual Meeting (2021).
- Qian, F., M. Yin, X.-Y. Liu, Y.-J. Wang, C. Lu, and G.-M. Hu, 2018, Unsupervised seismic facies analysis via deep convolutional autoencoders: Geophysics, 83, no. 3
- Schultz, P., S. Ronen, M. Hattori, and C. Corbert, 1994, Seismic-guided estimation of log properties, Part 2: Using artificial neural networks for non-linear attribute calibration: The Leading Edge, 13, 674–678.
- Simm, R., and M. Bacon, 2014, Seismic amplitude: An interpreter's handbook: Cambridge University Press.
- Vincent P., Larochelle H., Bengio Y., and Manzagol P.-A., "Extracting and Composing Robust Features with Denoising Autoencoders," Proc. Int'l Conf. Machine Learning, 2008.
- Weber R., Schek H.-J., Blott S.: A Quantitative Analysis and Performance Study for Similarity-Search Methods in High-Dimensional Spaces. VLDB Conference Proceedings, 1998.
- Yosinski J, Clune J, Bengio Y, and Lipson H. How transferable are features in deep neural networks?. In Advances in Neural Information Processing Systems 27 (NIPS '14), NIPS Foundation, 2014.