PSO-based ViT-Seismic: A Vision Transformer Approach for Gas Detection in Seismic Images

Domingos A. Dias Junior[1], Luana B. Cruz[2], João O. Bandeira Diniz[3], Aristófanes Corrêa Silva[1], Anselmo C. Paiva[1], Marcelo Gattass[4]

[1]Federal University of Maranhão
Av. dos Portugueses, SN, Campus do Bacanga, Bacanga
65085-580, São Luís, MA, Brazil

[2]Federal University of Cariri
Av. Ten. Raimundo Rocha, 1639
63048-080, Juazeiro do Norte, CE, Brazil

[3]Federal Institute of Maranhão
BR-226, SN, Campus Grajaú, Vila Nova
65940-00, Grajaú, MA, Brazil

[4]Pontifical Catholic University of Rio de Janeiro
R. São Vicente, 225, Gávea,
22453-900, Rio de Janeiro, RJ, Brazil

Abstract. *Seismic reflection is one of the geophysical methods most used in the oil and gas (O&G) industry for hydrocarbon prospecting. In particular, for some Brazilian onshore fields, such a method has been used for estimating the location and volume of gas accumulations. However, the analysis and interpretation of seismic data is time-consuming due to the large amount of information and the noisy nature of the acquisitions. In order to help geoscientists in those tasks, computational tools based on machine learning have been proposed considering Direct Hydrocarbon Indicators (DHIs). In this study, we present a methodology for detection of gas accumulations based on vision Transformer neural network (ViT) and Particle Swarm Optimization (PSO) scheme. In the best scenario, the proposed method achieved a sensitivity of 75.14%, a specificity of 96.14% and an accuracy of 95.60%. We present some tests performed on Parnaíba Basin which demonstrate that the proposed method is promising for gas exploration.*

1. Introduction

Seismic reflection is one of the most used geophysical methods in the oil and gas (O&G) industry to extract information related to geological structures, lithology and rock properties [Pochet et al. 2018, Di et al. 2018, Chevitarese et al. 2018]. Moreover, that method also has been used to estimate the location and volume of gas accumulations, contributing to reducing exploration risks. However, seismic reflection produces a large amount of information, thus it demands considerable time and effort from specialized teams to interpret the seismic data.

With the development of artificial intelligence, several works have proposed methodologies to extract seismic features based on machine learning techniques. For example, Deep Neural Networks have been used for seismic fault detection

[Guitton et al. 2017, Di et al. 2017, Pochet et al. 2018] and seismic facies classification [Zhao 2018, Wrona et al. 2018].

Recently, [Santos 2019] proposed a novel methodology to detect DHIs using seismic data and a Long Short-Term Memory (LSTM) neural network based on a seismic trace scale. In these studies, each seismic trace is divided into patches that are the entrance to the LSTM network along with the labeling of each patch.

Later on, [Santos et al. 2020] proposed the use of transfer learning techniques to expand existing classifier and apply it to different type of seismic surveys into Parnaíba Basin. Moreover, in order to check that methodology based on seismic trace, other networks based on different architectures were developed using an improved encoder-decoder LSTM [Andrade et al. 2021] and a Transformer neural network [Dias et al. 2021]. All those networks have been implemented in ALINE, a computational tool for the assessment of gas accumulations [Santos et al. 2021].

In this study, we propose an enhanced method for detecting potential gas accumulations using the Vision Transformer (ViT) Neural Network. ViT is a classification neural network based on the original Transformer, which was proposed, initially, to solve problems in the area of natural language processing (NLP) [Dosovitskiy et al. 2020a]. Moreover, in order to optimize the ViT model hyperparameters, we use the Particle Swarm Optimization (PSO) algorithm [Le et al. 2019, Júnior et al. 2021]. The results show that our proposal improves the accuracy of the Transformer neural network and increases its efficiency by reducing the spent computational time.

This paper is organized as follows: Section 2 describes the seismic field where the seismic images were acquired. Section 3 presents the proposed method. Section 4 describes the experiments conducted to validate our research. Finally, our conclusions and future works are presented in Section 5.

2. Field Description

The seismic data used as the object of study come from the Paleozoic Parnaíba Basin. The Basin is a classic oval-shaped intracratonic basin developed on a continental basement, during the South American Platform Stabilization Stage. It is located between the Amazonic Craton and the Borborema Province [Almeida et al. 2004, de Miranda et al. 2018]. It covers more than 600,000 km^2 with a depocenter reaching almost 3,500 m in thickness [de Miranda et al. 2018].

The main reservoirs in six of the seven existing fields at the Parnaíba Basin are the Poti Formation sandstones with good poroperm properties in the basin [de Miranda et al. 2018]. The major producing area is known as the 'Parque dos Gaviões' translated as 'Sparrow-hawk's Field' in a reference to the native Brazilian hawk species that the fields are named (Figure 1).

The available data consists of 380 seismic sections located at the Sparrow-hawk's Field area. The database was provided and labeled by Eneva S.A., a Brazilian energy company. These data were obtained at different time intervals with differences in climate, geology, acquisition process, and other external factors. Thus, the data are diversified and heterogeneous.

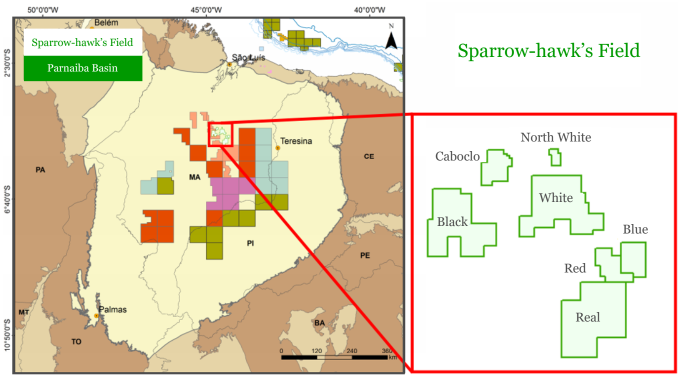

Figure 1. Sparrow-hawk's Field [de Miranda et al. 2018].

3. Proposed Method

First of all, we applied a data preparation and then, the Vision Transformer model is used to detect potential gas accumulations in the seismic images. Figure 2 illustrates each of these steps, and details are provided in the following sections.

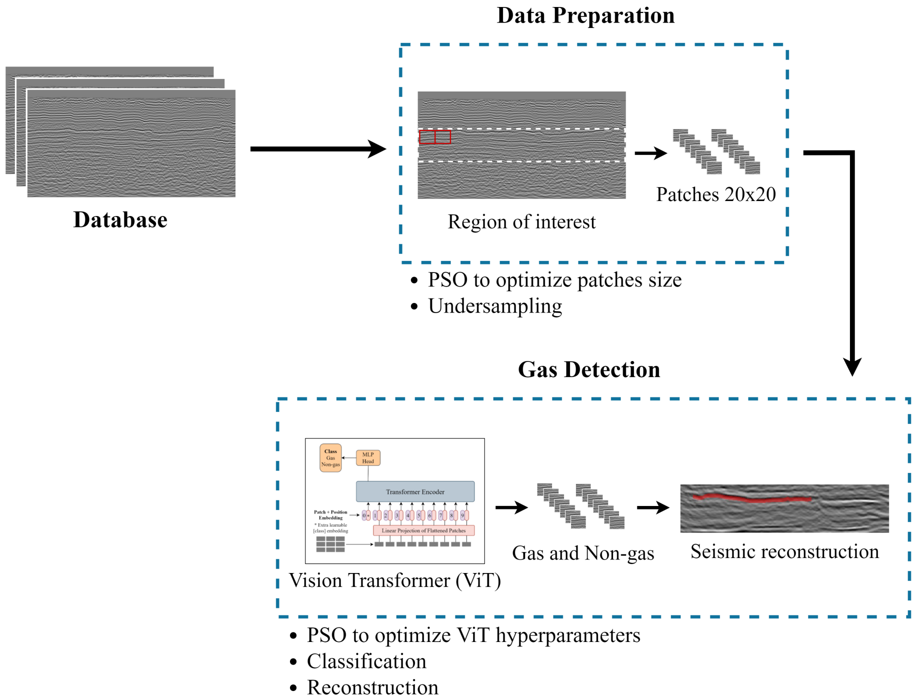

Figure 2. Steps of the proposed method.

3.1. Data Preparation

First, we perform a data preparation step in the seismic images. Based on field data, exploratory wells drilled, and inference, the ENEVA geoscientists delimited the Regions of Interest (ROI) that can contain gas accumulation. The ROI is individual for each image and delimits an area with the seismic patterns that the model must learn to detect the gas or non-gas separation structures. Then, we apply a sliding window process to extract patches over ROIs by sliding window size 20x20, and step equals 1. Finally, the Particle Swarm Optimization (PSO) defines the 20x20 window size for producing the best results.

After the patch generation step, there is an imbalance between gas and non-gas patches at an average ROI ratio of 1:234. This sample imbalance can negatively impact the model performance in learning the correct gas patterns. For this reason, we perform the undersampling technique [Drumnond and Holte 2003] in patches of the predominant class (non-gas) to exclude some random samples to obtain a 1:4 ratio of gas to non-gas samples. The 1:4 ratio produced the best results without compromising computational resources.

3.2. Gas Detection

After data preparation, the next step is to classify the seismic patches as “gas” or ”non-gas” using the ViT model. Figure 3 shows the ViT architecture employed in this study. The architecture consists of an Embedding layer, a stack of Transformer blocks, and a Multilayer Perceptron (MLP). The Embedding layer transforms a 2D image into flattened token sequences, keeping its positional information, to feed the stacked Transformer blocks. A standard Transformer encoder consists of multi-head self-attention layers alternating with MLP blocks. Besides, there is a Layernorm in the block beginning and residual connections at the end of each Transformer block. Finally, an MLP layer is responsible for classifying the samples based on the stacked Transformer blocks output.

ViT has some variations (ViT-Base, Vit-Large, ViT-Huge) that differ from each other due to some hyperparameters: Layers, Hidden Size D, MLP Size, Dropout e Heads. The Layers hyperparameter represents the depth of the network and indicates the number of stacked Transformer encoders. Hidden Size (D) is the dimension that the two-dimensional input samples will be flattened through a linear projection. MLP Size represents the number of neurons in the hidden layer. Dropout is used in MLP to solve the problem of over-adjusting training, and Heads is the number of attention layers present in the Transformer's encoder.

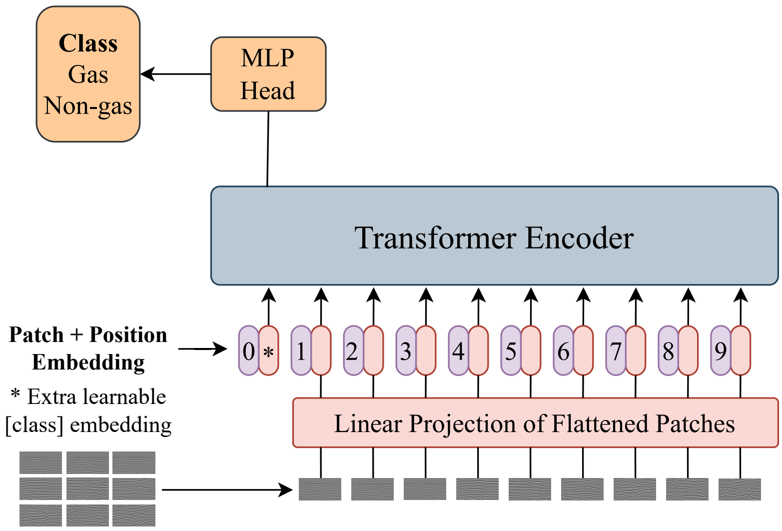

Figure 3. Vision Transformer architecture [Dosovitskiy et al. 2020b].

At last, we use PSO to optimize the ViT model's hyperparameters. We chose

PSO because it provides a high-quality solution in a shorter time, also presents more efficient agility features, and can be more efficient than other optimization techniques [Le et al. 2019, Júnior et al. 2021]. The results achieved in the hyperparameters optimization process are described in Table 1.

Table 1. Optimized hyperparameters of the ViT model using PSO.

Hidden Size D	Layers	Heads	MLP Size	Dropout
256	3	20	512	0.125

The model outputs a binary classification of gas and non-gas, and we use the result to rebuild the final seismic image. For this, we accumulate the output values associated with the same coordinate in the final image. Finally, the resulting image is normalized between 0 and 1. To evaluate the efficiency of the method used, the following validation metrics: accuracy (Acc), sensitivity (Sens), specificity (Spec), and area under roc curve (AUC) [Duda 1973].

4. Results and Discussion

In this section, we show the training environment, the result of each step, and the performance of the method in the case study. The proposed method was implemented by using the Python language. We mainly used the Keras deep learning library [Chollet et al. 2015] with tensorflow-gpu [Abadi et al. 2015] as the back-end. Also, we use a python library pyswarm [Miranda 2017] to perform the Particle Swarm Optimization. The computer used in the experiments consists of an Intel Core i7-9700K 4.20 GHz CPU, 24 GB of RAM, and Nvidia GeForce RTX 2070 super graphics card, running on the Windows 11 operating system. The split of the seismic images 2D dataset for the experiments is described in Table 2.

Table 2. Division of seismic images 2D.

Gaviões Dataset	Train	Validation	Test
Preto	15	2	4
Real	14	2	5
Branco	24	4	7
Vermelho	7	1	2
All	60	9	18

We can see that the datasets are distinct in the number of possible images for training, validation and testing. Thus, due to the amount of training samples, the results may vary due to the number of representative individuals in each dataset.

Next, we show the experiments carried out to validate the proposed method. First, the results are presented in all datasets separately, then the results are presented according to the use of PSO, and finally, the results compared to other approaches are presented.

4.1. Results per Datasets

After splitting the bases, the next phase is extracting patches from each ROI for each image from each dataset. For this, patches of 20×20 were extracted from each image that was the size estimated by the PSO. It is then under sampled to maintain a 1:4 ratio for each gas and non-gas patch (Section 3.1).

With the patches split, the next step is to train the ViT Transformer to classify these patches into gas and non-gas. As highlighted in Section 3.2, ViT hyperparameters have been optimized by PSO. Table 3 describes the results produced by applying the method to the described datasets.

Table 3. Results per Dataset.

Gavião Dataset	Sen	Spec	Acc	AUC
Preto	37.63%	89.29%	88.29%	63.46%
Real	50.13%	94.29%	93.29%	72.21%
Branco	58.02%	88.77%	88.29%	73.39%
Vermelho	67.49%	96.85%	96.20%	82.17%
All	75.14%	96.14%	95.60%	85.64%

We observed that in all datasets, specificity and accuracy metrics higher than 88% were produced. However, the Gavião Preto dataset produced a low sensitivity in relation to the others. This result can be attributed to poor and heterogeneous seismic data quality [Santos 2019].

On the other hand, we noticed that when we train all the datasets together, the results produce the best metrics. This is justified by the fact that there is an increase in the variability of the data, making the ViT better learn the patterns of differentiation between the classes, which, consequently, increases the generalization power of the network, producing metrics of 75.14% of sensitivity, 96.14% specificity, 95.60% accuracy and 85.64% AUC. Thus, we show the effectiveness of the method in gas detection using the PSO-optimized ViT Transformer.

4.2. Results: Vit with PSO and without

It is worth mentioning that the use of PSO for optimization of hyperparameters, since the search for these hyperparameters is something tiring and susceptible to great variation given the search space, when using PSO we can find them automatically and improve even more the ViT performance. Thus, in order to verify the efficiency of the PSO, we present the Table 4, where we show the results achieved by ViT with its default hyperparameters and optimized by the PSO in the all dataset.

Table 4. Results with and without PSO.

Gavião Dataset	Sen	Spec	Acc	AUC
Vit	63.81%	90.20%	89.44%	77.01%
Vit + PSO	75.14%	96.14%	95.60%	85.64%

We can see that the use of PSO provided a significant improvement in validation metrics. Where sensitivity has improved by more than 11%, this means that more gas regions are being found. Furthermore, the specificity improves by almost 6%, which shows that with the use of PSO the method produced fewer false positives. Thus, we emphasize that the use of PSO for ViT optimization was essential to produce promising results.

4.3. Comparison with other approaches

In this section, we present a comparison of the results achieved in the 'All dataset' with the work proposed by [Santos 2019] that uses LSTM and we also trained and tested a LeNet-5 [LeCun et al. 1998] network to validate the effectiveness of ViT in relation to a conventional CNN. The Table 5 displays the results.

Table 5. Results with and without PSO.

Gavião Dataset	Sen	Spec	Acc	AUC
LSTM	52.99%	96.69%	93.97%	74.84%
LeNet-5	30.54%	93.61%	90.57%	62.08%
ViT Transformer	75.14%	96.14%	95.60%	85.64%

We observed that the proposed method surpasses the other comparatives in relation to sensitivity. We highlight once again the sensitivity metric as being crucial, given the importance of gas detection. Compared to LeNet-5, our method outperforms all validation metrics, showing its generalization power compared to a conventional CNN. On the other hand, the work of [Santos 2019] presents metric of specificity slightly higher than the proposed method. However, it is worth noting that our method produces greater sensibility and accuracy in the gas class, which demonstrates greater robustness.

4.4. Case Study

To evaluate the results achieved in the proposed method, we define two case studies. In the first case, the model can detect the gas reservoir effectively. In the second case, the model presents some deficiencies in gas reservoir detection results.

In Figure 4, we can see three cases that had good results in detecting gas reservoirs. However, some false positives are generated (in red), the method can distinguish the aimed region (in blue), which can facilitate the analysis of the data by an expert. Thus, these cases demonstrate that the proposed method is promising for both quantitative and qualitative results. It is worth mentioning that several similar results were found across all datasets but could not be shown due to the page limit.

The second case study is illustrated in Figure 5. Then, in these cases, we can assume that the proposed method cannot detect some potential gas reservoirs. Although the model hit some aimed regions, most of the gas reservoirs regions were not detected (in green). Besides, ViT confuses the gas prediction with similar regions, generating false-positive predictions.

It is worth remembering that data analysis is not a trivial task. For this reason, it requires expert experience and is time-consuming. Therefore, we believe that the proposal, combined with the expert's knowledge in data analysis, can be a faster method of identifying potential gas reservoirs.

5. Conclusion

In this work a method for gas detection using seismic data was proposed. For this, a ViT Transformer network optimized by PSO was proposed. The proposed method used a 2D approach, with an architecture based on attention mechanisms, which was initially proposed to solve problems in the area of natural language processing, but has been applied to

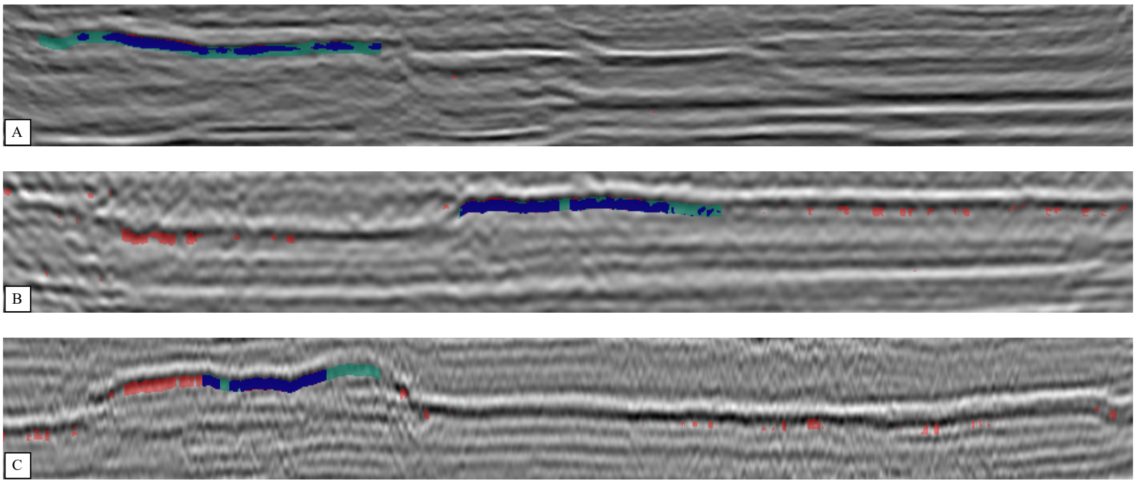

Figure 4. Case Study 1: (a), (b) and (c) represent three different seismic images. In red, it represents false positives. In blue, the true positives. In green, false negatives.

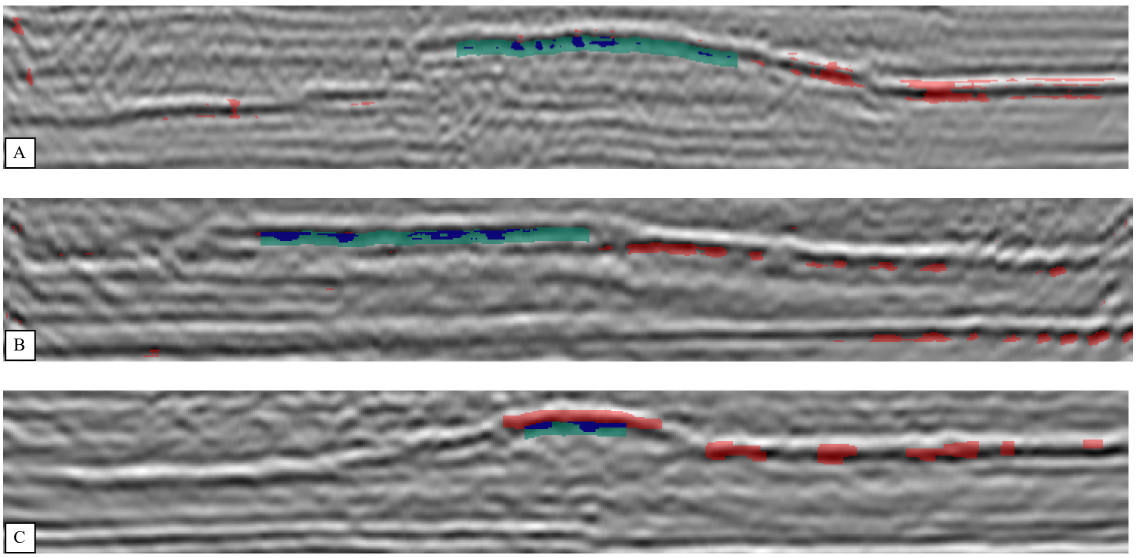

Figure 5. Case Study 2: (a), (b) and (c) represent three different seismic images. In red, it represents false positives. In blue, the true positives. In green, false negatives.

other areas such as image processing. The proposed method consists of an improvement and adjustment in the parameters of the ViT Transformer using an evolutionary algorithm to better adjust it to the recognition of patterns in seismic images.

The results achieved by the proposed method are promising. The method was efficient in gas detection, presenting a sensitivity superior to other networks consolidated in the literature. The use of PSO to optimize ViT hyperparameters also proved to be an important step, producing an improvement in all validation metrics. With this, it is believed that the proposed method can be crucial, combined with the practice of the specialist, for the gas detection.

As future work, we suggest the validation of the method in a more robust database, since given the limitation and heterogeneity of the data, the model may not have reached its full generalization. Also, an adaptation of ViT so that it works with a semantic segmentation network to improve the performance of the results already achieved by ViT-Seismic. Finally, another possible improvement would be to combine the 1D information with the 2D information achieved by ViT.

References

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G. S., Davis, A., Dean, J., Devin, M., et al. (2015). Tensorflow: Large-scale machine learning on heterogeneous systems. *Software available from tensorflow.org*.

Almeida, F. d., Carneiro, C. D. R., et al. (2004). Inundações marinhas fanerozóicas no brasil e recursos minerais associados. *Mantesso Neto, V.; Bartorelli, A.; Carneiro, CDR*, pages 43–60.

Andrade, F., Fernando Santos, L., Gattass, M., Quevedo, R., Michelon, D., Siedschlag, C., and Ribeiro, R. (2021). Gas reservoir segmentation in 2d onshore seismics using lstm-autoencoder. In *First International Meeting for Applied Geoscience & Energy*, pages 1651–1655. Society of Exploration Geophysicists.

Chevitarese, D. S., Szwarcman, D., e Silva, R. G., and Brazil, E. V. (2018). Deep learning applied to seismic facies classification: A methodology for training. In *Saint Petersburg 2018*, volume 2018, pages 1–5. European Association of Geoscientists & Engineers.

Chollet, F. et al. (2015). Keras. `https://keras.io`.

de Miranda, F. S., Vettorazzi, A. L., da Cruz Cunha, P. R., Aragão, F. B., Michelon, D., Caldeira, J. L., Porsche, E., Martins, C., Ribeiro, R. B., Vilela, A. F., et al. (2018). Atypical igneous-sedimentary petroleum systems of the parnaíba basin, brazil: seismic, well logs and cores. *Geological Society, London, Special Publications*, 472(1):341–360.

Di, H., Shafiq, M. A., and AlRegib, G. (2017). Seismic-fault detection based on multi-attribute support vector machine analysis. In *SEG Technical Program Expanded Abstracts 2017*, pages 2039–2044. Society of Exploration Geophysicists.

Di, H., Wang, Z., and AlRegib, G. (2018). Seismic fault detection from post-stack amplitude by convolutional neural networks. In *80th EAGE Conference and Exhibition 2018*, volume 2018, pages 1–5. European Association of Geoscientists & Engineers.

Dias, D., Diniz, P., Marin, L., Cipriano, C., Gattass, M., Santos, L., Quevedo, R., Michelon, D., Siedschlag, C., and Ribeiro, R. (2021). Automatic gas detection in land seismic using transformer neural network. In *17th International Congress of the Brazilian Geophysical Society*. Brazilian Geophysical Society.

Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., et al. (2020a). An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv preprint arXiv:2010.11929*.

Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., Uszkoreit, J., and Houlsby, N.

(2020b). An image is worth 16x16 words: Transformers for image recognition at scale. *CoRR*, abs/2010.11929.

Drumnond, C. and Holte, R. (2003). Class imbalance and cost sensitivity: Why undersampling beats oversampling. In *ICML-KDD 2003 Workshop: Learning from Imbalanced Datasets*, volume 3.

Duda, R. (1973). *Pattern classification and scene analysis*. Wiley-Interscience Publication, 512.

Guitton, A., Wang, H., and Trainor-Guitton, W. (2017). Statistical imaging of faults in 3d seismic volumes using a machine learning approach. In *SEG Technical Program Expanded Abstracts 2017*, pages 2045–2049. Society of Exploration Geophysicists.

Júnior, D. A. D., da Cruz, L. B., Diniz, J. O. B., da Silva, G. L. F., Junior, G. B., Silva, A. C., de Paiva, A. C., Nunes, R. A., and Gattass, M. (2021). Automatic method for classifying covid-19 patients based on chest x-ray images, using deep features and pso-optimized xgboost. *Expert Systems with Applications*, 183:115452.

Le, L. T., Nguyen, H., Zhou, J., Dou, J., Moayedi, H., et al. (2019). Estimating the heating load of buildings for smart city planning using a novel artificial intelligence technique pso-xgboost. *Applied Sciences*, 9(13):2714.

LeCun, Y., Bottou, L., Bengio, Y., and Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324.

Miranda, L. J. V. (2017). Pyswarms, a research-toolkit for particle swarm optimization in python.

Pochet, A., Diniz, P. H., Lopes, H., and Gattass, M. (2018). Seismic fault detection using convolutional neural networks trained on synthetic poststacked amplitude maps. *IEEE Geoscience and Remote Sensing Letters*, 16(3):352–356.

Santos, L., Jordao, F., Gattass, M., Quevedo, R., Lima, M. J., Michelon, D., Siedschlag, C., Ribeiro, R., and Pereira, S. (2021). Natural gas detection in onshore data using transfer learning from a lstm pre-trained with offshore data. Society of Exploration Geophysicists.

Santos, L. F., Gattass, M., Silva, A., Miranda, F., Siedschlag, C., and Ribeiro, R. (2020). Natural gas detection in onshore data using transfer learning from a lstm pre-trained with offshore data. In *SEG Technical Program Expanded Abstracts 2020*, pages 1190–1195. Society of Exploration Geophysicists.

Santos, L. F. T. (2019). Detector de assinaturas de gás em levantamentos sísmicos utilizando lstm. Master's thesis, Pontifícia Universidade Católica do Rio de Janeiro.

Wrona, T., Pan, I., Gawthorpe, R. L., and Fossen, H. (2018). Seismic facies analysis using machine learning. *Geophysics*, 83(5):O83–O95.

Zhao, T. (2018). Seismic facies classification using different deep convolutional neural networks. In *SEG Technical Program Expanded Abstracts 2018*, pages 2046–2050. Society of Exploration Geophysicists.