
**Submission code: 4PAR8KNL7V**

See this and other abstracts on our website: https://home.sbgf.org.br/Pages/resumos.php

# Application study of a Transformer Mask U-net model for basalt layer scattering noise attenuation on onshore seismic data

**Victor Ferreira (NCA UFMA), Mauricio Morais Almeida (NCA UFMA), Alan Araújo (NCA UFMA), Luis Sepulveda (Tecgraf PUC-Rio), Carlos Rodriguez (Tecgraf Institute/PUC-Rio), Deane Roehl (Tecgraf Institute/PUC-Rio), Gabriel Costa (NCA UFMA), Aristófanes Silva (NCA UFMA), Anselmo Paiva (NCA UFMA)**

# Application study of a Transformer Mask U-net model for basalt layer scattering noise attenuation on onshore seismic data



## Abstract

Seismic data from onshore basins with volcanic activity, such as the Paraná Basin in Brazil, often exhibit intense scattering noise due to basalt layers, impairing subsurface data. This study proposes a hybrid deep learning model combining Transformer, Mask U-Net, and Fourier Neural Operator components to attenuate low-frequency linear noise. The model is trained and evaluated using synthetic and real pre-stack seismic datasets. Results show significant noise suppression in synthetic data, achieving a PSNR of 44.68 and an SSIM of 0.98. Real data results indicate partial noise attenuation (PSNR of 23.84 and SSIM of 0.35), highlighting both the method's potential and the challenges in real-world applications.

## Introduction

Seismic data are an essential resource for various fields of geosciences and are widely used in exploring hydrocarbon reservoirs. However, the quality and representativeness of the data determine its usefulness in a significant way (Sepulveda et al., 2023).

In some onshore basins in Brazil extensive volcanic paleo activity is present, generating a complex system of dikes and sills, and also producing seismic data that can be considered of low quality, given the presence of noise. This noise may originate in the basalt layer , which often has irregular geometry at the top and bottom, causing seismic signals to be degraded in these geological settings (Costa et al., 2016). This noise seems to be concentrated in the low frequencies, and it is a problem because it prevents the identification of geological formations beneath the basalt layers and therefore hampers the search for hydrocarbon reserves.

Considering the problem mentioned above, this work aims to apply a method that integrates various Deep Learning (DL) techniques, such as Transformer, Mask U-net, and Fourier Neural Operator, to attenuate the noise in seismic data from regions with basalt layers.

## Method and Theory

This section describes the model created and the techniques used to attenuate noise in 2D pre-stack seismic data, using Machine Learning (ML) and DL resources. Initially, the datasets used in the analysis are presented.

- **Seismic Dataset**

Two datasets were used in this work, both in shot gathers presenting pre-stacked seismograms.

The first dataset comprises 462 synthetic seismograms, presenting 21 base seismograms with 21 different coherent noise levels (diffractions that seek to emulate the linear noise seen on real data) for each. All samples from the same base seismogram have the same signal amplitude; only the noise changes.

The second dataset contains 469 real seismograms from 2D seismic line 236-0062 (ANP-REATE, 2021) in Paraná Basin. To make this dataset useful for the ML training process, it was subjected to a processing pipeline to create an additional dataset, which is a noise-reducing version of the original base, as shown in (a,b) of Figure 1. This dataset is used as ground truth (target) for ML.

Figure 1 (c) presents the proposed pipeline; for both datasets, it was necessary to include a preprocessing step to apply a normalization between [-1,1] that takes into account the maximum value of the entire line.

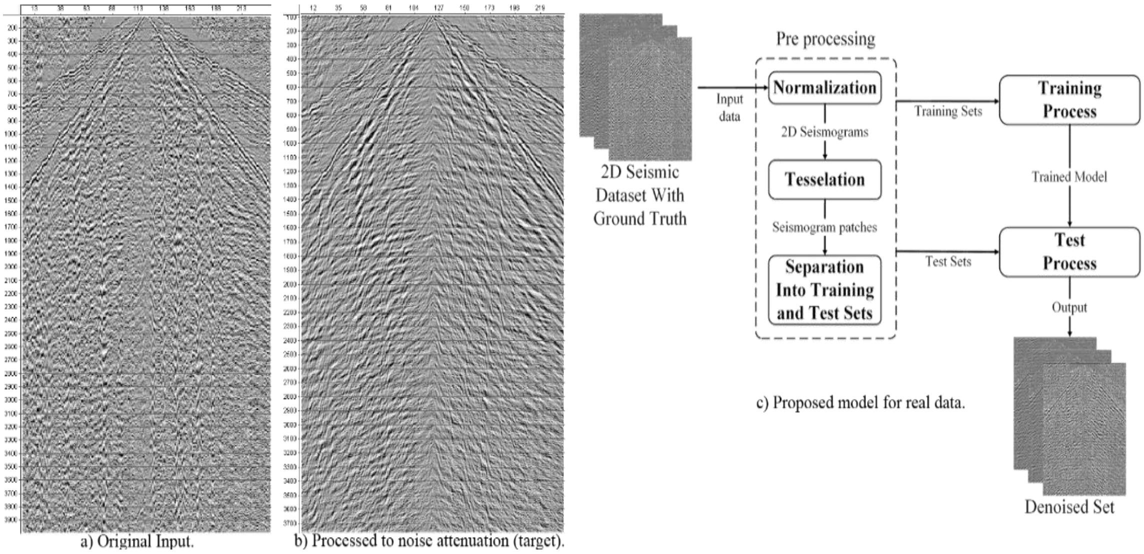


**Figure 1:** Real dataset: a) Original input, b) Seismogram processed to noise attenuation, c) Proposed model.

- **Transformer Mask U-net**

This is the network used to process both databases. It is based on the U-Net model, which was chosen for its good performance in addressing noise removal tasks (Vaswani et al., 2023). However, several layers with different techniques were integrated. Transformer is used in the network's bridge to focus learning on the most relevant features, which is possible thanks to this technique's self-attention (Vaswani et al., 2023). The mask forces the network to learn to relate spatial features based on neighboring data (You et al., 2025). Finally, the inclusion of Fourier Neural Operator (FNO) layers ensures that learning is focused on low frequencies (Li et al., 2021).

Figure 2 presents the proposed net architecture. It is clarified that all layers of the bridge are subject to a spatial change that depends on the learning process itself, taking the stagnation of the loss function as a measure of change. This particularity shifts from temporal space to frequency space and vice-versa when it is detected that the loss function does not improve over a certain number of epochs.

The proposed network has three depth levels, taking 480x16 seismic data patches as input, with more than 100 million trainable parameters and 100 iterations. The proposed architecture pays special attention to the data encoding stage, seeking to identify relevant features in the data for specific element elimination.

After receiving the patches, residual blocks are applied, followed by convolution and the masking process, which hides part of the data to force the network to reconstruct it based on its neighbors (this helps the network understand each element's relationship with its environment). At the next level, convolutional blocks that include normalization are applied, followed by FNO that focus on learning low frequencies, and finally residual blocks. At the next two levels, this structure is repeated. Finally, at the bridge, convolutional and residual blocks are applied consecutively, followed by additional masking, three consecutive transformer layers, and simple convolution and residual blocks.

The deconvolution process is comparatively simpler than the encoding process, at each level it receives a skip connection and then applies deconvolutions until it reaches the top level and reaches the same size as the input data

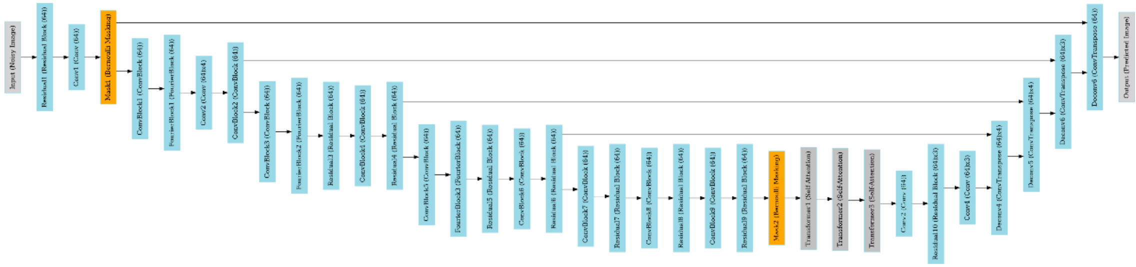

**Figure 2:** Proposed Net: blue color indicates convolution-based layers, orange shows the position of the masking process, and gray indicates Transformer layers. See text for explanation.

## Results

**Synthetic dataset:** 60%, 20%, and 20% of the data were randomly selected for training, validation, and testing, respectively. The metrics in the test set reached a PSNR of 44.68 ± 2.32 and an SSIM of 0.98 ± 0.01, indicating that the trained model manages to mute the noise in the seismograms almost completely.

Figure 3 shows two results. On a) it is clear that the noise has been almost completely muted, but b) shows that there are cases where the noise persists, although to a much lesser extent. The presence of remaining noise, in addition to the artifacts in the upper section, indicates that the model still can be improved.

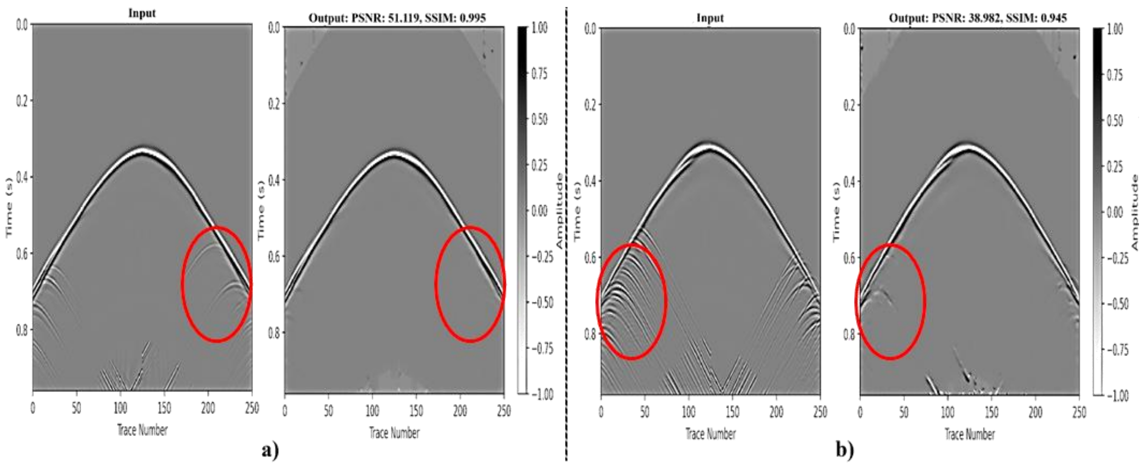


**Figure 3:** Result in synthetic data, a) Noise almost completely muted, b) Presence of remaining noise.

**Real dataset:** 50%, 20%, and 30% of the data were randomly selected for training, validation, and testing, respectively. The metrics in the test set reached a PSNR of 23.84 ± 0.73 and an SSIM of 0.35 ± 0.03. The results contrast with those obtained in the synthetic data set and demonstrate the complexity of the problem.

Figure 4 shows the result in a real seismogram, in which it can be seen that the output of the proposed model b) begins to make the modifications to get closer to the Target c), however, when analyzing the difference d) it is observed that the changes made are slight, which indicates that the model identifies part of the components to be eliminated.

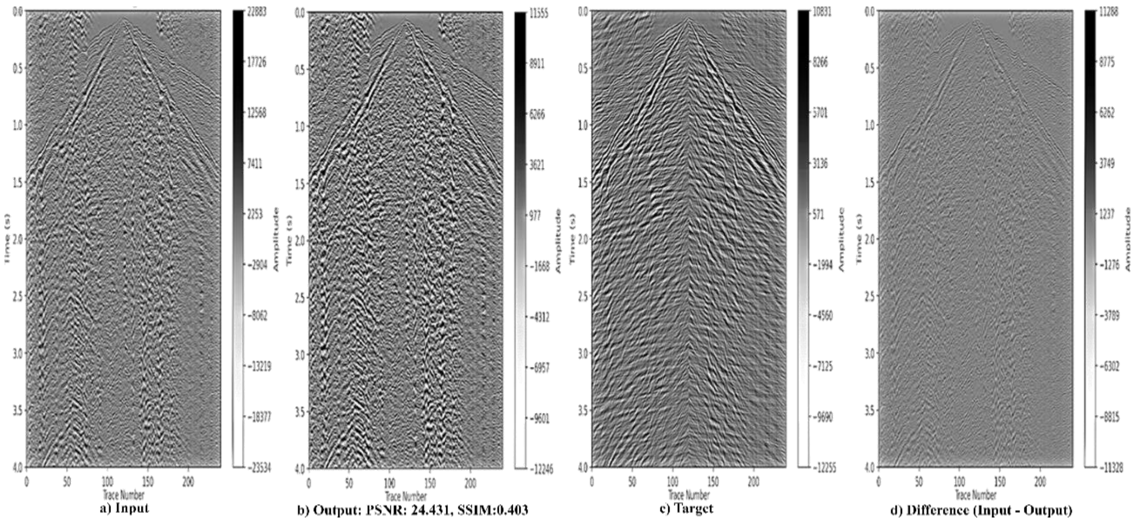


**Figure 4:** Result in real data, a) Original input, b) Output from ML model, c) Target (filtered seismogram), d) Difference between input and output.

## Conclusions

The results obtained on the synthetic dataset indicate that the model is capable of recognizing noise and removing it from seismograms. However, when analyzing the results with real data, it is evident that the synthetic data still does not accurately represent the properties of the problem. This also demonstrates that the model needs to increase its complexity to capture the features of noise from basalt layers.

## Acknowledgments

The authors would like to thank the Brazilian National Agency for Petroleum, Natural Gas and Biofuels (ANP) and CNPC Brasil for subsidizing this research project through the R&D Clause.

## References

ANP/REATE, 2021, Public Land Data Free Access. https://reate.cprm.gov.br/anp/TERRESTRE

Costa, D. F. B. D., W. H. D. Santos, S. Bergamaschi, and E. Pereira, 2016, Analysis of the geometry of diabase sills of the Serra Geral magmatism, by 2D seismic interpretation, in Guareí region, São Paulo, Paraná basin, Brazil: Brazilian Journal of Geology, 46, no. 4, 605–615.

Li, Z., N. Kovachki, K. Azizzadenesheli, B. Liu, K. Bhattacharya, A. Stuart, and A. Anandkumar, 2021, Fourier neural operator for parametric partial differential equations.

Sepulveda, L. F. M., M. Gattass, A. C. Silva, R. Quevedo, D. Michelon, C. Siedschlag, and R. Ribeiro, 2023, Seismic data classification for natural gas detection using training dataset recommendation and deep learning: Geoenergy Science and Engineering, 228.

Vaswani, A., N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin, 2023, Attention is all you need.

You, N., T. Lin, and W. Li, 2025, Masked-swinunet-powered microresistivity borehole image inpainting: GEOPHYSICS, 90, no. 2, D47–D59.