Estimating Initial Velocity Models for the FWI Using Deep Learning

Luan Campos[1]; Peterson Nogueira[1,2]; Erick Nascimento[1,3]

[1] *SENAI CIMATEC Supercomputing Center for Industrial Innovation*

[2] *UFBA/INCT-GP*

[3] *SENAI CIMATEC Researcher/Post-Graduate Program*

Abstract

Building satisfactory initial velocity models is a crucial task needed to achieve valuable results without spending too much computational resources when using the full waveform inversion (FWI) method. Many techniques are employed as an attempt to find an initial model, such as reflection tomography, migration-based velocity analysis, genetic algorithms and simulated annealing. Recently, researchers are also experimenting the use of deep learning methods to build velocity models, but not necessarily to be used with the FWI. This work is twofold: 1) to use a fully convolutional network (FCN) to estimate an initial velocity model and 2) to evaluate the resulting model on a multi-scale FWI implementation. The FCN was trained using synthetic velocity models randomly built as its target and their respective seismic data modelled by the finite-difference method as its input. The models produced by the FCN lacked some details, such as fault structures or positioning of some layers, when compared to their ground-truth counterpart, but that was fixed after using the FWI method. Thus, we conclude that the results produced by the FCN proved to be, in overall, satisfactory initial models for the FWI.

Introduction

The FWI is a high resolution method for subsurface parameters inversion and requires a satisfactory initial model since it tries to solve a non-linear problem (Virieux and Operto (2009)). In the oil and gas exploration industry, the initial velocity model for the FWI is commonly built by reflection tomography and migration-based velocity analysis. Most methods proposed for this procedure are based on time-travel tomography, but the possibility of avoiding such methods is desirable because they can be an exceptionally time-consuming task. Alternatives using global methods, such as genetic algorithms and simulated annealing, have been respectively proposed by Sajeva et al. (2016) and Datta and Sen (2016). However, the bigger the size of the velocity models or the population being considered, the higher the number of modelling steps and their computational cost. Lately, there has also been an increasing appeal of using deep learning methods to solve a diverse range of problems from different areas of knowledge, including geophysics applications (Lewis and Vigh, 2017); (Araya-Polo et al., 2018); (Wu et al., 2018).

In this context, we analyse the use of a FCN to generate initial models for the FWI. We implement the FWI method on the time domain using a multi-scale approach on the frequency domain with the finite-difference method to solve an acoustic wave equation. The FCN follows the architecture used by Wang et al. (2018) with modifications on its optimiser and metrics. We also applied a different positioning scheme of sources and receivers to simulate a more conventional seismic acquisition. Therefore, our work aims to estimate the initial model, instead of directly estimating the final model with the FCN as proposed by Wang et al. (2018), and to prove that the FCN can produce an optimal initial model for the FWI.

Fully Convolutional Network

Convolutional neural networks (CNNs) try to mimic some features of the visual cortex of the brain (Min et al., 2017) and they are well-suited to work on two-dimensional images. Even though the CNN was firstly proposed by LeCun et al. (1989), its usage has been gaining even more terrain since the implementation of the AlexNet (Krizhevsky et al., 2012), a variation of the original CNN which achieved a significant performance on the ImageNet database. Other variations of the conventional CNN were also proposed, such as the fully convolutional network (Long et al., 2015).

For this work we use an FCN closely related to the one discussed by Wang et al. (2018): an encoding section consisting of successive layers of 3x3 convolutions, using batch normalisation and the rectified linear unit (ReLU) activation function, and 2x2 max-pooling, a decoding section having 2x2 up-sampling layers in conjunction with 2x2 convolutions followed by 3x3 convolution layers, and a final 1x1 convolution layer to yield the predicted model applied in a section of the last 3x3 convolution output cropped to match the velocity model's size. On one hand, the convolution layers organise feature maps that perform filtering operations mathematically similar to a discrete convolution, whilst the pooling layers combine related features into one (LeCun et al., 2015). On the other hand, up-sampling layers substitute the pooling layers in the reverse operation to increase the resolution of the network's output. This FCN model was trained on two NVIDIA Tesla P100-SXM2 GPUs with 16GB RAM each

during 200 epochs using a batch size of 5 samples.

Our approach, when compared to Wang et al. (2018), has some core differences. Firstly, although using the same U-Net configuration and loss function as theirs, we used the Adam optimiser (Kingma and Ba, 2014) implementation available with Keras (www.keras.io) and a learning rate of 1×10^{-3} instead of the default Stochastic Gradient Descent (SGD). Furthermore, since predicting the velocity model can be seen as a regression problem, we evaluated other four metrics besides the loss function - mean squared error (MSE) -, namely: mean absolute error (MAE), R^2, Pearson coefficient r and factor of two. Another key difference is the fact that we use larger velocity models and our seismic data were modulated with a different configuration of sources and receivers, as explained further.

Synthetic Velocity Model

The velocity models represent a section of subsurface of 3000 meters in depth by 9225 meters in length and they were generated from a predefined configuration of minimum and maximum velocity, number of samples in the z (nz) and x (nx) axes, number of spacing intervals on both axes (dz and dx) and minimum and maximum number of layers. We set the minimum and maximum velocities to 1500 and 3500 m/s respectively, 150 samples of both z and x axes with spacing interval of 20 and 61.5 meters, respectively, and the number of layers varying from 8 to 12.
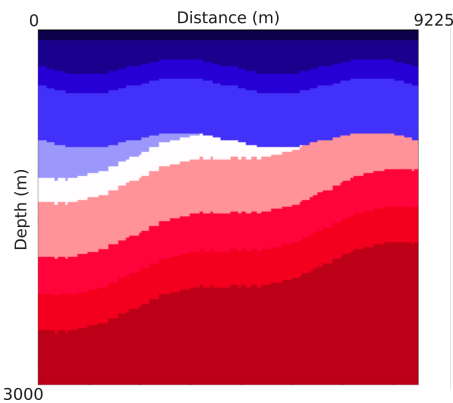

Figure 1: Velocity model generated with 10 layers, undulation, inclination, but no fault structures

For each velocity model generated, we randomly draw its number of layers and whether the layers should be undulated, inclined or contain fault structures. We randomly pick these parameters so the FCN could be trained with a large variety of models. The velocity of each layer increases in a rate of $\frac{(MAX\_VEL - MIN\_VEL)}{num\_layers}$. Figure 1 shows one of the generated velocity models.

Data Acquisition

The seismic data were generated using the finite-difference method of 32nd order for an acoustic wave equation with a Ricklet wavelet having a maximum frequency of 18.4

Hz filtered to be 12.20 Hz. We used 50 sources equally spaced $184.5\ m$ one from another on the x axis and 150 receivers spaced $61.5\ m$ and placed on the same level of the sources. We chose this configuration of sources and receivers because it is more likely to occur in real data acquisitions. Our number of samples in the time domain is 1500 (nt) with a sample interval of 0.002 (dt), totalling 3 seconds of sampling. The modelling step used in the FWI method follows the same configuration as the one used to model the seismic data, but using the frequency of 18.4 Hz instead of the 12.20 Hz.

Results

The evolution of the FCN was monitored every two epochs during the training step by predicting and plotting one of the models from the testing dataset. We can notice in Figure 2 that, although the result during the first epoch seems to be only composed of a water layer, the FCN model is able to gradually identify not only the beginning and ending of more layers, but also their undulation and inclination. However, two problems can be primarily pointed out: by the end of the training neither of fault structures, nor the beginning and ending of some layers were correctly identified.

The FCN took 18 hours to train all 200 epochs, but predicting 20 models from the testing dataset only took, approximately, 16 seconds. This means that once the seismic data is acquired, an initial velocity model can be inferred in less than one second. Figure 3 shows two examples of predicted models to aid our analysis of the problems previously stated. Table 1 shows the evaluated metrics with the fully trained model on the testing dataset. The goal is to minimise the MSE and MAE metrics, whilst maximising R^2, Pearson r and Factor of Two to as close as possible of 1. As we can see, the model achieved a value relatively low of approximately 65.6 for the mean absolute error, meaning that the velocities of the predicted models differ more or less 65.6 m/s from the true model, in average. Furthermore, we obtained desirable values of R^2 and Pearson r that are close to 1 and Factor of Two exactly 1. This indicates, as we can confirm in Figure 3, that the predicted model is highly correlated to its true counterpart. However, the high value of MSE justifies the lack of precision in determining the beginning and ending of inclined layers (as seen in models 3a and **??**) and identifying fault structures (models 3c and **??**).

Metric	Value
MSE	10188.0654
MAE	65.5954
R^2	0.9671
Pearson r	0.9840
Factor of two	1.0000

Table 1: Evaluation Metrics of the FCN Model

The next step was to feed one of the predicted models to the FWI method. Once again we chose the model depicted in 1. As it was previously noted, the predicted model does not contain the fault structures and some layers are incorrectly positioned. Both problems are respectively identified in Figure 4b by an ellipsis and two arrows.

(a) Epoch 1   (b) Epoch 75   (c) Epoch 151   (d) Epoch 199
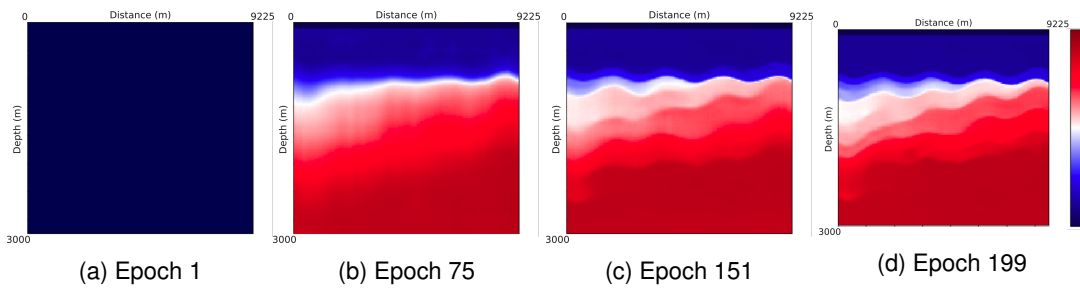
Figure 2: Prediction output during the training step at, respectively, the a) 1st, b) 75th, c) 151st and d) 199th epochs of the model introduced in Figure 1

(a) Predicted 1   (b) Ground-truth 1   (c) Predicted 2   (d) Ground-truth 2
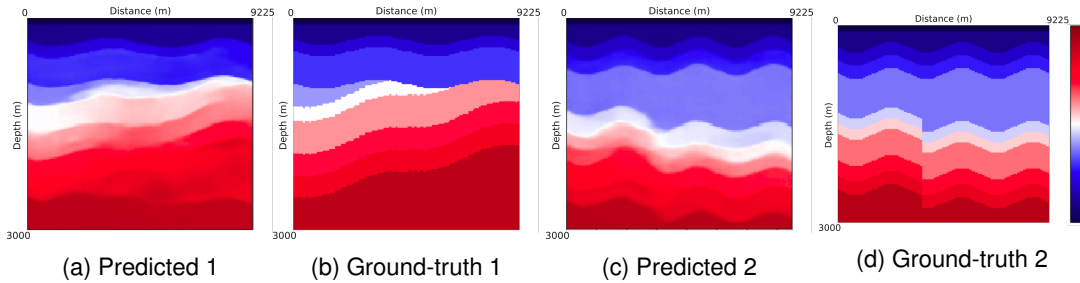
Figure 3: Results of two velocity models from the testing dataset along with their respective ground-truth model. a) The first model has undulated and inclined layers, b) the second model has undulated layers and fault sections

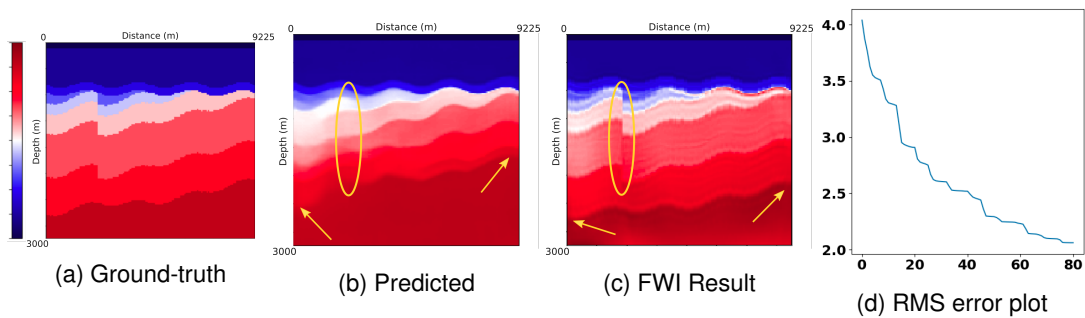(a) Ground-truth   (b) Predicted   (c) FWI Result   (d) RMS error plot

Figure 4: Comparison of (a) the ground-truth model, (b) the predicted model (FCN result), (c) the model after using the FWI method and (d) the decreasing rate of the RMS error on each iteration of the FWI

Nonetheless, this lack of details is expected because the FCN produces an initial velocity model, whilst the FWI is responsible for generating the final inverted model. As we can see in Figure 4d, we can confirm again that the predicted velocity model poses as a satisfactory initial model due to the low RMS error at the beginning of the FWI. Furthermore, the FWI was indeed able to further improve the velocity model by identifying the once absent fault structures and adjusting the position of some layers. This improvement is shown by the ellipsis and arrows in Figure 4c and by the decreasing rate of the RMS error in Figure 4d.

Conclusion

The FCN demonstrated to be a valuable tool for building optimal initial velocity models for the FWI method. Even though the initial velocity models generated by this technique lack details of fault structures and have

difficulties in determining the correct position of inclined layers, once the training and validation phases were performed, the method is pretty fast for inferring an initial velocity model, requiring quite few computational resources and less human interference to provide an initial model highly correlated to the ground-truth model, containing low error rate on the velocity values. Furthermore, since it generates an initial velocity model to be used as an input to the FWI technique, it is this combination of such methods that aims to adequately predict the final velocity model. So, using the initial model that has been inferred by the FCN, the FWI finally corrected the fault structures identification and inclined layer positioning. Further studies point to the improvement of this technique for building initial velocity models, the comparison of it with the genetic algorithm and particle swarm optimisation techniques and the usage of other deep learning methods and higher complexity models, such as the Marmousi.

References

Araya-Polo, M., J. Jennings, A. Adler, and T. Dahlke, 2018, Deep-learning tomography: The Leading Edge, **37**, 58–66.

Datta, D. and M. K. Sen, 2016, Estimating a starting model for full-waveform inversion using a global optimization method: Geophysics, **81**, R211–R223.

Kingma, D. P. and J. Ba, 2014, Adam: A method for stochastic optimization: arXiv preprint arXiv:1412.6980.

Krizhevsky, A., I. Sutskever, and G. E. Hinton, 2012, Imagenet classification with deep convolutional neural networks: Advances in neural information processing systems, 1097–1105.

LeCun, Y., Y. Bengio, and G. Hinton, 2015, Deep learning: nature, **521**, 436.

LeCun, Y., B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, and L. D. Jackel, 1989, Backpropagation applied to handwritten zip code recognition: Neural computation, **1**, 541–551.

Lewis, W. and D. Vigh, 2017, Deep learning prior models from seismic images for full-waveform inversion, *in* SEG Technical Program Expanded Abstracts 2017, 1512–1517, Society of Exploration Geophysicists.

Long, J., E. Shelhamer, and T. Darrell, 2015, Fully convolutional networks for semantic segmentation: Proceedings of the IEEE conference on computer vision and pattern recognition, 3431–3440.

Min, S., B. Lee, and S. Yoon, 2017, Deep learning in bioinformatics: Briefings in bioinformatics, **18**, 851–869.

Sajeva, A., M. Aleardi, E. Stucchi, N. Bienati, and A. Mazzotti, 2016, Estimation of acoustic macro models using a genetic full-waveform inversion: Applications to the marmousi modelgenetic fwi for acoustic macro models: Geophysics, **81**, R173–R184.

Virieux, J. and S. Operto, 2009, An overview of full-waveform inversion in exploration geophysics: Geophysics, **74**, WCC1–WCC26.

Wang, W., F. Yang, and J. Ma, 2018, Velocity model building with a modified fully convolutional network, *in* SEG Technical Program Expanded Abstracts 2018, 2086–2090, Society of Exploration Geophysicists.

Wu, Y., Y. Lin, and Z. Zhou, 2018, Inversionnet: Accurate and efficient seismic waveform inversion with convolutional neural networks, *in* SEG Technical Program Expanded Abstracts 2018, 2096–2100, Society of Exploration Geophysicists.