Machine learning for seismic processing: The path to fulfilling promises

Song Hou and Jérémie Messud, CGG*

Summary

Machine learning (ML) has garnered great attention within the field of seismic processing due to its vast achievements for quality and efficiency in the area of computer vision. Recent academic papers have demonstrated some potential for the use of machine learning in processing seismic signal, such as random and coherent noise removal, deblending, and interpolation. In this paper, we illustrate some uses of ML on real 3D seismic data and discuss the common challenges that need to be addressed in order to fulfill the promises of the deep neural network (DNN) for seismic processing. We also point out that, in some cases, the result of ML could be good enough for some fit-for-purpose applications. Finally, we summarize a few learnings based on our research and experiences in both the seismic processing and ML worlds.

Introduction

ML algorithms are usually categorized as supervised or unsupervised. DNN, a family of supervised learning, represents a way to efficiently parameterize a non-linear function that maps an input to a given output (or target). For seismic processing, the input often relates to the raw seismic data to be processed, while the target is the output obtained from the physics-based algorithms. DNN-based methods involve two phases: 1. Training: learning the mapping function (or DNN model) from known inputs to outputs (targets), i.e., supervised learning; 2. Inference: mapping the input to an unknown output with the trained model.

Unlike supervised learning, unsupervised learning only requires the input data without outputs (or target), such as principal component analysis (PCA), outlier detection, and clustering. Unsupervised learning is performed to learn about patterns in the data. It has wide applications in seismic data analysis and QC. Hou et al. (2019a) propose to use unsupervised machine learning technologies to help geophysicists analyze seismic data more efficiently without compromising detail. The same technologies can also be used for dispersion curve picking QC (Masclet et al., 2019), first break picking QC, and FWI cycle-skipping QC (Dinh et al, 2020).

Deep neural network for seismic processing

To illustrate the potential of using DNNs to process seismic data, we applied the U-Net (Ronneberger et al., 2015) architecture to the seismic deblending problem. Modern seismic data involves multiple sources firing with a high shooting rate, which causes the wavefield to overlap (blend) in the recorded data. Conventional physics-based deblending, the process of isolating the individual source wavefields, takes the source firing times, sequences, and physical signal behavior into account in order to separate and obtain cleaned shot records for each individual source. The training dataset was chosen from a sparse grid of sail lines. We validated the application of the DNN by comparing with results from physics-based deblending on validation data, which was a few kilometers away from the sail lines used in the training phase.

Figure 1 shows a comparison of the physics-based deblending and DNN deblending results. The U-Net model was first trained on a subset of physics-based deblended data and then applied to the validation data (Figures 1a and 1c). At first glance, the DNN model generates comparable results to the physics-based approach. Careful examination reveals low-frequency signal leakage of deeper events in the DNN result when we gain up the difference (physics-based result minus DNN result) by 20 dB (Figure 1d).

For a further test, instead of learning a single processing step, we tested the DNN capability for a fast-track processing workflow that combines the steps of deblending, denoise, and deghosting into one single DNN model. The level of non-linearity of the problem has dramatically increased, as the DNN model not only needs to remove coherent and random noise but also shape the seismic wavelets in deghosting. Figure 2 shows that the DNN is able to produce a comparable result to the physics-based approach, except that there is clear signal leakage after gaining up 20 dB, leading to a similar conclusion as in the previous case.

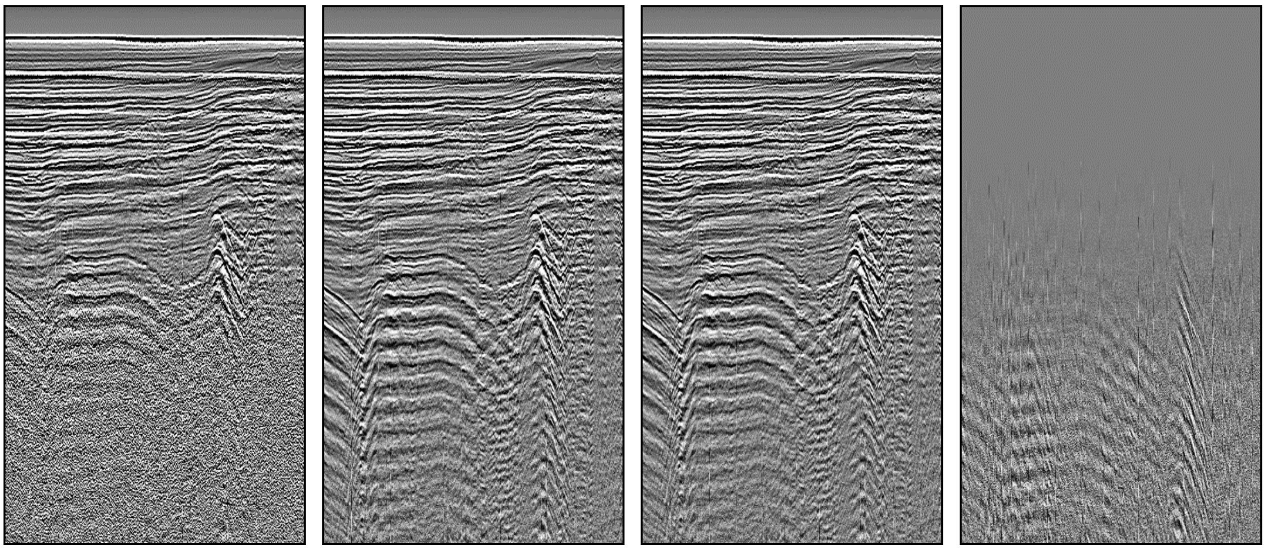
Both of these examples demonstrate the capability of DNN to learn most physics-based seismic processing methods. Next, we highlight a few common challenges to be addressed before bringing the DNN tool into production.

1. Signal fidelity preservation (Hou and Hoeber, 2020): DNNs for computer vision are used to find locations (for object detection) and object boundaries (for semantic segmentation), and determine classes mainly based on shapes and colors. The results are less sensitive to the pixel-level details from the inputs, and the decision is less impacted by minor imperfections. In contrast, seismic processing requires a high standard of signal fidelity as there are physical meanings for each wiggle that require in-depth analysis before making an impactful business decision. This means that the outputs are sensitive to not only the general shapes but also the signal amplitude and phase. Weak events in seismic data (e.g., diffractions, poorly illuminated targets) are normally much more important than the

10.1190/segam2021-3590137.1

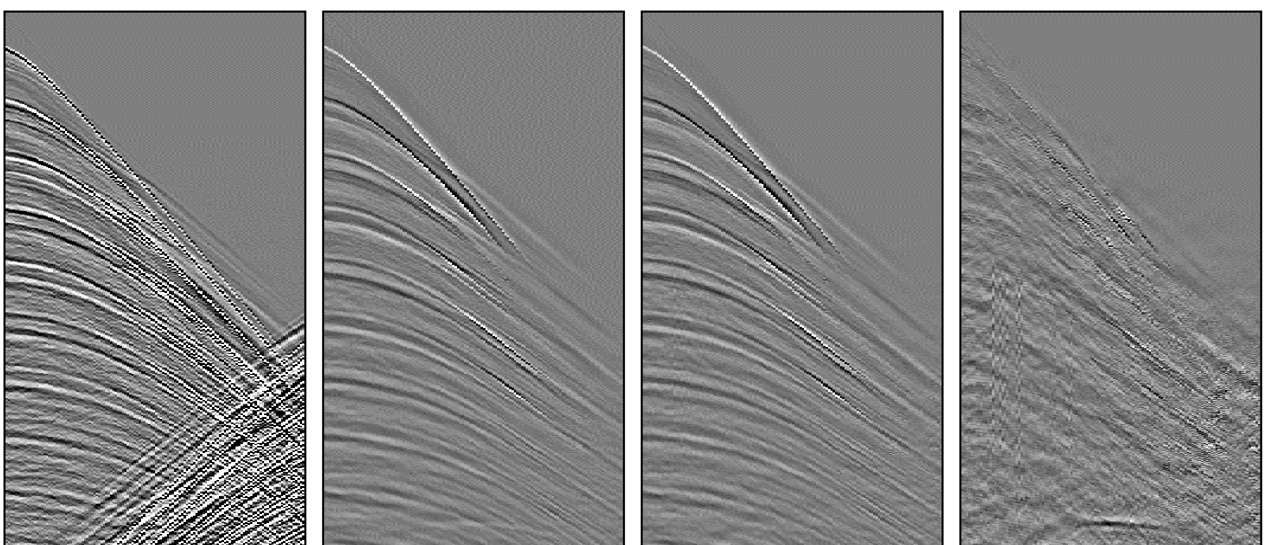
strong events. Non-stationarity of the amplitudes and wavelet (due to geological, divergence, and attenuation effects) add further complications that indicate even the best off-the-shelf DNN models from computer vision need to be further optimized for seismic data processing.

2. Training data set generation and selection: The choice of the training dataset is as important as or even more important than the choice of the DNN architecture. The accuracy in the training dataset affects the general accuracy of the model, while the statistical distribution



(a) Input (b) Physics-based deblending (c) DNN deblending (d) Difference (+20 dB)

Figure 1 Common channel gathers from field data: (a) input to deblending; (b) physics-based deblending of (a); (c) DNN deblending of (a); (d) difference of (b) minus (c) with 20 dB gain of amplitude.



(a) Input (b) Physics-based pre-processing (c) DNN pre-processing (d) Difference (+20 dB)

Figure 2 Shot records from field data: (a) Input to pre-processing; (b) physics-based pre-processing; (c) DNN pre-processing of (a); (d) difference of (b) minus (c) with 20 dB gain.

10.1190/segam2021-3590137.1

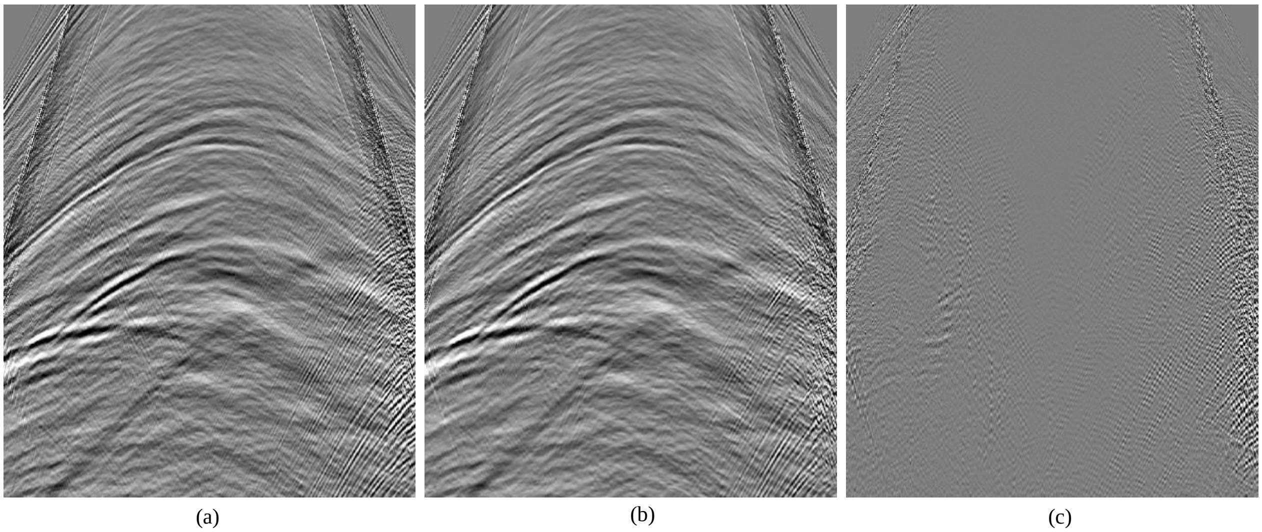


Figure 3 Synthetic shot gathers (a) from forward modeling at 12.5 m interval; (b) from forward modeling at 37.5 m interval followed by DNN interpolation from 37.5 m to 12.5 m interval; (c) difference between (a) and (b).

similarity between the training dataset and the full production dataset affects the performance of the DNN model. For example, if the target of the training dataset contains primary leakage, the output of the DNN model is expected to have a similar level of primary leakage as well. One can use synthetic data as the training dataset to have the best accuracy (i.e., with no primary leakage as the exact processing is known), but the model performance on field data will be heavily compromised due to the "domain shift" problem. Thus, to obtain the best DNN result in seismic processing, it is necessary to continue evolving the conventional processing toolbox in order to improve the training dataset quality, and have robust training data selection logic to ensure the model performance on real data is as accurate as possible.

3. Balancing the local and global performance: Physics-based algorithms often utilize iterative optimization for each spatial and/or temporal window of seismic data, which leads to the theoretically optimal solution for each input. However, these algorithms might suffer from unexpected noise, geological variation, or other edge-cases. On the contrary, DNN approximates a function for the whole dataset, here called the global optimal, and is more robust to data containing unexpected noise and artifacts. This is considered a primary advantage of DNN over conventional methods (Hou et al., 2019b). But this "global optimal performance" might imply a critical issue: it tends to compromise the quality of areas with rare features or abnormal geology, which are likely to be the areas of interest. As a result, it is necessary to develop proper metrics to evaluate the model with respect to the application dataset to ensure it captures authentic local geological variation found in the application dataset without overfitting the noise.

Although the accuracy of DNNs has not yet met the production standards of existing high-end seismic processing toolboxes, there are specific applications where the trade-off between DNN quality and computation efficiency can be exploited. For example, guided denoise requires generation of clean synthetic data and uses the kinematics (instead of amplitude) to guide the noise removal algorithms. However, it is computationally intensive to model the full synthetic data. Figure 3b shows an example of using a DNN to interpolate the synthetic shot gather from a sparse trace interval of 37.5 m to a dense trace interval of 12.5 m. It is comparable to the dense synthetic shot gather (Figure 3a) obtained purely via forward modeling. With the help of DNN, we save 2/3 of the runtime for synthetic modeling (at 37.5 m rather than 12.5 m), while the DNN runtime for interpolation is negligible by comparison.

Reduce human efforts with machine learning

For seismic processing technology, the goal is not only to achieve higher accuracy but also to reduce project turnaround time. Beneath the iceberg, the most time consuming steps during the processing flow are often those that require human intervention, e.g., first break picking, dispersion curve picking, QC. These processes rely mostly on visual perception and have less physics behind them. Although some rule-based methods were developed to automate these tasks, there are still significant human efforts involved. With modern high-density seismic surveys, where we collect tens of billions of traces, it becomes even more important to seek out efficiency gains and to reduce time spent on the more mundane processing aspects, in order to leave more time to focus on the critical aspects of the analysis.

Moreover, for some specific tasks, supervised ML, like DNN, can produce more robust solutions than manual picking and simple rule-based algorithms. This means less human intervention. Figure 4 shows an example of 3C ocean bottom node residual orientation correction. In a conventional rule-based node orientation workflow, the algorithm iteratively inverts the rotation matrix to minimize the mean absolute amplitude of the direct arrival in the transverse component. It does not inspect the polarity and patterns of the direct arrival amplitude on other components. As a result, it can sometimes be trapped into the local minima, which completely flips the input data polarity, as shown in Figure 4b. The geophysicist has to spend time to carefully QC the result to flag this problem. On the other hand, the DNN (Figure 4c) analyzes the amplitude maps of all components simultaneously and inverts for a more reliable rotation matrix. It is more robust to noise and does not suffer from the issue of reverse polarity, thus requiring less human intervention. Similarly, the DNN can be applied for other tasks, such as repositioning and clock drift detection.

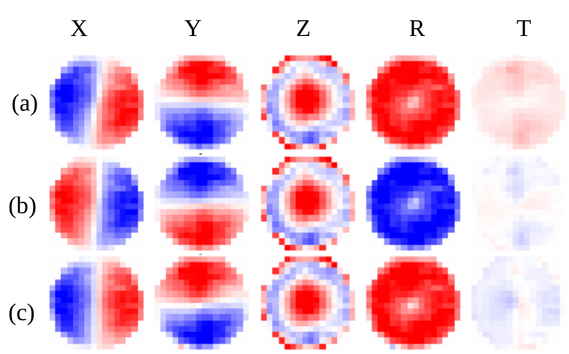


Figure 4. Amplitude maps along direct arrival of x-component (X), y-component (Y), vertical component (Z), radial component (R), and transverse component (T): (a) raw data with orientation error; (b) corrected via the rule-based algorithm; (c) corrected via DNN.

Discussion and Conclusions

Seismic processing, unlike regular computer vision routine tasks, requires very high standards for the preservation and extraction of signal fidelity, due to (1) most of the signals of our targets are extremely weak in the raw data; (2) any damage caused by one processing step may be amplified in later steps due to the non-linear nature of inversion algorithms; and (3) final processed results are quantitatively analyzed by multi-disciplinary teams to support impactful business decisions.

We have discussed various challenges and limitations associated with the application of DNN and ML to seismic processing. After the initial rush of excitement, we are now in a phase of realizing the limitations and inherent uncertainties with ML algorithms in seismic processing. First of all, DNN can be trained as an efficient approximation of processing algorithms, but the accuracy relies on the quality of the training dataset. As the training datasets are usually obtained via conventional processing flows, the physics-based algorithms still play a central role in seismic processing and should continue to be further developed. Secondly, even the best off-the-shelf DNNs need to be further optimized for seismic data processing in order to fulfill the high standard in signal fidelity (Messud and Chambefort, 2020). There are also other practical issues that need to be further addressed, such as training data selection, uncertainty qualification, and model interpretability. In our view, ML will not replace human expertise or geophysical algorithms, but it can augment human and geophysical algorithms to achieve more robust and efficient solutions, together.

Moreover, with the rapidly growing trace counts in newly-acquired seismic datasets, it becomes more and more challenging for geophysicists to comprehensively inspect and analyze all of the data. ML is able to leverage the "big data" effect and helps geophysicists perform large-scale data analysis more efficiently. Moreover, the example of node orientation QC demonstrates that DNN could produce a more robust solution than that of simple rule-based methods.

In summary, ML is clearly an exciting, fast-moving emerging technology. Further research and development is needed to address the key challenges before it can become a routine part of the seismic processing toolbox.

Acknowledgments

The authors thank CGG for permission to publish this work. We also thank CGG MCNV for providing the data from multiple projects. Special thanks to Chu-Ong Ting, Gordon Poole, Ricardo Diaz, Ewa Kaszycka, and our colleagues in the CGG AI Lab and Subsurface Imaging teams for discussions and suggestions.

REFERENCES

Dinh, H., G Cattini, P. Fallon, F. Hall, A. Holman, and S. Hollingworth, 2020, Machine learning applied to drive large scale FWI in the Central North Sea: Presented at Machine Learning in Geoscience Workshop, PESGB.

Hou, S., S. Angio, A. Clowes, I. Mikhalev, H. Hoeber, and S. Hagedorn, 2019, Learn to invert: surface wave inversion with deep neural network: 81st Conference and Exhibition Workshop Programme, EAGE, Extended Abstracts, WS10 02, doi: https://doi.org/10.3997/2214-4609.201901965.

Hou, S., S. Angio, H. Hoeber, V. Massart, L. Peng, R. Johnston, and R. Walters, 2019, Application of unsupervised machine learning to the processing of a land mega-survey: Subsurface Intelligence Workshop, EAGE, Extended Abstracts, SSI01, doi: https://doi.org/doi: 10.3997/2214-4609.2019X6104.

Hou, S., and H. Hoeber, 2020, Seismic processing with deep convolutional neural networks: Opportunities and challenges: 2020 Annual Conference & Exhibition Online, EAGE, Extended Abstracts, doi: https://doi.org/10.3997/2214-4609.202010647.

Masclet, S., T. Bardainne, V. Massart, and H. Prigent, 2019, Near surface characterization in southern Oman: multi- wave inversion guided by machine learning: 81st Conference and Exhibition, EAGE, Extended Abstracts, Tu R16 05, doi: https://doi.org/10.3997/2214-4609.201900968.

Messud, J., and M. Chambefort, 2020, Understanding how a deep neural network architecture choice can be related to a seismic processing task: First Digitalization Conference and Exhibition, EAGE, Extended Abstracts, WeDIGI09, doi: https://doi.org/10.3997/2214-4609.202032076.

Ronneberger, O., P. Fischer, and T. Brox, 2015, U-Net: Convolutional networks for biomedical image segmentation: International Conference on Medical Image Computing and Computer-Assisted Intervention, 234–241.