Framework and standalone applications of machine learning in seismic processing

Tony Martin, Bagher Farmani, Morten Pedersen and Elena Klochikhina, PGS

Summary

Machine learning, and the variations thereof, have been around for a considerable time, perhaps as far back as the early 19th century with Bayes' work on probability theory. It is then a bit of a surprise that these techniques have not made more headway in seismic processing, a data rich industry that should naturally suit greater automation using machine learning algorithms.

Mathematical functions used in seismic processing are highly evolved and designed to improve the data quality by removing noise, correctly positioning data, or enhancing the data quality. The decade's long development of these applications means they are extremely good at solving the challenges each one individually addresses.

So how can machine learning help when the industry already has highly evolved and effective tools? There are two possible categories for applications, one to supplement the existing functions in a framework that enables greater autonomy, and the second to replace tools that are not as effective as they could be, allowing a greater diversity of testing-free applications (generalization). In both cases, the goal is improved data quality, faster. We present examples that might benefit seismic processing, as well as comment on the challenges faced with these methods.

Introduction

Machine learning (ML) applications in seismic processing are an ever increasing research investment, yet very few are being actively used in commercial seismic processing projects. An indication of the increasing research in machine learning can be seen in Figure 1, which plots the number of submissions of papers linked to machine learning against all submissions, for the Society of Exploration Geophysicists' GEOPHYSICS journal. Total submissions have increased by 2.4 times since 2006, whilst those specifically targeting ML (including papers mentioning machine learning, artificial intelligence, data analytics or automation) have grown by more than 23 times, and now represent 15% of all submissions.

Early ML work in seismic processing targeted quality control (QC) tasks like impulsive noise attenuation, classifying noise, signal and leakage, prior to defining a 3D space suitable for unsupervised clustering to define whether the underlying mathematical algorithm had been successful in its task (Spanos and Bekara, 2013; Bekara, 2014). These processes evolved to incorporate other techniques, like principal component analysis, after significantly increasing the feature space, in an effort to try and further replicate what a geophysicist might do when checking if a noise attenuation process was successful (Martin et al., 2015).

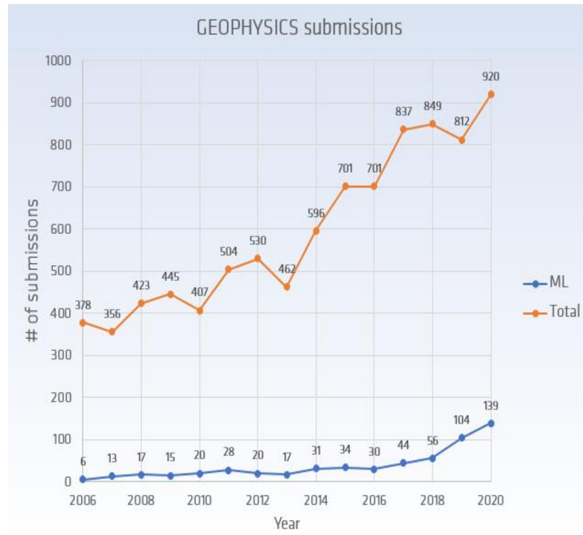

Figure 1: Year on year submissions to the Society of Exploration Geophysicists' GEOPHYSICS journal. Total submissions are mapped against the ever-growing number of those refering to machine learning, artificial intelligence, data analytics and automation.

There are numerous signal processing steps that impact data quality, and are heavily dependent on highly evolved mathematical processes, including data interpolation algorithms. There are several reasons why seismic data sampling might not be complete or optimal, from acquisition-related weather conditions to cost effective geometry designs. Some amelioration of these can be achieved by algorithms designed to reconstruct missing data, however, they have limitations and use assumptions, meaning results are sometimes sub-optimal. There have been several research projects that have developed ML mechanisms to solve the data reconstruction challenge, such as support-vector regression (SVR) (Jia and Ma, 2017). Unlike classification approaches, SVR is effective for continuous records and for transferring nonlinear regression problems with low dimensionality to linear problems in high dimensional space; mapping incomplete input data to fully populated output data, based on a volume of training points, so that different data may be interpolated using the trained model. Other applications of ML to data interpolation include dictionary learning (Turquais et al., 2019) and more

advanced deep convolutional neural network methods (Jaikla et al., 2021).
The usability of seismic data is heavily dependent on the accurate positioning of energy, through data migration. Building the velocity model is a time-consuming step in a processing project. Most models are derived using inverse systems that are under-determined. This may lead to uncertainty in the imaging. Additionally, many of the processes are highly nonlinear, again leading to model and imaging ambiguity. Accelerated and automated model building can be achieved in a stochastic framework with feedback, a crude system built around existing tools (Martin and Bell, 2019), however, this kind of approach still relies on the same underlying assumptions of conventional methods. A number of examples of neural network driven approaches, where efficacy and transferability is dependent on the training regime and hyperparameters, can be found in the literature (Øye and Dahl, 2019; Yang and Ma, 2019; Zheng et al., 2019)

In the following Examples section we outline two applications of ML to noise attenuation; the first is a innovative framework around an existing algorithm that utilizes both supervised and reinforcement learning to process at the ensemble level, the second uses a deep convolutional neural network for removing migration related coherent noise, and demonstrate the models ability to effectively generalize to different data sets.

Example one – the framework approach

The first example uses an automatic attribute driven image classification system to perform a labelling of data for an impulsive noise attenuation process; samples are labelled as signal, noise or signal and noise (Farmani and Pedersen, 2020). For the latter classification, there are two scenarios:

1) 'denoised' data with residual noise, or
2) Noise estimate with signal leakage

The image segmentation classification used a U-Net architecture (Ronneberger et al., 2015), with input tiles of 336 by 336 samples. The model component has 21 by 21 samples in the lowest resolution layer, where the first encoder had 16 filters. The supervised training was conducted using labelled tiles with over 5000 used for each classified class. Tiles were split, with 70% being used for training, whilst verification used 20% and 10% was used for testing.

The mixed classification approach (fourth signal and noise category) improves the determination of leakage, whether signal or noise, and is demonstrated in Figure 2, where the results of three increasingly harsh applications of denoise are input to the classifcation, enabling a more accurate measure of the denoising effectiveness. Equally importantly, it permits the inclusion, within the framework, of a reinforcement loop (Farmani and Pedersen, 2020), initially determining the successfulness of the application/parameterization, and subsequently followed by an automated re-run with automatic parameter modification, on a gather by gather basis, to optimize denoise performance. This kind of framework, built around an existing signal processing tool, is well suited to the use of a data analytics approach to initial parameter estimation (Martin et al., 2020), and may enable ever more automation, by bypassing parameter testing, therefore accelerating turnaround.

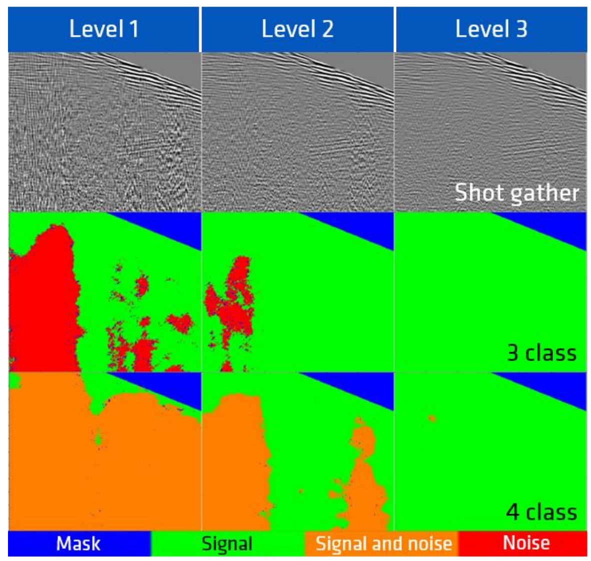

Figure 2: Four-way automated image classification system, defining whether energy in the data is masked, signal, noise or signal and noise. The additional class improves the classification and permits the reinforcement automation.

The inclusion of the reinforcement system, where classification leads to a re-run of the underlying algorithm to high-grade the denoise performance at the ensemble level, is demonstrated in Figure 3 using a common channel data set. The 'Production denoise' panel show a single pass denoise result which is fed into the ML framework. The 'Automated denoise' panel demonstrates the high grading of the data quality following classification and automatic reparameterization and rerunning, the difference of which is shown in the 'Automation benefit' panel.

Example two – the algorithm replacement approach

In the second example coherent noise in the image domain is attenuated using a deep convolutional neural network (CNN). Coherent noise is challenging to remove as it can often have characteristics similar to the underlying signal we want to preserve, and consequently processing algorithms based on seismic characteristics or patterns, often categorize the noise as signal and therefore do not remove it, or at least not very effectively. Developing a CNN tool to improve the efficiency of the denoising is a highly desirable goal. The architecture used in the following example contains contraction, bottleneck, and expansion branches as part of a U-Net architecture. Training was performed on approximately 100 000 samples, augmented with flips, crops, scaling and filtering. The input consisted of tiles of migrated data that had been subsampled on input resulting in migration related coherent noise, whilst the idealized output was optimally sampled for migration. Hyperparameters were adjusted during the training phase to minimize the prediction error (Klochikhina et al., 2020). Figure 4 shows the results of an application of the trained network, for a deep water data example from offshore Brazil. Orange arrows in the upper panel highlight the offending noise, whilst blue arrows in the bottom panel show how the CNN has removed the noise that is still present when a conventional approach is used (turquoise arrows in the middle panel).

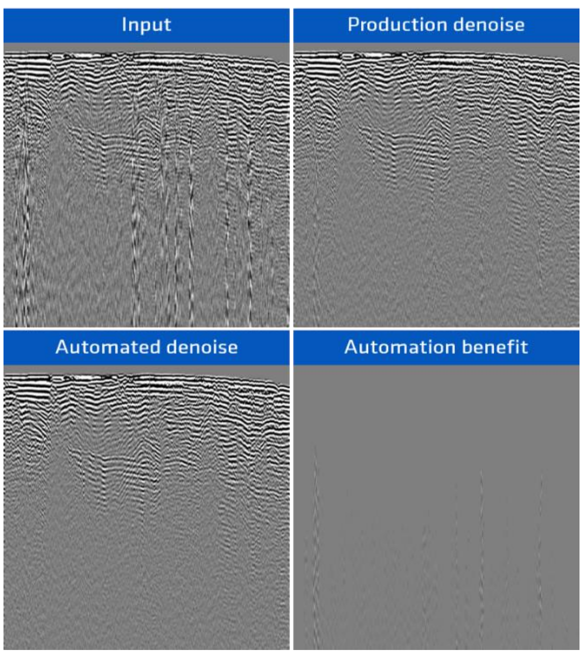

Figure 3: Images showing the improvement in denoise performance through use of a machine learning framework that includes automated image classification and reinforcement enabled feedback. The top row shows a traditional ‘single survey parameter set’ application, whilst the bottom left panel demonstrates the machine learning result, which had better signal preservation and noise attenuation (bottom right panel).

In this first example the model generalized well from the initial training based on synthetic data, the challenge of over-fitting had been averted. This is primarily due to the variability of the training data used, facilitated through the augmentation process; the frequency, amplitude and noise level of the real field data set example are accommodated in the training data set.

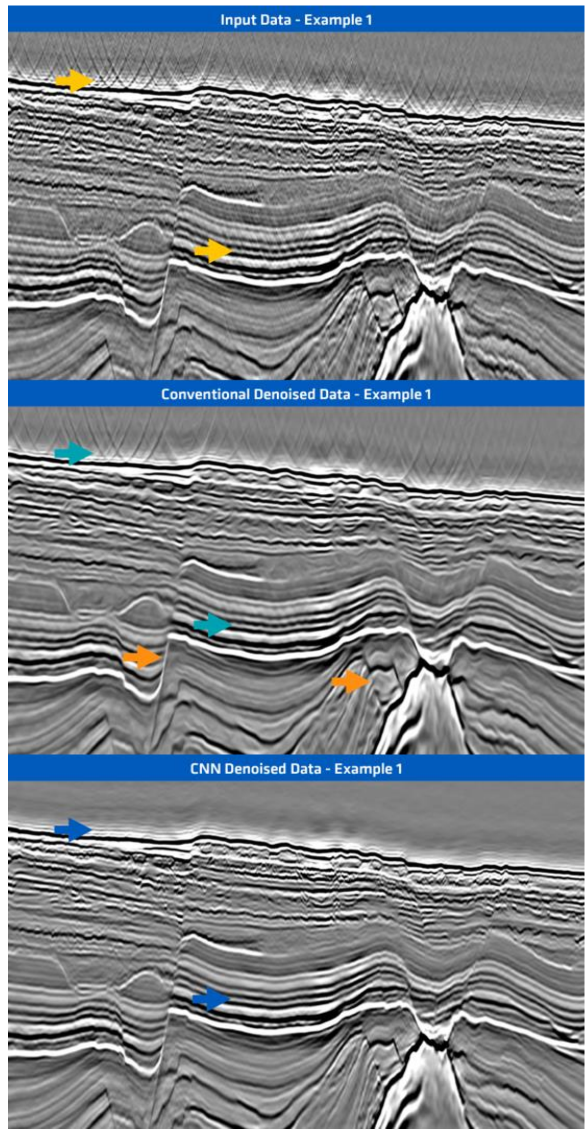

Figure 4: Offshore Brazil example of the application of a trained CNN to remove coherent image domain noise. The noise (top panel orange arrows) is attenuated by the CNN application (blue arrows bottom panel). The result is superior to a conventional apprach (middle panel).

This is also confirmed by an application to a different field seismic data set, a shallow water example from the North Sea (Figure 5), where the seismic and noise characteristics are very different to the first example from offshore Brazil. The upper panel highlights the noise caused by suboptimal destructive interference in the migration process for a deep high impedance contrast (orange arrows). The bottom panel shows how the application of the trained CNN attenuates this noise (blue arrows) with minimal impact on the desirable reflectivity, despite it having similar seismic 'characteristics' as the noise.

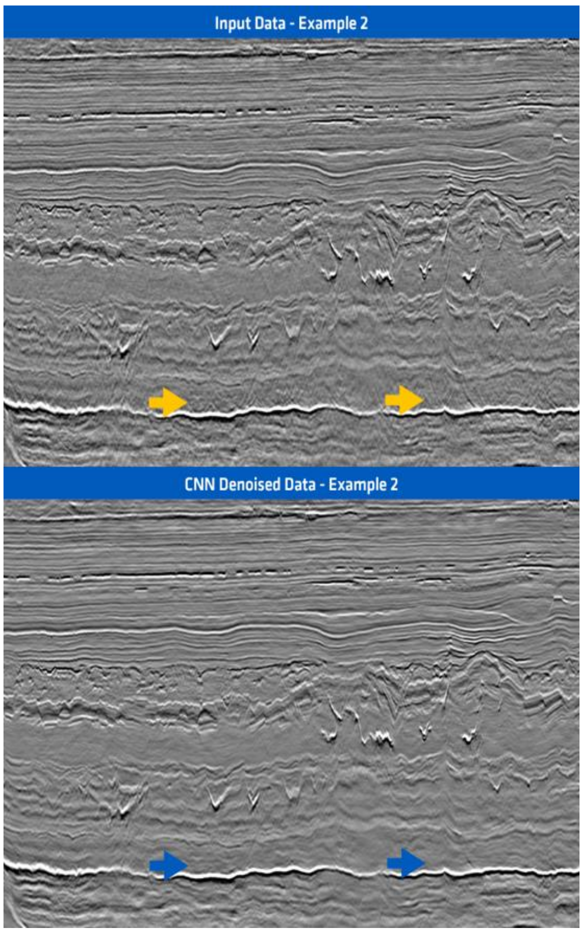

Figure 5: North Sea example of the application of a trained CNN to remove coherent image domain noise. The noise (top panel orange arrows) is attenuated by the CNN application (blue arrows bottom panel). The result shows that the network generalizes in an effective way from the initial training.

Discussion and conclusions

With success stories like those outlined and referenced here, why is it that there are not more examples of ML being actively used in seismic processing? It may be that the current and historical applications, whose continual development has spanned years of effort, are perpetually optimized and highly effective; it is easy to cherry-pick some good ML application examples, but are they consistently better—do the results meet the hype? It could be that effective multi-dimensional training is a significant undertaking, and that hyperparameter testing is still a form of testing, so will these approaches ever really help reduce turnaround? Even though seismic data processing experts are at the pinnacle of data science, there might be a perception that different expertise is required in research divisions to properly develop ML tools? Conceivably, integrating ML libraries and codes within existing seismic processing platforms might not be as easy as anticipated, and new tools are constantly evolving; which does one use? Could it be that there are cultural issues with ML driven automation, or maybe even a lack of financial imperative? Perhaps it always takes this long, and our collective memories of the time it takes from conception to delivery have been warped. Whatever the reason, ML development in seismic processing is slow, but steady.

Summary

Machine learning is an increasingly popular area of seismic processing research, although little evidence exists of day-to-day use. We have presented two examples of ML methods to help in data processing. The first, a framework built around an existing mathematical algorithm, enables automated classification of the process' effectiveness, followed by automated re-running enabled through reinforcement, with high-graded parameters. This type of approach could reduce turnaround, by eliminating laborious QC and parameter testing, and result in an improved denoising performance. The second, a standalone application, and direct replacement of an existing denoise tool, demonstrates how ML methods may supplant existing tools where they are not as effective as they might be.

Acknowledgements

The authors would like to thank PGS for permission to present the results, and for the data examples used in this paper.

REFERENCES

Bekara, M., 2014, Toward an automatic swell noise attenuation process: 76th EAGE conference and Exhibition, Th ELI1 04, doi: https://doi.org/10.3997/2214-4609.20141442.

Farmani, B., and M. Pedersen, 2020, Application of a convolutional neural network to classification of swell noise attenuation: 90th SEG Annual Meeting, Extended Abstracts, 2868–2871, doi: https://doi.org/10.1190/segam2020-3425046.1

Farmani, B., and M. Pedersen, 2020, Extended attributes for machine learning denoise process: First step towards automation: 82nd EAGE Conference and Exhibition, Extended Abstracts.

Jaikla, C., E. Alkan, C. Sutton, Y. Cai, S. Mannava, A. Gala, P. Devarakota, D. Knott, L. Chernis, H. Badat, G. Madiba, D. Segonds and D. Hohl, 2021, Deep learning techniques revolutionize E&P – Two practical applications: 1st EAGE Workshop on Optimizing Turnaround and Project Performance, Tu07, doi: https://doi.org/10.3997/2214-4609.202130009.

Jia, Y., and J. Ma, 2017, What can machine learning do for seismic processing? An interpolation application: Geophysics, **82**, no. 3, V163–V177, doi: https://doi.org/10.1190/geo2016-0300.1.

Klochikhina, E., S. Crawley, S. Frolov, N. Chemingui, and T. Martin, 2020, Leveraging deep learning for seismic image denoising: First Break, **38**, 41–48, doi: https://doi.org/10.3997/1365-2397.fb2020048.

Martin, T., and M. Bell, 2019, An innovative approach to automation for velocity model building: First Break, **37**, 57–65, doi: https://doi.org/10.3997/1365-2397.n0033.

Martin, T., A. Long, S. Butt, and K. Baharom, 2020, Seismic processing parameter mining – The past maybe the key to the present: First Break, **38**, 55–58, doi: https://doi.org/10.3997/1365-2397.fb2020043.

Martin, T, C. Saturni, and P. Ashby, 2015, Using machine learning to produce a global automated quantitative QC for noise attenuation: 85th SEG Annual Meeting, Extended Abstracts, 4790–4794, doi: https://doi.org/10.1190/segam2015-5826237.1.

Øye, O. K., and E. K. Dahl, 2019, Velocity model building from raw shot gathers using machine learning: 81st EAGE Conference and Exhibition, Extended Abstracts, doi: https://doi.org/10.3997/2214-4609.201900039.

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.

Spanos, A., and M. Bekara, 2013, Using statistical techniques to improve the QC process of swell noise filtering: 75th EAGE conference and Exhibition, We 14 13, doi: https://doi.org/10.3997/2214-4609.20130884.

Turquais, P., W. Söllner, and M. Pedersen, 2019, Parabolic dictionary learning: A method for seismic data reconstruction beyond the linearity assumption: 81st EAGE conference and Exhibition, We_R04_13, doi: https://doi.org/10.3997/2214-4609.201901190.

Yang, F., and J. Ma, 2019, Deep-learning inversion: A next-generation seismic velocity model building method: Geophysics, **84**, no. 4, R583–R599, doi: https://doi.org/10.1190/geo2018-0249.1.

Zheng, Y., Q. Zhang, A. Yusifov, and Y. Shi, 2019, Applications of supervised deep learning for seismic interpretation and inversion: The Leading Edge, **38**, 526–533, doi: https://doi.org/10.1190/tle38070526.1.