geoconvention
Calgary • Canada • May 13-17 2019

A comparison of Machine Learning methods for seismic inversion to estimate velocity and density

Yexin Liu
SoftMirrors Ltd., Calgary, Canada; yexinliu@softmirrors.com

Summary

Seismic inversion is a method that infers the reservoir rock properties, such as density, impedance and porosity, from seismic data. However because of seismic band-limited, other information, such as well logs and geological data, are needed for the inversion. How to incorporate all these information is crucial for any inversion methods. In this study, we will use advanced Machine Learning methods to incorporate the seismic data, well logs and geological data to propagate the reservoir rock properties from well locations to any seismic space. In order to compare the Machine Learning methods for seismic inversion, we have studied Deep Convolutional Neural Networks (CNN), Bayesian-based Support Vector Machine (BSVM) and the joint methods of CNN and BSVM. All these three methods we utilized for seismic inversion are direct, one-step inversion, which can incorporate of a wider range of seismic data, well logs, geological data and experience to infer the reservoir rock properties instead of traditional two-step process of estimating impedance and then converting impedance to reservoir properties.

In our case study we trained the Machine Learning methods using a real dataset which consists of 4641 density and 4646 velocity logging samples, which extracted from the wireline loggings and associated seismic AVO attributes and then used another 131 new samples for testing. The BSVM method trains and infers the rock properties sample-by-sample, but the small images are needed to feed into CNN for the rock property inversion. The small image can be generated from seismic AVO attributes and other features. The joint methods of CNN and BSVM will take the advantages of CNN's automatic attributes extraction and then these attribute maps of fully connected layer will be fed into BSVM for training and prediction. Our results demonstrated that the BSVM can achieve best training and prediction results with small mean absolute errors and the joint methods of CNN and BSVM can outperform the CNN method. Mathematically we found BSVM can perform better for a small dataset and CNN needs big data to achieve better results. Our seismic inversion usually has limited samples from well logs, so BSVM can effetely maps the samples onto a high dimension space, in which no information will be lost, but CNN uses the convolutional operator to extract attributes and then pooling/dropout layer to avoid over-fitting, in which some information could be lost and then degrades the performance for the seismic inversion.

Bayesian-based Support Vector Machine (BSVM) method

Due to its excellent performance in dealing with sparse data and its good empirical performance, the Bayesian-based Support Vector Machine (BSVM) has been widely used in Artificial Intelligent areas, especially the Machine Learning community (Vapnik, 1992; Cortes and Vapnik, 1995; Wenzel, 2017). We (Liu and Sacchi, 2003, Liu, 2017 and Liu, 2018) presented the use of BSVM methods for seismic Quantitative Interpretation (QI) to incorporate seismic attributes, petrophysical properties and geological understanding and then to make the predictions of reservoir properties, such as porosity, density and pore pressure. The method should not only be geophysical consistent but also geological and petrophysical consistent for seismic inversion to estimate the reservoir rock properties.

Given a set of seismic attribute vectors $\{x_{n,}, n = 1, \ldots, N\}$ along with the corresponding petrophysical data $\{t_n, n = 1, \ldots, N\}$. The SVM makes the training and prediction of the petrophysical data based on a function of the form:

$$t(x_j) = \sum_{n=1}^{N} \omega_n K(x_j, x_n) + \omega_0 \qquad (1)$$

Where $\{\omega_n\ n = 1, \ldots N\}$ are the model weights which can be estimated after SVM training, and $\{K(x_j, x_n\}$ denotes the kernel functions (Schölkopf, 1999).

BSVMs are known to achieve the best performance for regression problems, but it is restricted to rather small dataset. We use the method of inducing point Gaussian Process (Wenzel et al, 2017) to extend BSVMs to practically handle large real world datasets.

Deep Convolutional Neural Network (CNN)

BSVM becomes less practical to handle the big data, but Deep CNN has the capability to address the issues using a convolution operator to generate the feature maps of the thousands of labeled samples, which are more efficient for representing the space distribution of these samples. Figure 1 is our CNN workflow for seismic inversion. In this workflow, the input includes seismic images generated from seismic AVO attributes and their associated well logs around well trajectories. Unlike most traditional Machine Learning algorithms, Deep Learning networks, such as CNN, perform automatic feature extraction without human intervention. Because the deep CNN is prone to over-fitting, one often inserts pooling layers between convolution layers to reduce the feature maps, and the dropout techniques to avoid the over-fitting. In the fully connected layers the linear regression will be applied to make the prediction of the rock physical properties, such as density, velocity and porosity.

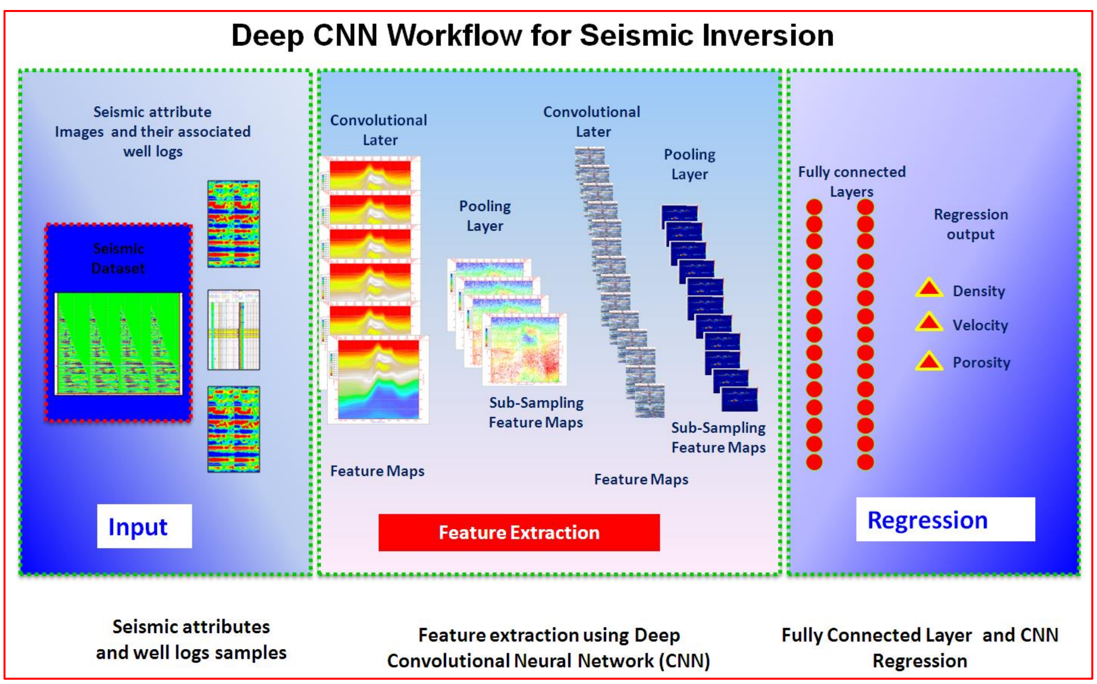


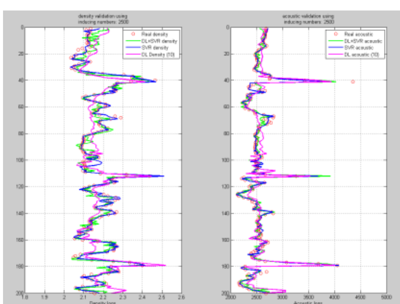
Figure1: Deep Convolutional Neural Network (CNN) workflow for seismic inversion

Joint method of CNN and BSVM

In our study, we found the BSVM can achieve good performance with very small absolute average errors for seismic inversion, but BSVM becomes less practical to handle five thousands or more samples. The CNN has the advantages of handling the big data containing millions of samples and of automatic feature extractions, but it is still a challenge to reduce the average errors between the predicted well logs and real well logs to meet the requirement for seismic inversion (Das et al, 2018). We are inspired to combine the two Machine Learning methods (Tang, 2013) for seismic inversion, in which we can use CNN for automatic attribute extraction and BSVM for better performance for regression. We propose the joint methods of CNN and BSVM for seismic inversion. Firstly we used CNN methods for feature extraction and well log property prediction and then the CNN outputs the feature maps of last fully connected layer. The outputted feature maps from CNN are more efficient for representing the space distribution of the thousands of input samples. Secondly the BSVM will use the feature maps as its input to map them to some high dimensional space where it reveals the hidden feature using kernel functions for seismic inversion.

Examples

We used a real dataset with 4641 training samples and 131 test samples to compare our Machine Learning methods for density and 4646 samples for velocity inversion. The density and velocity training and testing samples were generated from real well logs and associated seismic AVO attributes along these well trajectories. Three Machine Learning methods, Bayesian-based SVM, Deep CNN and Joint methods of SVM and CNN, were trained using the 4641 samples and then calculated the training errors and predicted errors. Figure 2 (a) shows an example of three methods validation on the training dataset and Figure 2 (b) shows the example of three methods prediction on the testing dataset.


(a)

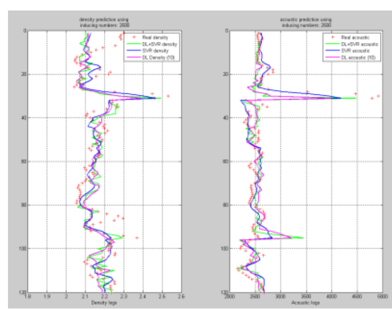

(b)

Figure 2 density and velocity prediction results from (a) training samples and (b) testing samples

The BSVM method maps the AVO attributes to some high dimensional space where it reveals the hidden features of these samples. It demonstrates good performance for the inversion. The average training errors with respect to the mean of density is about 0.86% and the average testing errors is about 2.5%. The AVO attributes were fed into CNN and then feature maps were automatically extracted using convolution operator. We used L2-norm to calculate the cost between real density and predicted density. The joint methods of CNN and BSVM will use the feature maps of CNN's fully connected layer and then use the SVM's kernel function to map these features onto a high dimensional space to build the model to predict the density and other reservoir properties. The training and prediction errors for the density and velocity inversion show in Table 1. In our study we found BSVM with kernel functions can achieve best training and prediction results for the density and velocity inversion, and the joint methods of CNN and

BSVM can improve the CNN prediction and training results. Furthermore, mathematically BSVM methods are effective for a small dataset for training and prediction and CNN has the capability of handling the big data containing millions of samples.

Table 1 density and velocity inversion errors from machine learning methods

Density and Velocity samples from real well logs and seismic AVO attributes	SVM with 2500 inducing points. Mean Percentage Errors (%)	Deep CNN Mean Percentage Errors (%)	Joint Methods Mean Percentage Errors (%)
4641 training samples (density)	0.86%	2.21%	1.28%
4646 training samples (velocity)	1.63%	4.03%	2.20%
131 testing samples (density)	2.50%	2.82%	2.88%
131 testing samples (velocity)	4.30%	5.29%	5.17%

Usually there are limited samples from well logs and associated seismic attributes for seismic inversion, the BSVM method with kernel functions easily performs better than the Deep CNN method, which needs millions of samples. Also we found the joint methods of CNN and BSVM can improve the CNN's prediction and training performance, but it is still a challenge to outperform the BSVM for seismic inversion with limited training samples. The reasons may be that CNN can capture a mearnful geological pattern from the input attributes using convolition operator, but some special information could be lost while feature maps were automactially extrated and other pooling/dropout techniques were applied to avoid over-fitting, which all will affect the CNN's performance. But in BSVM we used kernal functions to maps the input features and usually no any information will be lost.

Conclusions

Three Machine Learning methods, Bayesian-based SVM, Deep CNN and Joint methods of SVM and CNN, were compared for seismic inversion to estimate velocity and density. In the case study, we used a real dataset with 4641 training samples and 131 testing samples from real well logs and seismic AVO attributes for density inversion. The results of three Machine Learning methods demonstrated that BSVM method outperforms other two methods and achieves about 0.83% average percentage errors with respect to the mean of density from training samples and 2.5% from testing samples. The joint methods of CNN and BSVM show that the methods can effectively improve the results of CNN method. Mathematically the BSVM can achieve better performance in small dataset and the CNN can achieve better performance for big dataset containing millions of dataset. Fortunately in our seismic inversion study, there are limited samples from well logs and BSVM uses the kernel function to map the input attributes to high dimensions space, in which no information will be lost. However the CNN uses the conventional operator to extract the attributes and pooling/dropout layers to avoid the over-fitting, in which some special information could be lost and then degrades the performance. Also we propose the method of inducing point Gaussian Process to reduce BSVM training cost and make the methods practical for thousands of training samples.

References

Cortes, C. and Vapnik, V., 1995, Support-vector networks, Machine Learning, 20, 273-297

Das, V., Pollack, A., Wollner, U. and Mukerji, T., 2018, Convolutional Neural Network for Seismic Impedance Inversion, SEG International Exposition and 88th Annual Meeting

Liu, Y., 2018, Machine learning to enhance the vertical resolution of seismic geostatistical inversion**,** Geoconvention 2018, Canada

Liu, Y., 2017, Applications of Machine Learning for Seismic Quantitative Interpretation, Geoconvention 2017, Canada

Liu, Y., Sacchi, M., 2003, Propagation of borehole derived properties via a Support Vector Machine (SVM), CSEG Recorder, Canada, December 2003, 54-58

Schölkopf, B., Burges J., Smola, A., 1999, Advances in kernel methods: Support vector learning: MIT Press.

Tang, Y., 2013, Deep Learning using Linear Support Vector Machines reference style. International Conference on Machine Learning: Challenges in Representation Learning Workshop. Atlanta, Georgia, USA.

Vapnik, V.N., 1982, Estimation of Dependences Based on Empirical Data: Springer Series in Statistics, New York: Springer-Verlag.

Wenzel, F., Galy-Fajou, T., Deutsch, M., and Kloft, M., 2017, Bayesian nonlinear support vector machines for big data, in Proceedings of the European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases, 2017.