Stratigraphy estimation from seismic data using deep learning

Fantine Huot, Robert Clapp, Bruce Power, and Joseph Stefani

ABSTRACT

Seismic interpretation of deep stratigraphy is a challenging and time-consuming task. We present a modular and scalable data preparation pipeline and deep learning framework to assist with this task. Deep neural networks require large amounts of labeled data to train reliable models. Since the earth is intrinsically unlabeled, herein we generate synthetic data to create hundreds of thousands of labeled examples. We derive field data stratigraphy statistics from well logs from the Wilcox formation in the Gulf of Mexico and use Markov chains for data augmentation. We leverage these statistics to generate synthetic 3-D earth models and their corresponding seismic image volumes. We generate multiple seismic images using source wavelets of decreasing frequency. We train a deep neural network for image segmentation to estimate the stratigraphy from these seismic sections. We demonstrate that while the accuracy of the stratigraphy estimation decreases with the seismic data bandwidth, we can increase the accuracy at lower frequencies using transfer learning.

INTRODUCTION

Sand and shale have similar elastic properties, hence a similar seismic response, making them hard to distinguish from seismic data alone. Moreover, seismic images are limited in resolution by the wavelength of the seismic wave. The higher seismic frequencies are attenuated as the wave propagates through the Earth, and seismic images are usually much lower resolution than the fine stratigraphy. As a consequence, seismic interpretation of deep stratigraphy is a challenging and time-consuming task.

Herein we investigate whether we can leverage deep neural networks to help us interpret stratigraphy content from seismic data. We examine the constraints and limits on the seismic data quality and the granularity of the interpretation.

Within the last five years, deep learning has had a dramatic impact on computer vision (Krizhevsky et al., 2012; He et al., 2015), speech recognition (Deng et al., 2010; Seide et al., 2011; Dahl et al., 2012; Hinton et al., 2012), and image segmentation (Cireşan et al., 2012; Sermanet et al., 2013; Farabet et al., 2013; Couprie et al., 2013). However, deep neural networks rely on large amounts of labeled data to achieve their performance. A rough rule of thumb is that a supervised deep learning algorithm will

generally achieve acceptable performance with around 5,000 labeled examples per category and will match or exceed human performance when trained with a dataset containing at least 10 million labeled examples (Goodfellow et al., 2016).

The challenge in applying deep learning algorithms to seismic data is that the earth is intrinsically unlabelled. Seismic interpretation is often uncertain and the identified geobodies have fuzzy boundaries. As a consequence, there are very few ground truth labels in reflection seismology. To address this issue, we generate synthetic data to create labeled ground truth examples (Huot et al., 2018). We first derive field data statistics from wells logs from the Wilcox formation in the Gulf of Mexico. We use these statistics to populate Markov chains and create realistic synthetic 3-D earth models with similar geology. We generate the corresponding seismic image volumes with different source wavelet frequencies. We obtain seismic image and labeled stratigraphy pairings. We then train an image segmentation deep neural network to estimate stratigraphy from seismic sections with different seismic bandwidth. We demonstrate that we can increase the accuracy at lower frequencies using transfer learning.

1-D APPROACH: A WILCOX FIELD CASE STUDY

Well logs provide 1-D lithology and stratigraphy information. In this study, we use data from more than 60 wells from the Wilcox formation in the Gulf of Mexico, as shown in Figure 1. We label the logs by chunks into five stratigraphy categories, ranging from low net-to-gross to high net-to-gross:

1. Unimodal thick shale,
2. Low net-to-gross thin-bedded shale and sandstone,
3. Bimodal blocky shales and sandstones,
4. High net-to-gross thin-bedded sandstone and shale,
5. Very high net-to-gross blocky sandstone.

These five classes of stratigraphy are defined such as to cover the different types of stratigraphy encountered in the Wilcox formation. Note that classes 1 and 2 are mostly shale, while 3 and 4 are mostly sand. Classes 1 and 5 correspond to thick layers, while 2 and 4 are thin-bedded. Class 3 corresponds to channelized systems.

We augment the dataset using Markov chains. We create a macro-level Markov chain to define the transitions between the different categories of stratigraphy. Five fine-level Markov chains then define the transitions within each class of stratigraphy. This methodology allows us to create random vertical earth profiles with field data statistics.

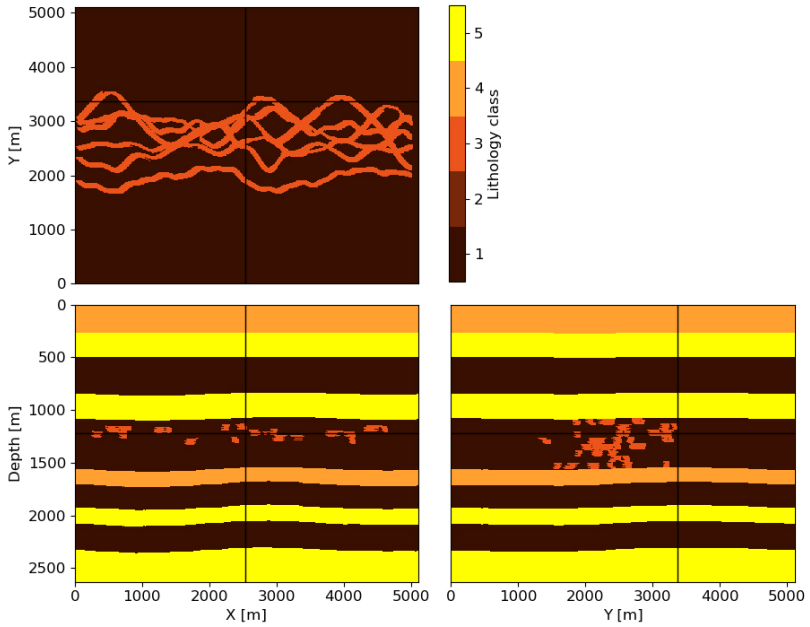
Figure 1: Some of the well locations from the Gulf of Mexico used in this study (courtesy of www.offshore-energy.biz). [**NR**]

However, the response of lithology in a sand-shale system is non-unique and highly unpredictable. Without the volume context in seismic data, accurate estimation is difficult, so we proceed to extending this study to 3 dimensions.

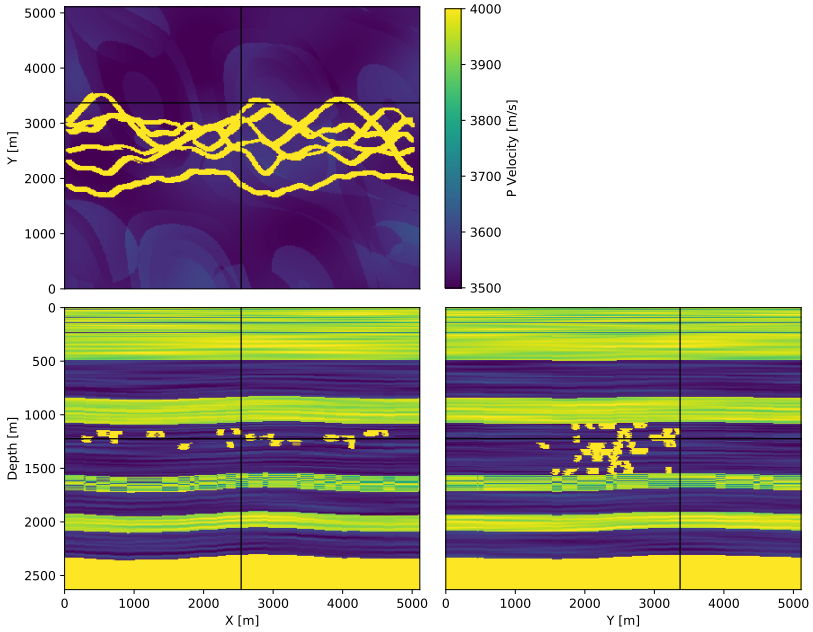
EXTENSION TO 3-D: SYNTHETIC DATA GENERATION

We use the aforementioned 1-D field data statistics to generate realistic synthetic multi-property 3-D earth models. We use a simplified basin modeling approach to create the earth volumes computationally efficiently. Our code is written in C++ using Threading Building Blocks and wrapped into a Python API using PyBind11. We use a Docker container for platform-independent deployment.

We use the macro-level Markov chain to define the transitions between the different classes of stratigraphy, and then use the five fine-level Markov chains to populate the various earth properties within each class. We use simplified deposition dynamics to create the stratigraphy and incorporate random channelized systems according to the Markov chain classes. We introduce random small-amplitude compression, deformation and faulting events. Figure 2 presents an example of a randomly generated 3-D earth model. We then generate multiple seismic images, using a 100 Hz, 80 Hz, 60 Hz and 40 Hz source Ricker wavelet. We extract random 2-D slices from the stratigraphy models and seismic images to generate hundreds of thousands of seismic sections with labeled stratigraphy classes.



(a)



(b)

Figure 2: An example of randomly-generated synthetic 3-D earth model using the described methodology. The 3-D volume is collapsed into three 2-D views, where each 2-D panel corresponds to the location of the black lines shown on the other 2-D panels. (a) The 3-D extension of the five classes of stratigraphy defined by the Markov-chain-generated vertical profile. (b) Same model populated with P-wave velocity values according to the field data statistics. [**CR**]

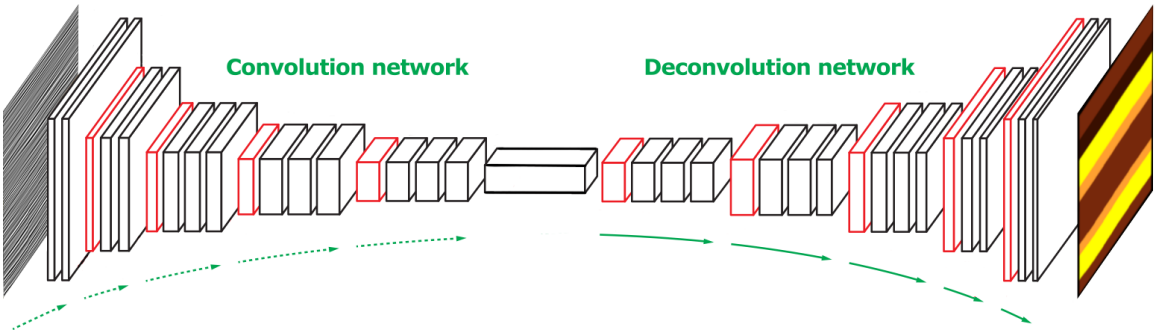


Figure 3: We use a convolution-deconvolution deep neural network to perform image segmentation to estimate the different classes of stratigraphy. It generates a feature representation through a series of convolution, pooling and rectification operations. Given a feature representation obtained from the convolution network, a dense stratigraphy type segmentation map is constructed through multiple series of unpooling, transposed convolution and rectification operations. [**NR**]

NEURAL NETWORK IMAGE SEGMENTATION

We train a convolution-deconvolution deep neural network (Noh et al., 2015) to estimate the different classes of stratigraphy from the seismic images. Figure 3 illustrates the overall configuration of the network. Our network is composed of two parts, a convolution and a deconvolution network. The convolution network acts as a feature extractor that transforms the input seismic image into a multidimensional feature representation, while the deconvolution network performs dense image segmentation from the feature representation extracted from the convolution network. The final output of the network is a probability map of the estimations of the different classes of stratigraphy.

All the convolution and transposed convolution layers have 3×3 kernels, while the pooling and unpooling layers have 2×2 kernels. We define our convolutional residual blocks as follows:

- 2-D convolution
- Batch normalization (Ioffe and Szegedy, 2015)
- Rectified linear unit (ReLU)
- 2-D convolution
- Batch normalization
- Addition of the input of the residual block to the output of the residual block

Our convolution network has an initial 2-D convolution and ReLU layer, followed by four convolutional residual blocks, each followed by a max-pooling layer.

Our deconvolution network is a mirrored version of the convolution network, with transposed convolutions instead of convolutions and unpooling layers instead of pooling layers. While the convolution network reduces the size of activations through feed-forwarding, the deconvolution network enlarges the activations through the combination of unpooling and transposed convolution operations.

We use a softmax with cross-entropy loss function, and eventually select the class with highest probability as label for each location of the input seismic image.

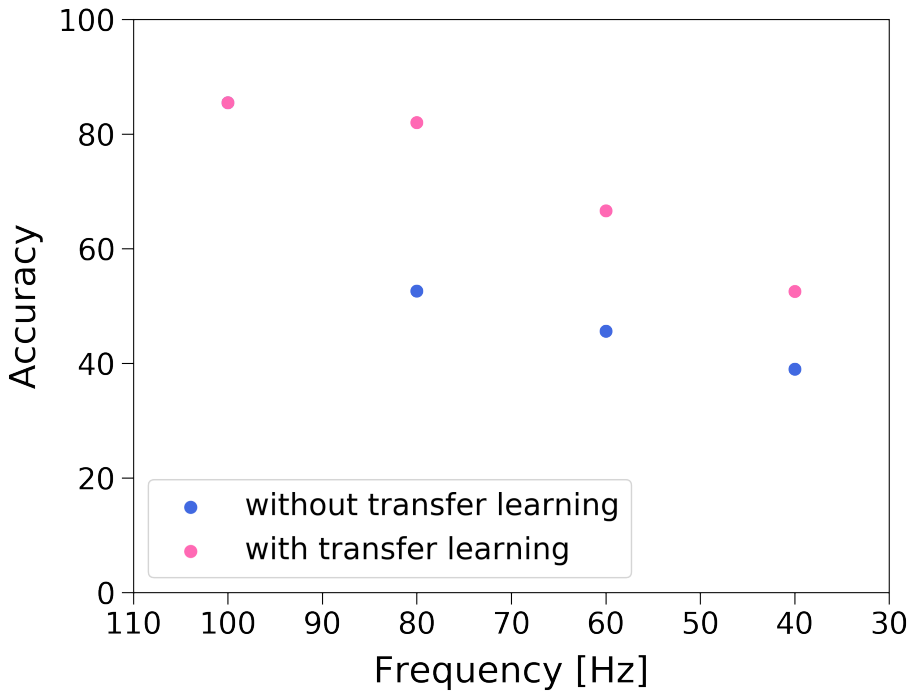


Figure 4: Segmentation accuracy on seismic data of varying bandwidth. While we obtain reasonable accuracy on high frequency seismic data, the neural network gets stuck in poor local minima when trained and tested on lower frequency data. We employ a two-stage training method to address this issue and pre-train the network with high frequency examples and then fine-tune the pre-trained network with lower frequency examples. [**NR**]

While we obtain reasonable accuracy on high frequency seismic data, the neural network seems to get stuck in poor local minima when trained and tested on lower frequency data as shown in Figure 4. We address this issue by transfer learning. We pre-train the network with high frequency examples and then fine-tune the pre-trained network with lower frequency examples.

From the preliminary segmentation results presented in Figure 5, we see that deep learning can provide some assistance in stratigraphy estimation from 2-D seismic data. The network performs well at distinguishing coarse bedding from fine bedding. The most common error comes from confusing sand and shale sections which are essentially just mirror images of each other from an impedance perspective. We obtain

high accuracy on channel detection, which is promising for geobody segmentation. Pushing the accuracy further on synthetic data might lead to overfitting, and we aim to transfer this problem to field data from the Wilcox formation.

CONCLUSIONS

We develop a modular and scalable data preparation pipeline and machine learning framework, leveraging synthetic data generation to create hundreds of thousands of labeled examples with field data statistics. We demonstrate that while the accuracy of the stratigraphy estimation decreases with the seismic data bandwidth, we can increase the accuracy at lower frequencies using transfer learning. Future steps of this project consist of transitioning to 3-D segmentation and applying the proposed methodology to field data in the Gulf of Mexico. Since the described workflow is entirely modular, it can be extended to lithology estimation or geobody detection for different geological contexts.

ACKNOWLEDGEMENTS

This work was conducted at Chevron Energy Technology Company in Houston, TX.

REFERENCES

Cireşan, D., U. Meier, J. Masci, and J. Schmidhuber, 2012, Multi-column deep neural network for traffic sign classification: Neural Networks, **32**, 333–338.

Couprie, C., C. Farabet, L. Najman, and Y. LeCun, 2013, Indoor semantic segmentation using depth information: arXiv preprint arXiv:1301.3572.

Dahl, G. E., D. Yu, L. Deng, and A. Acero, 2012, Context-dependent pre-trained deep neural networks for large-vocabulary speech recognition: IEEE Transactions on Audio, Speech, and Language Processing, **20**, 30–42.

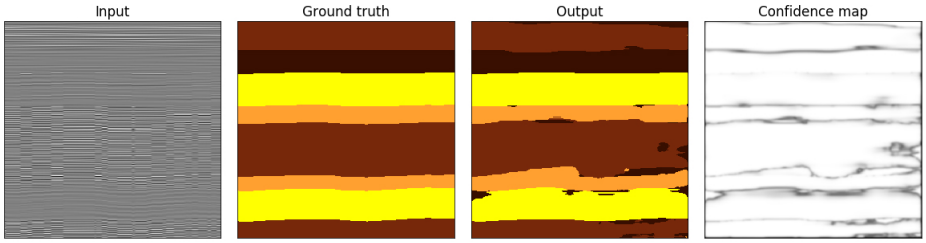
Deng, L., M. L. Seltzer, D. Yu, A. Acero, A.-R. Mohamed, and G. E. Hinton, 2010, Binary coding of speech spectrograms using a deep auto-encoder.: Interspeech, Citeseer, 1692–1695.

Farabet, C., C. Couprie, L. Najman, and Y. LeCun, 2013, Learning hierarchical features for scene labeling: IEEE transactions on pattern analysis and machine intelligence, **35**, 1915–1929.

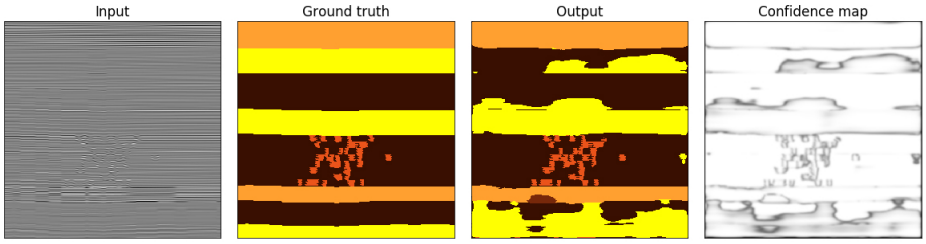
Goodfellow, I., Y. Bengio, and A. Courville, 2016, Deep learning: MIT Press.

He, K., X. Zhang, S. Ren, and J. Sun, 2015, Delving deep into rectifiers: Surpassing human-level performance on imagenet classification: Proceedings of the IEEE international conference on computer vision, 1026–1034.

Hinton, G., L. Deng, D. Yu, G. E. Dahl, A.-R. Mohamed, N. Jaitly, A. Senior, V. Vanhoucke, P. Nguyen, T. N. Sainath, et al., 2012, Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups: IEEE Signal Processing Magazine, **29**, 82–97.

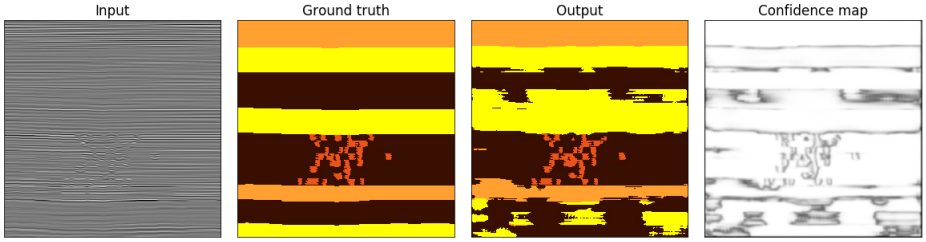


(a)



(b)

(c)



(d)

Figure 5: Preliminary segmentation results. From left to right, we show the seismic section given as input to the neural network, the ground truth stratigraphy classes, the estimated stratigraphy classes obtained as output of the network, and a confidence map that indicates which portions of the section where associated with high probabilities. White indicates high confidence, while black indicates low confidence. (a) and (b) correspond to seismic sections that were modeled using a 100Hz Ricker wavelet. (c) and (d) correspond to seismic sections that were modeled using a 60Hz Ricker wavelet. [**NR**]

Huot, F., B. Biondi, and G. Beroza, 2018, Jump-starting neural network training for seismic problems, *in* SEG Technical Program Expanded Abstracts 2018: Society of Exploration Geophysicists, 2191–2195.

Ioffe, S., and C. Szegedy, 2015, Batch normalization: Accelerating deep network training by reducing internal covariate shift: arXiv preprint arXiv:1502.03167.

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

Noh, H., S. Hong, and B. Han, 2015, Learning deconvolution network for semantic segmentation: Proceedings of the IEEE international conference on computer vision, 1520–1528.

Seide, F., G. Li, and D. Yu, 2011, Conversational speech transcription using context-dependent deep neural networks.: Interspeech, 437–440.

Sermanet, P., K. Kavukcuoglu, S. Chintala, and Y. LeCun, 2013, Pedestrian detection with unsupervised multi-stage feature learning: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 3626–3633.