SEISMIC INTERPRETATION WITH THE AID OF DEEP LEARNING

by

HAO WU

BO ZHANG, COMMITTEE CHAIR
NATASHA T. DIMOVA
YONG ZHANG
ZHE JIANG
ALAIN PLATTNER

A DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy
in the Department of Geological Sciences
in the Graduate School of
The University of Alabama

TUSCALOOSA, ALABAMA

2019

ABSTRACT

Nowadays one of the biggest challenges for geoscientists is effectively extracting the useful information from massive geo-datasets. Deep learning algorithms have become incredibly good at analyzing and identifying pieces of objects from massive data. The application of deep learning in seismic exploration has become one of the hottest research topics in recently two-years. My dissertation focuses on developing new workflows for seismic data processing and interpretation with the aid of deep learning algorithms.

Picking the first arrival of seismic data is one of the most time-consuming task in seismic data processing. The first arrival segments the seismic traces into two parts. Each part of the seismic traces can be viewed as a unique object. I automatically identify the two objects of the seismic trace by using a state-of-art pixel-wise convolutional image segmentation method. The boundary of the two objects is regarded as the first arrivals of seismic data. Noise filtering is another important step in seismic data processing. I proposed to filter the noise in seismic data by integrating deep learning and variational mode decomposition. My new method does not require prior information about the noise which is one of the compulsory inputs for image de-noising using deep learning. My method not only effectively removes the random noise in the seismic image but also the coherence noise such as migration artifacts which is beyond the capability of current filtering methods.

The process of seismic horizon interpretation can be treated as dividing the seismic traces into several segments. I proposed a workflow to perform semi-automated horizon interpretation

method by using the encoder-decoder convolutional neural network. There are two main parts of my workflow. The first part is segmenting the seismic traces into different parts using deep learning and treat the boundary of two nearby parts as the horizon. The second part is refining the horizons using a two-step filtering. My method does not require seismic attributes such as the dip and azimuth of a seismic reflector as the inputs which are compulsory for current horizon picking algorithms.

DEDICATION

This dissertation is dedicated to everyone who helped me and guided me through the trials and tribulations of creating this manuscript. In particular, my family and close friends who stood by me throughout the time taken to complete this degree program.

ACKNOWLEDGMENTS

Pursuing this Ph.D. degree program is a life-changing experience for me. This dissertation is impossible to be finished without the support of my family, my friends and all the people worked with me through the past three years.

First, I want to thank my advisor Dr. Bo Zhang who give me a chance to be his Ph.D. student. He taught me how to do research and write a scientific paper. He taught me how to make a clear research presentations. The most important thing that I have learned from him is how to think as an independent researcher, which will be a benefit for my future career.

I would also like to show my greatest gratitude to, Dr. Natasha T. Dimova, Dr. Yong Zhang, Dr. Zhe Jiang, Dr. Alain Plattner and Dr. Delores M. Robinson, who serve on my Ph.D. committee. I want to thank Dr. Dimova's encouragement and help to write a proposal to GSA. I want to thank Dr. Zhang's research guidance and scientific writing suggestions. I want to thank Dr. Jiang, his encouragement give more confidence to explore the connection between geoscience and computer science. I greatly appreciate the support received from Dr. Robinson. She helped me go through the scientific writing issues of this dissertation. I also greatly appreciate the research guidance from Dr. Fangyu Li and Dr. Tao Zhao, they provide much support when I begin to learn the deep learning techniques.

I appreciate the time and support from fellow students and friends including Naihao Liu, Yihuai Lou, Rongchang Liu, Nanke Wu, Wenhua Wang, Yan Ye, Man Lu and Shuo Chen. They provided essentially help during these three years.

Finally, I extended my deepest thanks to my parents for everything that they have done for me. It is impossible to finish my Ph.D. degree program without their love, courage and devotion.

CONTENTS

LIST OF FIGURES

CHAPTER 1
INTRODUCTION

In recent years, deep learning has achieved great success on computer science and attracting enormous attention on different disciplines. It has been applied to the tasks of image processing, speech recognition and natural language processing. Since the seismic data can be treated as images, my dissertation focuses on adapting and improving the deep learning based algorithms for seismic exploration.

Micro-seismic events picking is a key step in seismic processing and imaging. Manual picking is the most common but the most time-consuming interpretation technique. In recent decades, many approaches have been proposed to pick the first arrival of micro-seismic events automatically. One of the most widely used methods is the short-term average/long-term average (STA/LTA) (Baer et al., 1987). STA/LTA first computes the moving average energy ratio of the short time window to long time window. STA/LTA then chooses the peak that is larger than the user-defined threshold as the first arrival. Baer and Kradolder (1987) improved the algorithm of the STA/LTA by integrating a characteristic function (CF) and dynamic peak threshold into the algorithm. Saragiotis et al. (1999) used higher-order-statistics (HOS) in a redundant wavelet transform domain (R-WT) to pick the first arrival automatically. Zhang et al. (2003) used the Akaike information criteria (AIC) to detect and pick the first arrival automatically. However, these methods above usually encounter a problem in a really noisy environment. Other automatic first arrival picking methods include multi-window algorithm (Chen et al., 2005), modified energy ratio (MER) (Wong et al., 2009), transformed spectrogram approach (Song et al., 2010), kurtosis ratio (S/L-Kurt) method (Li et al., 2014), adaptive filtering (Li et al., 2017) and

clustering analysis (Chen, 2017). Song et al. (2010) employed the correlation-based algorithm that is capable of detecting the first arrival with low signal to noise ratio. Chen (2017) utilized an unsupervised machine learning algorithm for first arrival picking and obtained a more robust result than the state-of-art STA/LTA even in a really noisy environment.

Seismic denoising not only lowers the effect of subjectivity in seismic interpretation but also improve the reliability of seismic inversion. In recent decades, numerous seismic denoising approaches have been developed and widely applied in practice. Methods for seismic denoising can be generally classified into four categories: The first category is based on building a prediction filter to remove the noise of seismic data. The commonly used algorithms within the first category includes *f-x* predictive filtering (Canales, 1984), *t-x* predictive filtering (Abma et al., 1995), the forward-backward prediction approach (Wang, 1999), the polynomial fitting-based approach (Liu et al. 2011), and non-stationary predictive filtering (Liu et al., 2012). The second category projects the seismic data to a transformed domain and rejects the noise by applying a bandpass filter to the transformed data. Finally, obtain the denoised seismic data by projecting the filtered data back to the time domain. The commonly used algorithms within the second category include Fourier transform (Chen et al., 2014), curvelet transform (Herrman et al., 2008), seislet transform (Fomel et al., 2010), shearlet transform (Kong et al., 2015), Radon transform (Trad et al., 2002; Xue et al., 2016), wavelet transform (Donoho et al., 1994), and dictionary learning based sparse transform (Elad et al., 2006). The third category decomposes the seismic traces into a set of components. Then, examine the time or frequency features of each decomposed component. Finally, obtain the "clean" seismic traces by rejecting the components which are regarded as "noise". The commonly used algorithms within the third category include empirical mode decomposition (EMD) (Huang et al., 1998; Bekara et al., 2009), variational

mode decomposition (VMD) (Dragomiretskiy et al., 2014; Li et al., 2017), and singular value decomposition (SVD) based approaches (Bekara et al., 2007). Yuan et al. (2018) proposed a novel inversion-based denoising method. The method has the advantages of preserving 3-D spatial edges and low-frequency signals. The fourth category is based on the rank-reduction reconstruction of seismic data. The commonly used algorithms within the fourth category include Cadzow filtering (Trickett, 2008), and singular spectrum analysis (Vautard et al., 1992; Oropeza et al., 2011).

The seismic horizons can be treated as the stratigraphic boundaries that represent the depositional environments and geological features. Manual interpretation is the most familiar but time-consuming interpretation technique. In recent decades, seismic horizon interpretation has been automated to some extent by others. Zeng et al. (1998) proposed an interpolation method, which first manually picking several reference strata slices and then build a surface volume by interpolating the interpreted seismic surfaces. However, the interpolated surfaces may not follow the local discontinues, like the fault and unconformities. Stark (2003) computed the unwrapping instantaneous phase to generate a relative geological time (RGT) volume and produce multiple horizons simultaneously. The seismic horizons that generated by RGT volume can provide a quality measure along the unconformities, but it cannot follow the fault gap. Wu and Zhong (2012) improved this method by using the graph-cut phase unwrapping, which performs well at strong discontinues structure zone. Lomask et al. (2006) first calculated the local dips over the entire seismic volume and transferred them into the time shift and then apply the least-square method to track the seismic horizons automatically. Parks (2010) proposed a slope-based flattening method, which use the structure tensor to estimate the dip and track seismic horizons automatically. Fomel (2010) use the algorithm of predictive painting to build a 3D seismic

surface volume. Wu and Hale (2015) proposed an automatic horizons tracking method by adding control points in the complex structure zone. Wu and Fomel (2018) proposed a novel automatically seismic horizon interpretation method with local slope and multi-grid correlation that can effectively deal with faults without any human effort or very few control points.

In chapter 2 of my dissertation, I propose a novel workflow to automatically pick the first arrival of micro-seismic by using a state-of-art pixel-wise convolutional image segmentation method. I first form the training data by randomly selecting part of the micro-seismic traces and manually pick the time index of the first arrivals. I next segment the selected traces into two parts according to the time index of manual picking and assign each part a label accordingly. I then build an encoder-decoder convolutional neural network architecture and use the training data and training label as the input. I next obtain the trained network hierarchy by learning the segmented training data and labels. Finally, I predict the first arrivals of micro-seismic events by applying the trained network hierarchy to the rest of the micro-seismic traces. Both the synthetic and field data examples demonstrate that my proposed method successfully identifies the first arrivals. The predicted first arrival result obtained by using my proposed method is superior to the result obtained by using the traditional method of short-term average and long-term average (STA/LTA).

In chapter 3 of my dissertation, I propose to attenuate the seismic noise by integrating variational mode decomposition (VMD) and convolutional neural network. There are four main steps in the proposed workflow. The first step is decomposing the seismic volume into different intrinsic mode function and a residual component by using VMD. The second step is adding user-generated white noise to each decomposed component. The third step is building the neural network hierarchy to learn the feature of additive white noise. The last step is obtaining denoised

seismic data by applying the well-trained network to the original seismic data. I use both synthetic and field data examples to illustrate the robustness and superiority of the proposed method.

In chapter 4 of my dissertation, I propose a novel semi-automated seismic horizon interpretation method using an encoder-decoder convolutional neural network (CNN). To efficiently learn the waveform pattern that bounded by two adjacent horizons, I use a variable size for the convolution filters which differs from the current CNN based image segmentation. I form the training data by extracting the seismic traces on a user-defined coarse grid. The interpreted horizons for the seismic traces on the coarse grid separate the seismic traces into several segments. I then form the training labels by assigning a unique symbol to each segment. The training data and labels are then used to train a network hierarchy. Next, we obtain the seismic horizons over the whole seismic survey by applying the trained network to the rest of seismic traces. In addition, I designed a two-step smooth filter to remove the spike but kept the sharp fault, which can improve the accuracy of the predicted horizons. Two field data examples demonstrate that my proposed method can accurately generate seismic horizons for each seismic trace if I only have interpreted horizons on a coarse grid of the seismic survey. Finally, I made some test to find out the limitation of our proposed method that includes the minimum interval distance of the horizons and the maximum noise resisted ability. In the last chapter, I draw general conclusions and suggest possible future work direction.

REFERENCES

Abma, R., and J. Claerbout, 1995, Lateral prediction for noise attenuation by t-x and F-X techniques, Geophysics, vol. 60, no. 6, pp. 1887-1896.

Baer, M., and U. Kradolfer, 1987, An automatic phase picker for local and teleseismic events, Bulletin of Seismology Society of American, vol. 77, pp. 1598-1612.

Bekara, M., and M. van der Baan, 2007, Local singular value decomposition for signal enhancement of seismic data, Geophysics, vol. 72, no. 2, pp. V59-V65.

Bekara, M., and M. van der Baan, 2009, Random and coherent noise attenuation by empirical mode decomposition, Geophysics, vol. 74, no. 5, pp. V89-V98.

Canales, L., 1984, Random noise reduction: 54th Annual International Meeting, SEG, Expanded Abstract, pp. 525-527.

Chen. Y., 2017, Automatic microseismic event picking via unsupervised machine learning, Geophysical Journal International, vol. 212, pp. 88-102.

Chen, Y., J. Ma, 2014, Random noise attenuation by f-x empirical-mode decomposition predictive filtering, Geophysics, vol. 79, no. 3, pp. V81-V91.

Chen, Y., W. Yu, and T. Pock, 2015, On learning optimized reaction diffusion processes for effective image restoration, Process of IEEE Conference, Computer Vision Pattern Recognition, pp. 5261-5269

Chen, Z., and R. Stewart, 2005, Multi-window algorithm for detecting seismic first arrivals, CREWES.

Donoho, D. L., and I. M. Johnstone, 1994, Ideal spatial adaptation by wavelet shrinkage, Biometrika, vol. 81, pp. 425-455.

Dragomiretskiy, K., and D. Zosso, 2014, Variational mode decomposition, IEEE Transactions on Signal Processing, vol. 62, pp. 531-544.

Elad, M., and M. Aharon, 2006, Image denoising via sparse and redundant representations over learned dictionaries, IEEE Transaction on Image Processing, vol. 15, no. 12, pp. 3736-3745.

Fomel, S., 2010, Predictive painting of 3D seismic volumes, Geophysics, vol. 75, no. 4, pp. A25-A30.

Fomel, S., and Y. Liu, 2010, Seislet transform and seislet frame, Geophysics, vol. 75, pp. V25-V38.

Hermann, F. J., U. Boniger, and D. J. Verschuur, 2007, Non-linear primary multiple separation with directional curvelet frames, Geophysical Journal International, vol. 170, pp. 781-799.

Huang, N., Z. Shen, S., Long, M. Wu, H. Shih, Q. Zheng, and H. Liu, 1998, The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis, Proceedings of the Royal Society of London A: Mathmatical, Physical and Engineering Sciences, vol. 454, pp. 903-955.

Kong, D., and Z. Peng, 2015, Seismic random noise attenuation using shearlet and total generalized variation, Journal of Geophysics and Engineering, vol. 12, no. 6, pp. 1024-1035.

Lomask, J., A. Guitton, S. Fomel, J. Claerbout, and A. A. Valenciano, 2006, Flattening without picking, Geophysics, vol. 71, no. 4, pp. P13-P20.

Li. F, J. Rich, and K. Marfurt, 2014, Automatic event detection on noisy microseismogram, SEG expanded abstract, pp. 2363-2367.

Li. F, and W. Song, 2017, Automatic arrival identification system for real-time microseismic event location, SEG expanded abstract, pp. 2934-2939.

Li, F., B. Zhang, S. Verma, and K. Marfurt, 2017, Seismic signal denoising using thresholded variational mode decomposition, Exploration Geophysics, vol. 49, no. 4, pp. 450-461.

Liu, G., X. Chen, J. Du, and J. Song, 2011, Seismic noise attenuation using nonstationary polynomial fitting, Applied Geophysics, vol. 8, pp. 18-26.

Liu, Y, N. Liu, C. Liu, 2012, Adaptive prediction filtering in t-x-y domain for random noise attenuation using regularized nonstationary autoregression, Geophysics, vol. 80, no. 1, pp. V13-V21.

Oropeza, V., and M. Sacchi, 2011, Simultaneous seismic data denoising and reconstruction via multichannel singular spectrum analysis, Geophysics, vol. 76, no. 3, pp. V25-V32.

Parks, D., 2010, Seismic image flattening as a linear inverse problem, M.S. thesis, Colorado School of Mines.

Saragiotis, C. D., L. J. Hadjilenontiadis, and S. M. Panas, 1999, A high-order statistics-based phase identification of three-component seismograms in a redundant wavelet transform domain, Proceeding of IEEE Signal Processing Workshop on High-Order Statistics, pp. 396-399.

Song, F., H. S. Kuleli, M. N. Toksoz, E. Ay, and H. Zhang, 2010, An improved method for hydrofracture-induced microseismic event detection and phase picking, Geophysics, vol. 75, no. 6, pp. A47-A52.

Stark, T. J., 2003, Unwrapping instantaneous phase to generate a relative geologic time volume, 73rd SEG Annual International Meeting Expanded Abstracts, pp. 1707-1710.

Trad, D., T. Ulrych, and M. Sacchi, 2002, Accurate interpolation with high-resolution time variant Radon transforms, Geophysics, vol. 67, no. 2, pp. 644-656.

Trickett, S., 2008, F-xy cadzow noise suppression, CSPG CSEG CWLS Convention, pp. 303-306.

Vautard, R., P. Yiou, and M. Ghil, 1992, Singular-spectrum analysis: A toolkit for short, noisy chaotic signals, Physica D: Nonlinear Phenomena, vol. 58, no.1, pp. 95-126.

Versteeg, R., 1994, The Marmousi experience: Velocity model determination on a synthetic complex data set, The Leading Edge, vol. 13, pp. 927-936.

Wang, Y., 1999, Random noise attenuation using forward-backward linear prediction, Journal of Seismic Exploration, vol. 8, pp. 133-142.

Wong, J., L. Han, J. Bancroft, and R. Stewart, 2009, Automatic time-picking of first arrivals on noisy microseismic data, CSEG Conference Abstract, vol. 1, pp. 1-4.

Wu, X., and D. Hale, 2015, Horizon volumes with interpreted constraints, Geophysics, vol. 80, no. 2, pp. IM21-IM33.

Wu, X., and G. Zhong, 2012, Generating a relative geologic time volume by 3D graph-cut phase unwrapping method with horizon and unconformity constraints, Geophysics, vol. 77, no. 4, pp. O21-O34.

Wu, X., and S. Fomel, 2018, Least-squares horizons with local slopes and multi-grid correlations, Geophysics, vol. 84, no. 4, pp. IM29-IM40.

Wu, X., and Y. Shi, S. Fomel, and L. Luming, 2018, Convolutional neural networks for fault interpretation in seismic images, 88th Annual International Meeting, SEG, Expanded Abstracts, pp. 1946-1950.

Wu, Z., and N. Huang, 2004, A study of the characteristics of white noise using the empirical mode decomposition method, Proceedings of The Royal Society, vol. 460, pp. 1597-1611.

Wu, Z., and N. Huang, 2009, Ensemble empirical mode decomposition: a noise-assisted data analysis method, Advances in Adaptive Data Analysis, vol. 1, pp. 1-41.

Xue, Y., J. Yang, J. Ma, and Y. Chen, 2016, Amplitude-preserving nonlinear adaptive multiple attenuation using the high-order sparse radon transform, Journal Geophysics and Engineering, vol. 13, pp. 207-219.

Yuan, S., S. Wang, C. Luo, and T. Wang, 2018, Inversion-based 3-D seismic denoising for exploring spatial edges and spatio-temporal signal redundancy. IEEE Geoscience and Remote Sensing Letters, vol. 15, no. 11, pp. 1682–1686.

Zeng, H., M. M. Backus, K. T. Barrow, and N. Tyler, 1998, Stratal slicing; Part 1, Realistic 3D seismic model, Geophysics, vol. 63, no. 2, pp. 502-513.

Zhang, H., C. Thurber, and C. Rowe, 2003, Automatic p-wave arrival detection and picking with multiscale wavelet analysis for single-component recordings, Bulletin of Seismology Society of American, vol. 93, no. 5, pp. 1904-1912.

CHAPTER 2
SEMI-AUTOMATIC FIRST ARRIVAL PICKING OF MICRO-SEISMIC EVENTS BY USING PIXEL-WISE CONVOLUTIONAL IMAGE SEGMENTATION METHOD

Hao Wu[1], Bo Zhang[1], Fangyu Li[2], Naihao Liu[3]
[1]University of Alabama, Department of Geological Science.
[2]University of Georgia, College of Engineering.
[3]Xi'an Jiaotong University, School of Electronic and Information Engineering.
This paper was published by SEG journal Geophysics in 2019

ABSTRACT

Micro-seismic imaging plays an important role in hydraulic fracture detection and the first arrival picking of micro-seismic events is the bedrock of micro-seismic imaging. Manual picking is the most reliable but also the most time-consuming method for the detection of the first arrival of micro-seismic events. Accurate and efficient first arrival picking in a real noisy environment is a challenge for most of the automatic first arrival picking methods. We propose a novel workflow to automatically pick the first arrival of micro-seismic by using a state-of-art pixel-wise convolutional image segmentation method. We first form the training data by randomly selecting part of the micro-seismic traces and manually pick the time index of the first arrivals. We next segment the selected traces into two parts according to the time index of manual picking and assign each part a label accordingly. We then build an encoder-decoder convolutional neural network architecture and use the training data and training label as the input. We next obtain the trained network hierarchy by learning the segmented training data and labels. Finally, we predict the first arrivals of micro-seismic events by applying the trained network hierarchy to the rest of the micro-seismic traces. Both the synthetic and field data examples demonstrate that our proposed method successfully identifies the first arrivals. The

predicted first arrival result obtained by using our proposed method is superior to the result obtained by using the traditional method of short-term average and long-term average (STA/LTA).

INTRODUCTION

Micro-seismic events picking is a key step in seismic processing and imaging. Manual picking is the most common but most time-consuming interpretation technique. In recent decades, many approaches have been proposed to pick the first arrival of micro-seismic events automatically. One of the most widely used methods is the short-term average/long-term average (STA/LTA) (Baer et al., 1987). STA/LTA first computes the moving average energy ratio of the short time window to long time window. STA/LTA then chooses the peak that is larger than the user-defined threshold as the first arrival. Baer and Kradolder (1987) improved the algorithm of the STA/LTA by integrating a characteristic function (CF) and dynamic peak threshold into the algorithm. Saragiotis et al. (1999) used higher-order-statistics (HOS) in a redundant wavelet transform domain (R-WT) to pick the first arrival automatically. Zhang et al. (2003) used the Akaike information criteria (AIC) to detect and pick the first arrival automatically. However, these methods above usually encounter a problem in a real noisy environment. Other automatic first arrival picking methods include multi-window algorithm (Chen et al., 2005), modified energy ratio (MER) (Wong et al., 2009), transformed spectrogram approach (Song et al., 2010), kurtosis ratio (S/L-Kurt) method (Li et al., 2014), adaptive filtering (Li et al., 2017) and clustering analysis (Chen, 2017). Song et al. (2010) employed the correlation-based algorithm that is capable of detecting the first arrival with low signal to noise ratio. Chen (2017) utilized an unsupervised machine learning algorithm for first arrival picking and obtained a more robust result than the state-of-art STA/LTA even in a real noisy environment.

Deep learning is a subset of machine learning and has gained attention in various fields. Convolution neural network (CNN) proposed by LeCun et al. (1989) is one of the most popular and widely used deep learning algorithms. Many CNN based algorithms have been proposed for image segmentation and objects recognition, including LeNet (LeCun et al., 1998), AlexNet (Krizhevsky et al., 2012), RCNN (Girshick et al., 2014), GoogleNet (Szegedy et al., 2015) and VggNet (Simonyan et al., 2015). Long et al. (2015) proposed an end-to-end convolutional neural network that called fully connected neural network (FCN), which enhances the capability of image segmentation from regional level to pixel level. A general end-to-end pixel-wise image segmentation architecture can be broadly regarded as an encoder network followed by a decoder network. The encoder is a pre-trained classification network and decoder projects the discriminative features learned by the encoder semantically. He et al. (2015) introduced a novel architecture with a skip connection into the deep neural network. Skip connections are extra connections between nodes in different layers of a neural network that skips one or more layers of nonlinear processing, which makes the training of very deep network possible. Badrinarayanan et al. (2015) proposed the SegNet for pixel-wise image segmentation. Comparing with FCNs, SegNet has two advantages: (1) the upsampling layer in the decoder efficiently keeps high-frequency details in the segmentation; (2) the convolution layer can memorize the indices of image feature by using convolution layer instead of using fully connected layer.

In this paper, we propose a novel micro-seismic first arrival picking method based on the convolutional neural network. First, we briefly introduce the application of SegNet for image segmentation. We then describe our proposed workflow for the first arrival picking of micro-seismic events. Finally, we show the application of our proposed workflow in the synthetic and

field data examples, which demonstrate the effectiveness and superiority of our proposed method over the traditional method of STA/LTA.

THEORY

SegNet

SegNet is a CNN based algorithm that used for the pixel-level image segmentation. Figure 2.1 shows a typical architecture of SegNet, which consists of an encoder network and a decoder network. The encoder network consists of a sequence of non-linear processing layers and each layer contains four operations. The first, second, third, and fourth operators are convolution filter (Conv), batch-normalization regularizer (BN), ReLU activation operator, and max-pooling operator, respectively. The decoder network also consists of a sequence of non-linear layers that are corresponding to the layers in the encoder network. Each layer includes upsampling operator, convolution filter, batch-normalization regularizer, and ReLU activation operator. The layers in the encoder and decoder networks are connected using "skip-connection". The last step of the network is an element-wise classification layer.

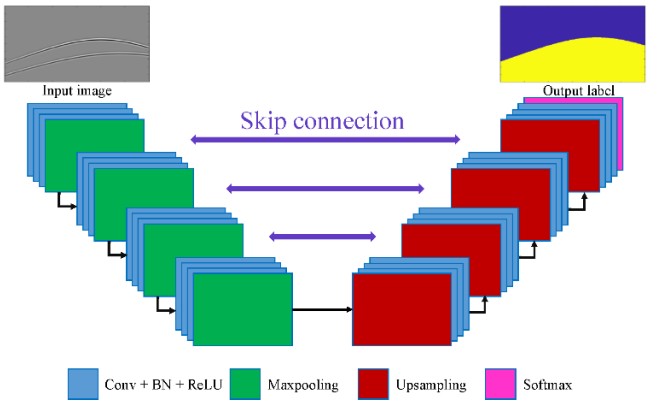


Figure 2.1. The overall architecture of the proposed deep convolutional end-to-end neural network. It has two main parts: an encoder network and a decoder network.

Obtaining the trained network is the most important step in the process of image segmentation using the encoder-decoder neural network. First, we need a set of training images to determine the parameters of the operators in the architecture. Each image in the training set has been manually segmented, and each segment in the image has a unique label. The objective of the training process is to determine the parameters of the operators in the encoder and decoder parts. The purpose of the encoder is passing the main features of the image to decoder while the purpose of the decoder is reconstructing the feature of the image. We iteratively update the parameters of operators by minimizing the loss function. The built network is then used to segment those images that need to be segmented.

First arrival picking using SegNet

The first arrival of micro-seismic events is picked by detecting the onset time of micro-seismic arrivals in the refracted signals from the wellbore. Figure 2.2 shows a typical micro-seismogram recorded by a geophone and the first arrival is indicated by the red arrow.

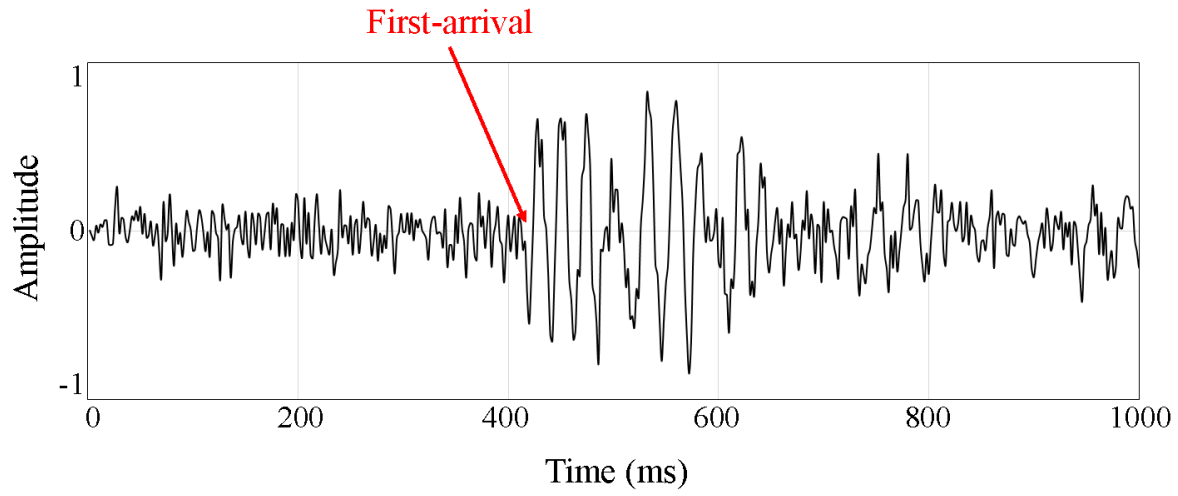


Figure 2.2. Micro-seismogram record. From this Micro-seismic trace, we can detect the first arrival that indicates by the red arrow.

A micro-seismic trace is treated as a 1D image and the time index of first arrival segments the 1D image into two parts. The lower part contains the information related to the micro-seismic events, while the upper part does not. In this manner, the first arrival picking of micro-seismic events can be regarded as detecting the boundary (a pixel) that divides the 1D seismic image into two sequences. SegNet is a pixel-level image segmentation method and is capable of obtaining an accurate boundaries for multiple objects. In addition, SegNet is superior to other CNN based image segmentation methods that detect the objects by using a series of the rectangular region. As a result, we propose a novel workflow by employing SegNet to pick the first arrival of micro-seismic events.

Workflow

The workflow begins by converting the micro-seismic trace into 1D gray-scale image. The function that transfers the seismic amplitude to the gray image is defined as follow:

$$\mathbf{z} = 255 \times \left(\frac{\boldsymbol{s} - mean(\boldsymbol{s})}{2 \times (max(\boldsymbol{s}) - min(\boldsymbol{s}))} + \frac{1}{2} \right), \tag{1}$$

where $\boldsymbol{s}$ is the input seismic trace and $\mathbf{z}$ is the gray output image.

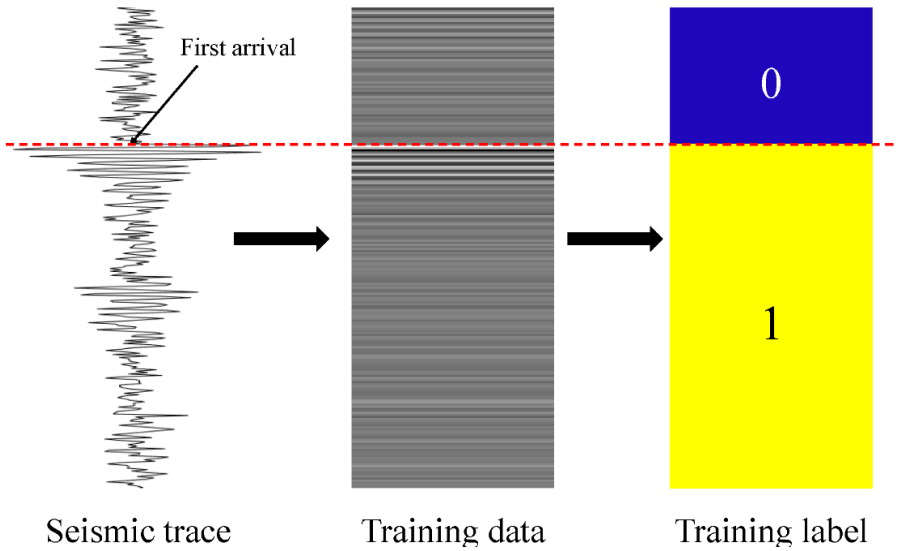


Figure 2.3. The basic workflow to build the train data and train label.

We next segment the micro-seismic trace into two sequences according to the time index of manually picked first arrival. The upper and lower sequences are labeled with 0 and 1, respectively. Any symbols can be assigned to the upper and lower parts of the segmented seismic trace as long as the symbols of the two sequences are different from each other. Figure 2.3 illustrates the process of data preparation. The first, second, and third panels in Figure 2.3 are the seismic trace, seismic image, and corresponding segmented label. The red dash line in Figure 2.3 indicates the time index of the first arrival. The gray images and corresponding labeled images are the inputs for the SegNet to train the proposed neural network.

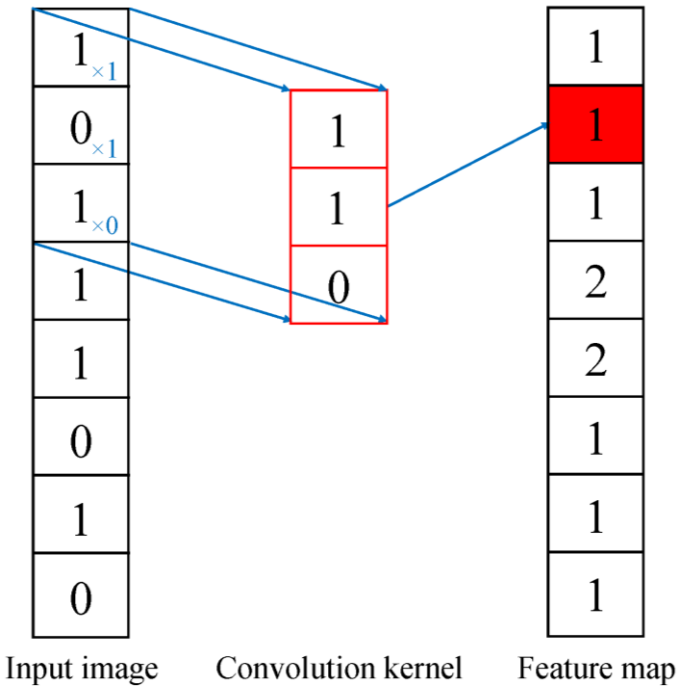


Figure 2.4. Example of convolution operation on a one-dimensional signal.

After preparing the input data, we build a deep convolutional encoder-decoder neural network using a sequence of nonlinear processing layers (encoder and the corresponding decoder) followed by a softmax classification layer. The encoder consists of several layers that are convolution filter, batch-normalization regularizer, ReLU activation operator and max-

pooling operator. Figures 2.4, 2.5, and 2.6 together illustrate the encoder-decoder process by using one-dimensional input data.

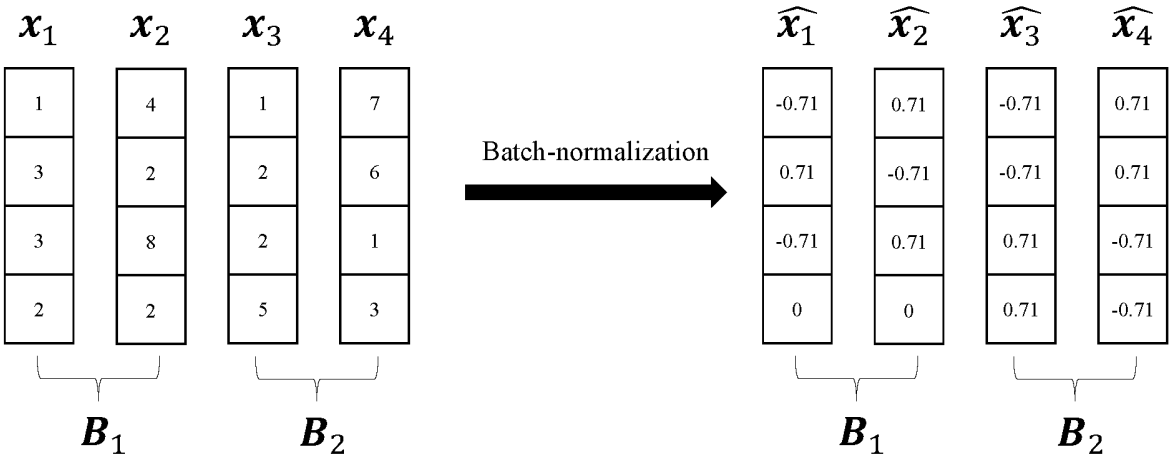


Figure 2.5. Illustration of batch-normalization regularizer on a one-dimensional signal.

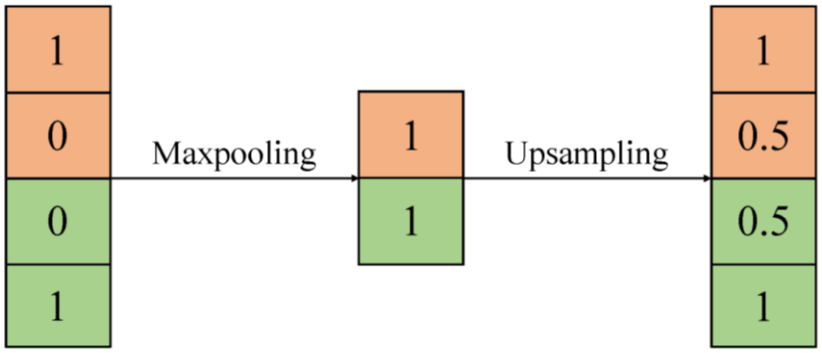


Figure 2.6. Illustration of batch-normalization regularizer on a one-dimensional signal.

The first step in the encoder part is convolution and the objective of the convolution layer is extracting the feature of images to form the feature maps. Consider the input data in this paper is 1D gray image, so the equation of convolution layer is defined as follow:

$$\boldsymbol{x} = \boldsymbol{b} + \boldsymbol{w} * \mathbf{z} \tag{2}$$

where * denotes the convolution operation, $\boldsymbol{b}$ is a bias term to be learned, $\boldsymbol{z}$ is the input data, $\boldsymbol{w}$ is the convolution filter, and $\boldsymbol{x}$ is the output feature maps. Figure 2.4 shows an example of the convolution operation on the one-dimensional signal.

The output feature maps $\boldsymbol{x}$ is normalized by using the batch-normalization (BN) (Ioffe et al., 2015) algorithm, which consists of four steps:

i. Define a batch $\boldsymbol{B} = \{\boldsymbol{x}_1, \boldsymbol{x}_2, \dots, \boldsymbol{x}_m\}$ according to the user defined batch size *m*. We randomly select seismic traces from the input training seismic traces to form a batch. Figure 2.5 illustrates the process of batch definition and the batch size is 2. Seismic traces of number 1 and 2 together form the first batch while number 3 and 4 together form the second batch.

ii. Compute the mean value μ_B and variance σ_B^2 of the seismic traces within each batch

$$\mu_B = \frac{1}{m}\sum_{i}^{m} \boldsymbol{x}_i \tag{3}$$

$$\sigma_B^2 = \frac{1}{m}\sum_{i}^{m} (\boldsymbol{x}_i - \mu_B)^2 \tag{4}$$

iii. Generate the normalized vector $\hat{x}_i$

$$\hat{\boldsymbol{x}}_i = \frac{\boldsymbol{x}_i - \mu_B}{\sqrt{\sigma_B^2 + \varepsilon}}, \tag{5}$$

where ϵ is a small positive constant value to avoid a case with zero variance value. We use the normalized vector $\hat{\boldsymbol{x}}_i$ to update the gradient at the current layer.

iv. The scaled and shifted normalized vector $\boldsymbol{c}_i$ is the input for the next layer:

$$\boldsymbol{c}_i = \gamma \hat{\boldsymbol{x}}_i + \beta, \tag{6}$$

where γ and β are the parameters to be learned. To pass the main features of the current layer to the next layer, a rectified linear unit (ReLU) activation operator is applied to the scaled and shifted vector:

$$\hat{\boldsymbol{c}}_i = max(0, \boldsymbol{c}_i) \tag{7}$$

The max-pooling operation is the final step in each layer of the encoder part. The upsampling is the first step in each layer of the decoder part. The upsampled result is used as the

input for constructing the feature maps. Because the input data are a 1D signal, the size of the max-pooling and upsampling operators is 2×1. Figure 2.6 illustrates the processing of max-pooling and upsampling for a 1D signal.

Next, we use the cross-entropy loss function to compute the loss between the ground truth and predicted probability distribution. The loss function is defined as:

$$h_\theta(\hat{\boldsymbol{c}}_i) = \frac{1}{1+e^{-\theta^T \hat{\boldsymbol{c}}_i}}, \tag{8}$$

$$\boldsymbol{J}(\theta) = -\frac{1}{m}\sum_{i}^{m} \boldsymbol{y}_i \log(h_\theta(\hat{\boldsymbol{c}}_i) + (1-\boldsymbol{y}_i)\log(1-h_\theta(\hat{\boldsymbol{c}}_i)), \tag{9}$$

where $h_\theta(\hat{c}_i)$ is the hypothesis function of $\hat{c}_i$, m is the total number of input data, and $\boldsymbol{y}_i$ is the corresponding training label.

We then employ the adaptive moment estimation (Adam) algorithm (Kingma et al., 2015) to iteratively update the parameters that need to be learned in the proposed deep convolutional neural network. The algorithm of Adam consists of the following steps:

i. Initialize the convolution filter $\boldsymbol{w}$, and bias terms $\boldsymbol{b}$ with random numbers.

ii. Initialize the 1st moment vector $\boldsymbol{p}_0$ and 2nd moment vector $\boldsymbol{q}_0$. Set the exponential decay rates β_1 and β_2 for the moment estimation within the range from 0 to 1.

iii. Get the gradient $\boldsymbol{g}_t$ of the loss function at time step t, which is given by:

$$\boldsymbol{g}_t = \nabla_\theta \boldsymbol{J}(\theta_{t-1}). \tag{10}$$

iv. Update the 1st moment vector $\boldsymbol{p}_t$ and 2nd moment vector $\boldsymbol{q}_t$, which is defined as,

$$\boldsymbol{p}_t = \beta_1 \cdot \boldsymbol{p}_{t-1} + (1-\beta_1)\boldsymbol{g}_t, \tag{11}$$

$$\boldsymbol{q}_t = \beta_2 \cdot \boldsymbol{q}_{t-1} + (1-\beta_2)\boldsymbol{g}_t^2, \tag{12}$$

v. Compute the corrected moment vectors $\widehat{\boldsymbol{p}_t}$ and $\widehat{\boldsymbol{q}_t}$:

$$\hat{\boldsymbol{p}}_t = \boldsymbol{p}_t / 1-\beta_1^t \,, \tag{13}$$

$$\hat{\boldsymbol{q}}_t = \boldsymbol{q}_t / 1-\beta_2^t \,, \tag{14}$$

vi. Update the gradient, convolution kernel ***w,*** and bias term ***b***.

$$\theta_t = \theta_{t\text{-}1} - \alpha \cdot \hat{\boldsymbol{p}}_t / \left(\sqrt{\hat{\boldsymbol{q}}_t} + \varepsilon \right), \tag{15}$$

$$\boldsymbol{w}_t = \boldsymbol{w}_{t\text{-}1} - \alpha \cdot \frac{\partial \boldsymbol{J}(\theta_{t\text{-}1})}{\partial \boldsymbol{w}_{t\text{-}1}} \,, \tag{16}$$

$$\boldsymbol{b}_t = \boldsymbol{b}_{t\text{-}1} - \alpha \cdot \frac{\partial \boldsymbol{J}(\theta_{t\text{-}1})}{\partial \boldsymbol{b}_{t\text{-}1}} \,, \tag{17}$$

The training process will keep updating the parameters until the loss function or epoch time meet the user-defined threshold. We finally apply the trained network to the rest of micro-seismic traces and extract the boundary as the predicted first arrival of micro-seismic events.

SYNTHETIC AND FIELD DATA TEST

Prepare input data for SegNet

Figure 2.7 shows a two-layer velocity model with the yellow dots representing the micro-seismic events source and the upside-down green triangles representing the geophones. Figure 2.8 shows the synthetic micro-seismic data set generated by centralized reverse time migration (RTM) source location method. Our synthetic example has 300 traces and the time sampling interval is 1*ms*. The blue dots in Figure 2.8 are the manual picking of the micro-seismic gather and these manual interpretation function as the ground truth for validating our workflow. We randomly select 30 seismic traces (Red color in Figure 2.9a) as the training dataset (Figure 2.9b). We next convert the selected micro-seismic traces into gray images (Figure 2.9c) and label the training traces (Figure 2.9d) according to the time index of manual picking.

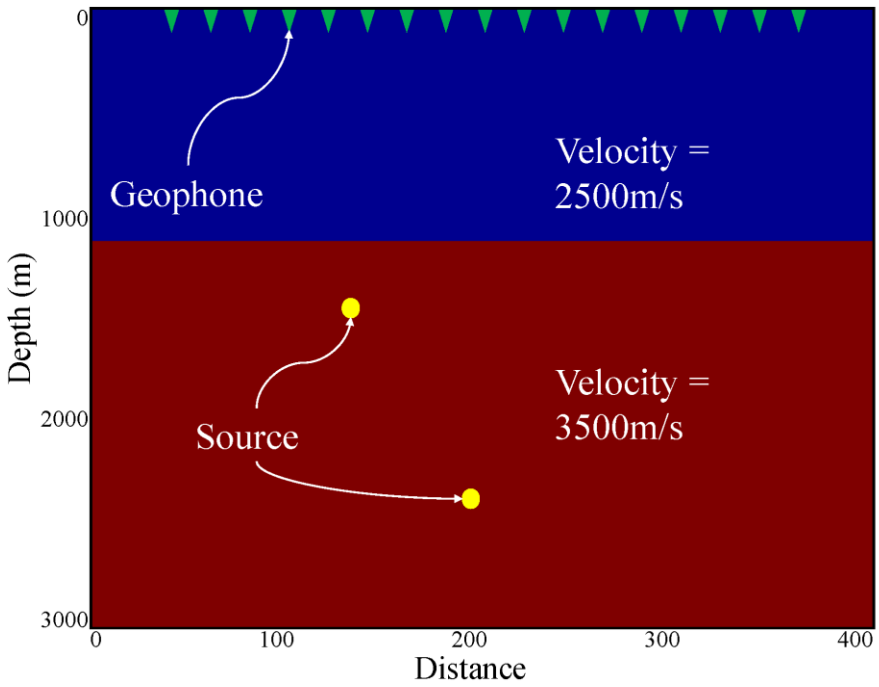


Figure 2.7. Two-layer velocity model for synthetic example.

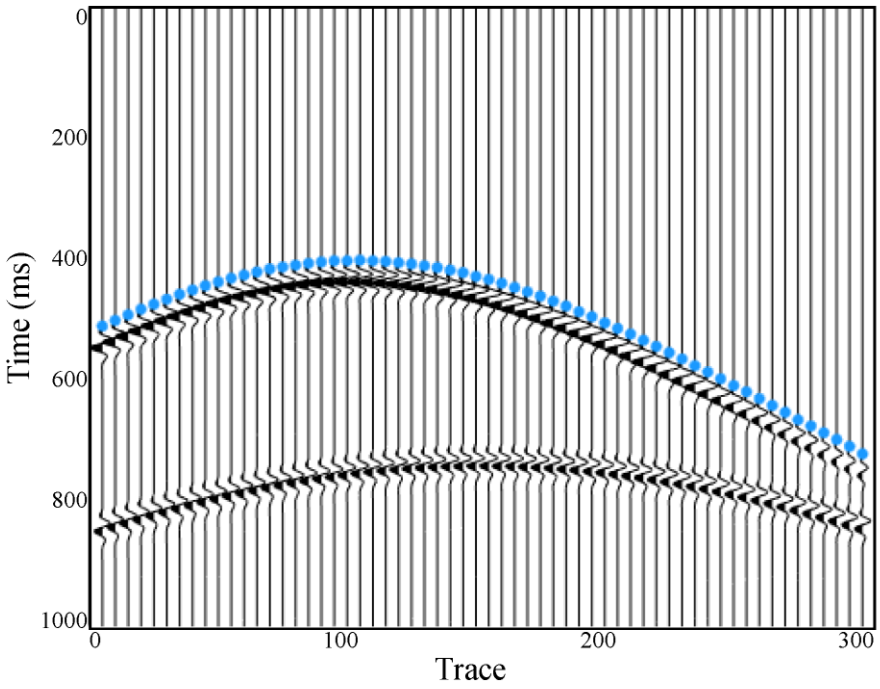


Figure 2.8. Manually picked first arrival of the synthetic micro-seismic gather. We use this manually picked first arrival as the ground truth model.

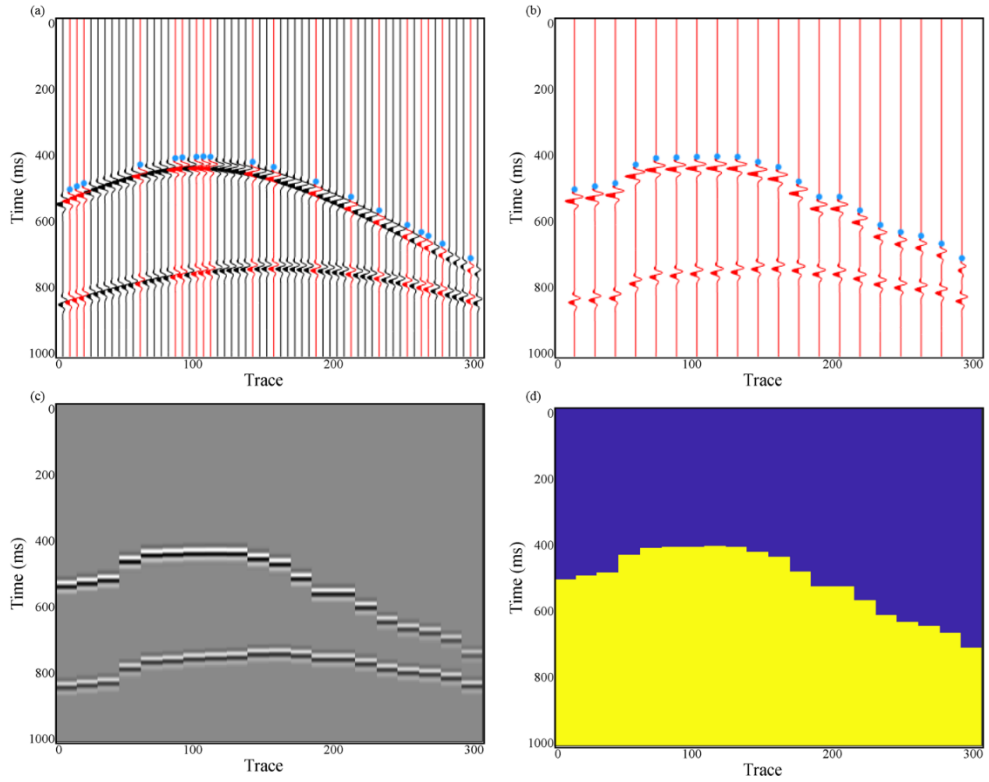


Figure 2.9. Training data and training label for synthetic example. (a) Randomly selected seismic traces over the whole seismic survey. (b) Organized selected seismic traces. (c) Convert the selected seismic traces to grey images. (d) Label the training data to build the training label.

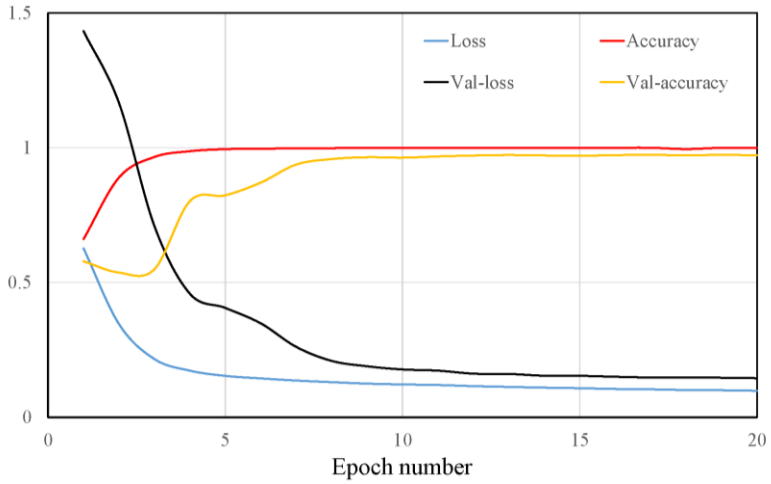


Figure 2.10. Illustrating the loss and accuracy of training processing on noise-free synthetic example.

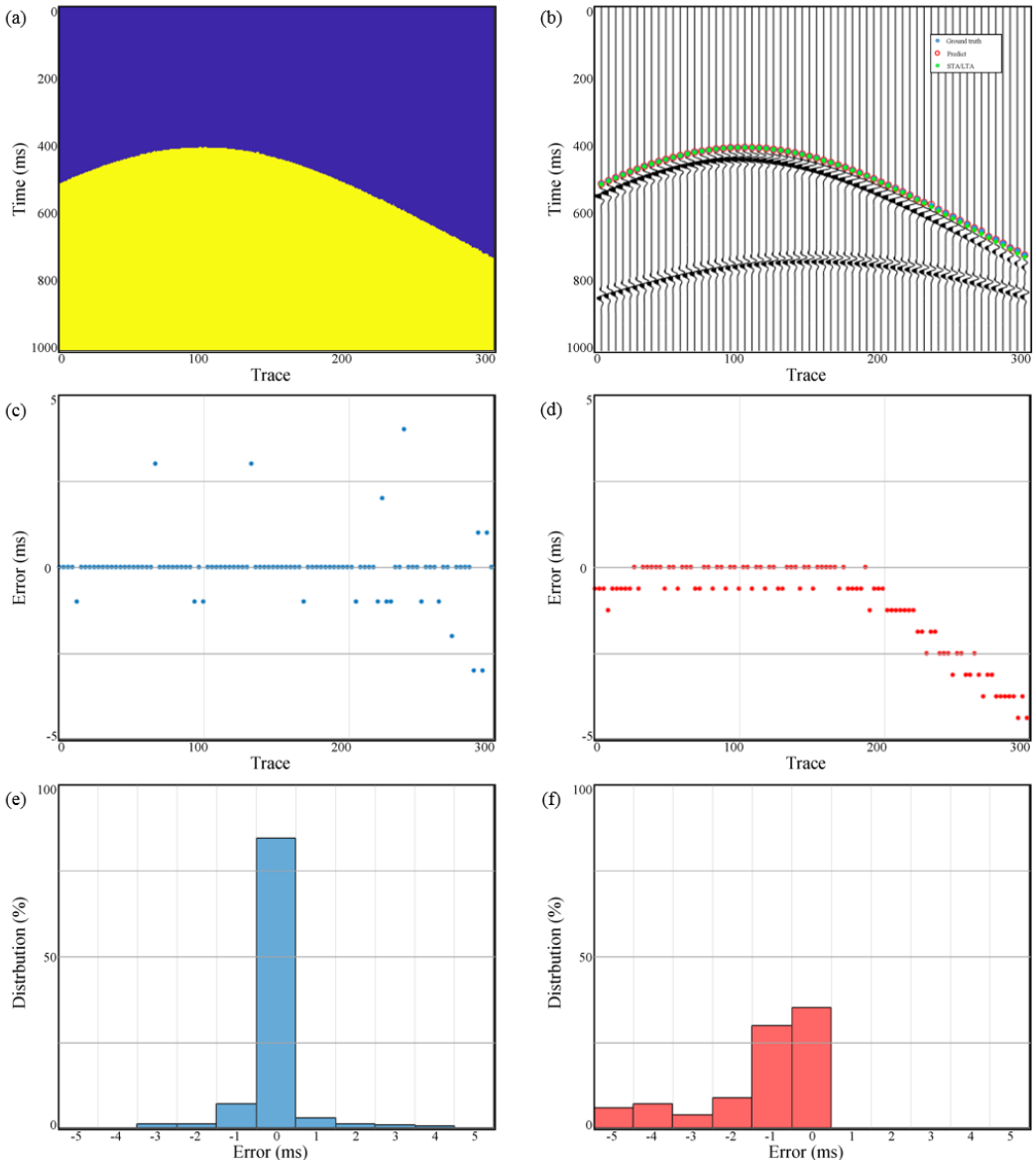


Figure 2.11. The predicted result of noise-free synthetic data. (a) The predicted label through our proposed method. (b) The ground truth first arrival (blue dots), first arrival picking by using STA/LTA (green dots) and predicted first arrival (red circles) overlaid the synthetic data. (c) The difference between the ground truth and predicted first arrival. (d) The difference between the ground truth and the picked first arrival using STA/LTA. (e) The distribution of the difference between the ground truth and predicted first arrival. (f) The distribution of the difference between the ground truth and the picked first arrival using STA/LTA.

Build a deep encoder-decoder neural network

We next build a deep convolutional encoder-decoder neural network. Six layers existed in the encoder network with the convolution kernel size of 32×1, 24×1, 16×1, 8×1, 5×1, and 3×1, respectively. In the decoder network, there are six layers corresponding to the structure of the encoder network with the convolution kernel size of 3×1, 5×1, 8×1, 16×1, 24×1, and 32×1, respectively. The validation data set is different at each iteration and provide the unbiased evaluation. Figure 2.10 shows the training loss and accuracy varying with parameters updating loops (epochs). Our training process converges after 20 epochs and the validation accuracy in the training process is above 99%. The training and validation accuracy is defined as:

$$accuracy = \frac{m}{n} \tag{18}$$

where m denotes the amount of matched pixels between the predicted and ground truth labels, and n denotes the total amount of pixels.

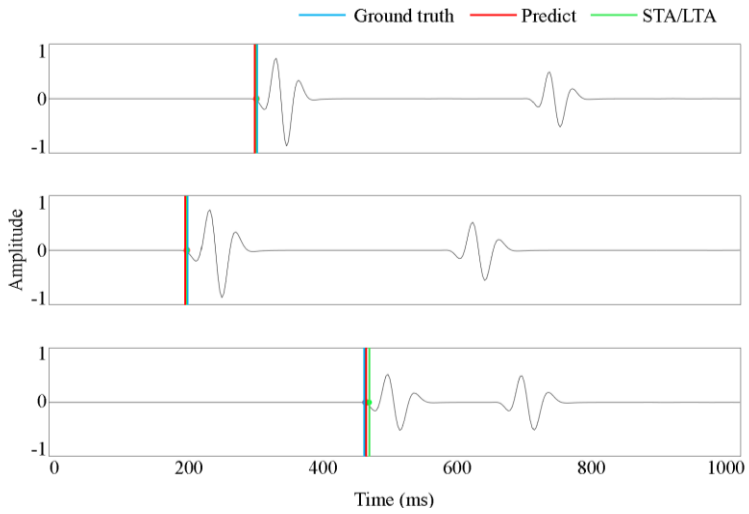


Figure 2.12. Comparison between the predicted and ground truth first arrival for three representative micro-seismic events of noise-free synthetic data test.

Predict the first arrival of micro-seismic events for noise-free synthetic

To demonstrate the performance of our proposed method, we first test a noise-free synthetic example and compare with the traditional method of STA/LTA. Figure 2.11a shows the predicted labels and the boundary between these two labels are treated as the time index of first arrival picking. The red circles and green dots in Figure 2.11b are the first arrival picking results by using our proposed method and STA/LTA, respectively. Both of these two methods obtained a very good match with the ground truth indicated by blue dots in Figure 2.11b. To further evaluate the performance of our proposed method and STA/LTA, we compare the difference between the manual picking with our method (Figure 2.11c) and STA/LTA (Figure, 2.11d). We then compute the absolute average difference between predicted and manually picked first arrival of micro-seismic events:

$$d = \frac{\sum_{i=1}^{N} |G_i - P_i|}{N} \tag{19}$$

where d denotes the absolute average difference, G_i denotes the ground truth first arrival of the micro-seismic event for the i^{th} trace, P_i denotes the first arrival of micro-seismic events for the i^{th} trace by using our proposed method or STA/LTA, N is the total trace number of the testing data. The absolute average difference by using our proposed method and STA/LTA are 0.25ms and 1.95ms, respectively. The absolute average difference indicates that the average picking error obtained by our proposed method is much smaller than the sampling interval of the micro-seismic events. Figure 2.11e and 11f shows the distribution of the difference by using our proposed method and STA/LTA, respectively. Around 85% of the micro-seismic events have an accurate prediction for the time index of the first arrival. The variance of the distribution is 0.52, which has shown a good stability of our predicted result. Figure 2.12 shows the comparisons

between the predicted and ground truth first arrival for three representative micro-seismic events. Note the good match between the ground truth and predicted the first arrival.

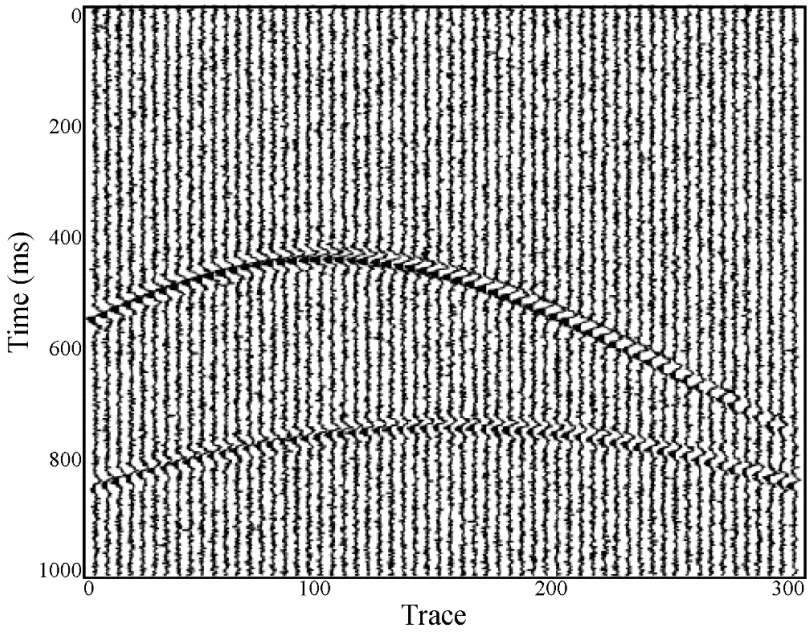


Figure 2.13. Noisy synthetic micro-seismic example.

Predict the first arrival of micro-seismic events for noisy synthetics

To illustrate the effectiveness and superiority of our proposed method over the traditional method of STA/LTA, we test a noisy synthetic example (Figure 2.13) where the signal to noise ratio (SNR) equals to 1. We repeat the procedure of the noise-free synthetic example and the predicted label is shown in Figure 2.14a. The first arrival picking results by using our proposed method and STA/LTA are indicated by red circles and green dots shown in Figure 2.14b. Note the good match between the result by using our proposed method and the ground truth (blue dots), while the method of STA/LTA has failed to pick the correct first arrival. Figures 2.14c and 2.14d show the difference with the ground truth by using our proposed method and STA/LTA, respectively. The absolute average differences are 0.41*ms* and 229.57*ms* that corresponding to our

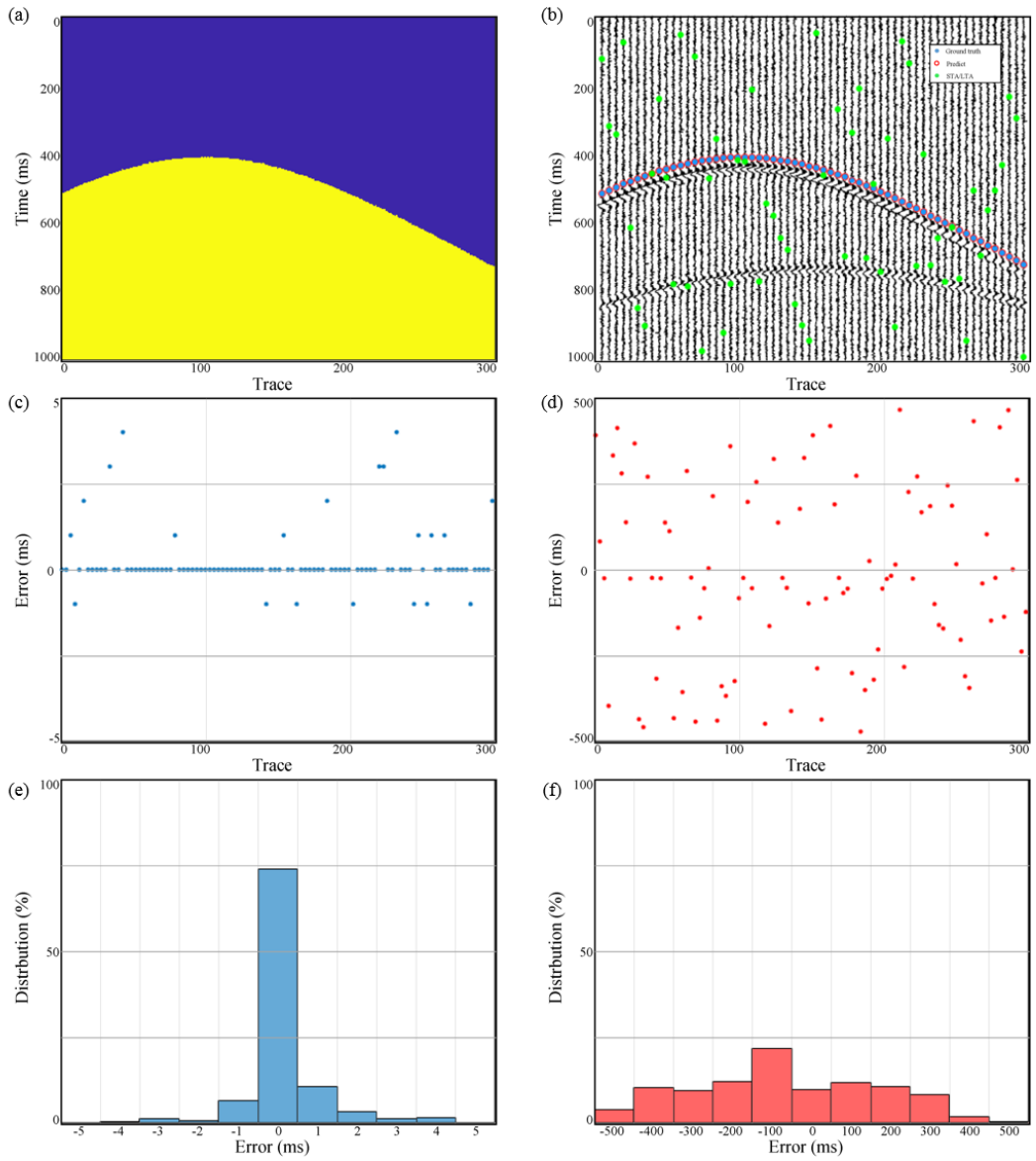


Figure 2.14. The predicted result in a noisy synthetic example. (a) The predicted label through our proposed method. (b) The ground truth (blue dots), first arrival picking using STA/LTA (green dots) and predicted first arrival (red circles) overlaid the synthetic data. (c) The difference between the ground truth and predicted first arrival. (d) The difference between the ground truth and the picked first arrival using STA/LTA. (e) The distribution of the difference between the ground truth and predicted first arrival. (f) The distribution of the difference between the ground truth and the picked first arrival using STA/LTA.

method and STA/LTA. Figures 2.14e and 2.14f show the distribution of the difference by using our proposed method and STA/LTA, respectively. The variance of the distribution by using these two methods are 0.87 and 780.59, which our proposed method has shown good stability for the accurate first arrival picking. Figure 2.15 shows the comparisons between the predicted, STA/LTA and ground truth first arrival for three representative micro-seismic events. Note the good match with the ground truth by using our proposed method and the significant difference by using STA/LTA.

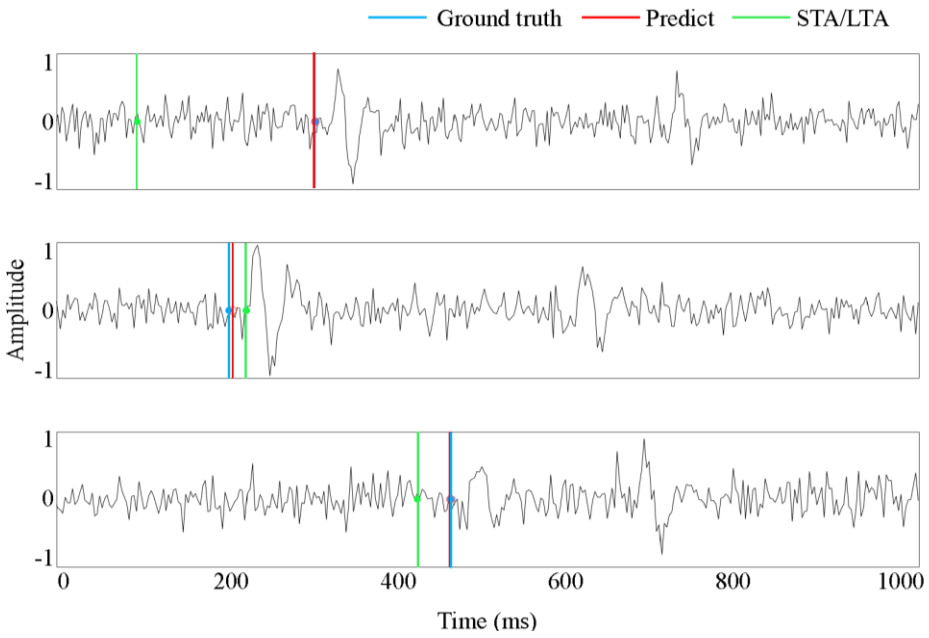


Figure 2.15. Comparison between the predicted and ground truth first arrival for three representative micro-seismic events of noisy synthetic data test.

Predict the first arrival of micro-seismic events for various SNR synthetics

To illustrate the robustness of our proposed method, we test a more challenging synthetic example to demonstrate the performance. Figure 2.16 shows a synthetic example that generates from a different layer-cake model. The SNR is various from trace to trace and the average SNR

is equal to 1. Figures 2.17c and 2.17d show the difference with the ground truth by using our proposed method and STA/LTA, respectively. The absolute average differences are 0.403*ms* and 230.75*ms* that corresponding to our method and STA/LTA. Figures 2.17e and 2.17f show the distribution of the difference by using our proposed method and STA/LTA, respectively. The variance of the distribution by using these two methods are 0.86 and 778.61, which our proposed method has shown good stability for the accurate first arrival picking. Figure 2.18 shows the comparisons between the predicted, STA/LTA and ground truth first arrival for three representative micro-seismic events. Note the good match with the ground truth by using our proposed method and the significant difference by using STA/LTA.

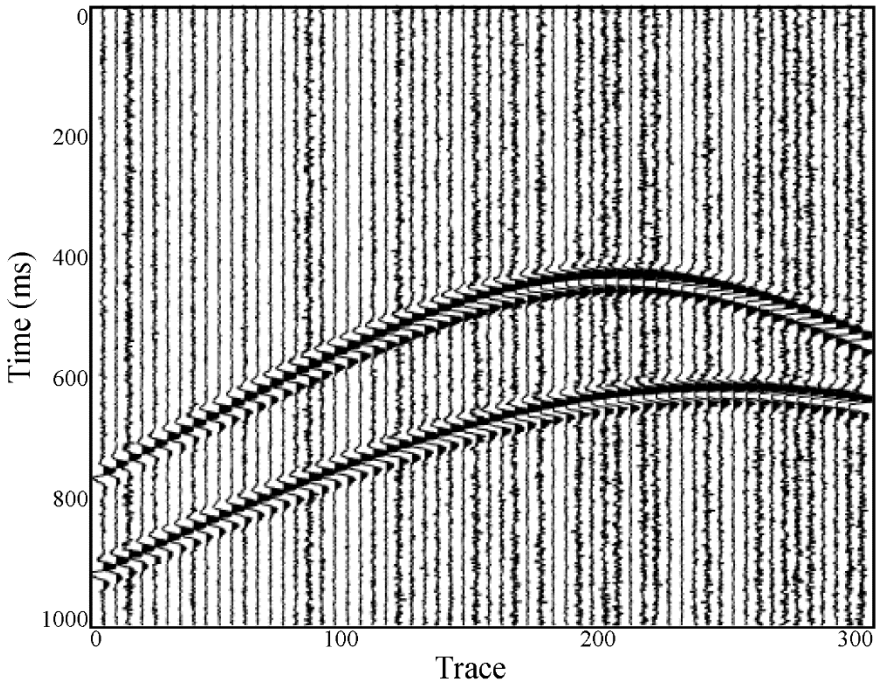


Figure 2.16. Noisy synthetic micro-seismic gather from a different model and the SNR is various from trace to trace.

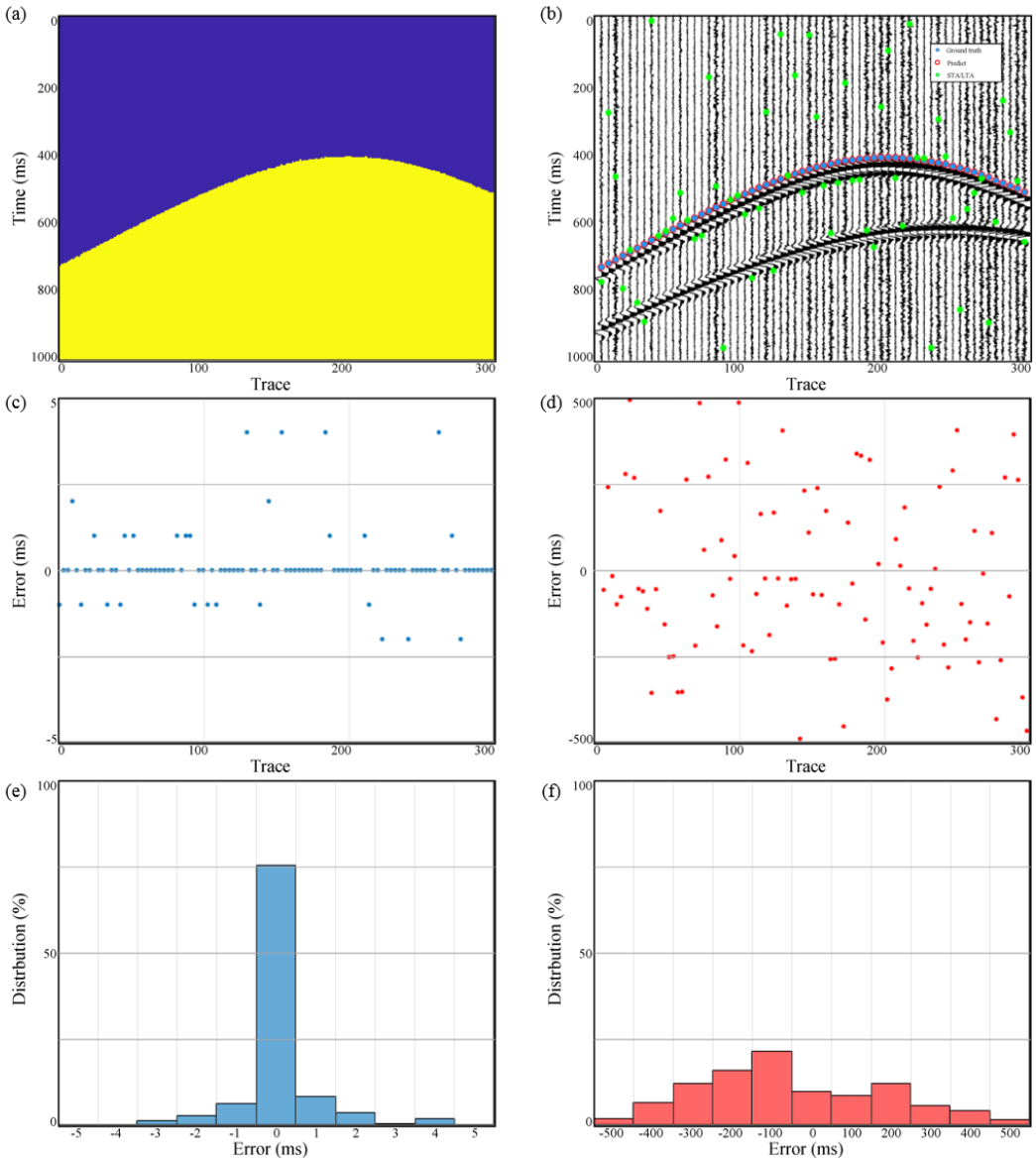


Figure 2.17. The predicted result of a different noisy synthetic example. (a) The predicted label through our proposed method. (b) The ground truth (blue dots), first arrival picking using STA/LTA (green dots) and predicted first arrival (red circles) overlaid the synthetic data. (c) The difference between the ground truth and predicted first arrival. (d) The difference between the ground truth and the picked first arrival using STA/LTA. (e) The distribution of the difference between the ground truth and predicted first arrival. (f) The distribution of the difference between the ground truth and the picked first arrival using STA/LTA.

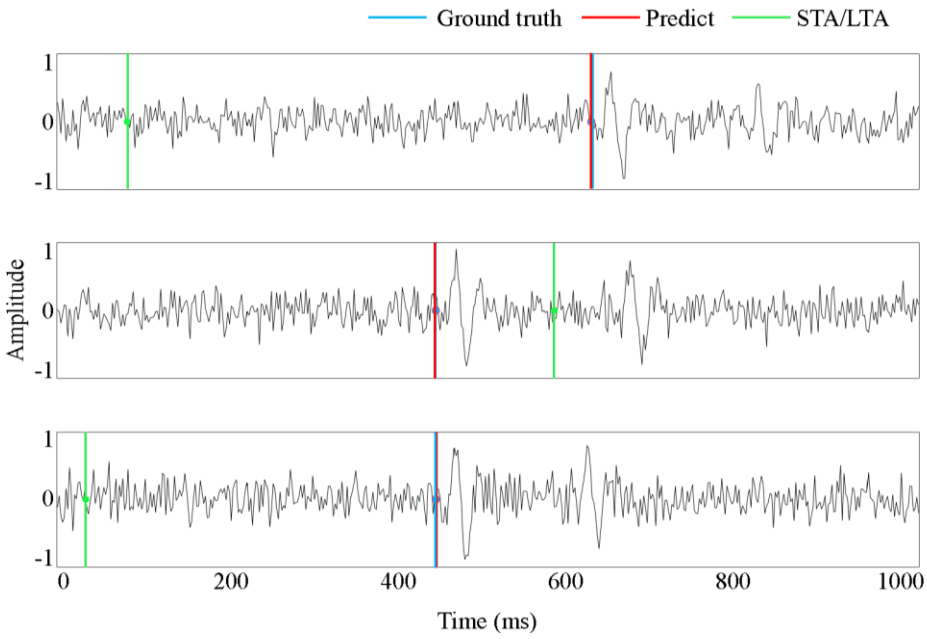


Figure 2.18. Comparison between the predicted and ground truth first arrival for three representative micro-seismic events of a different noisy synthetic data test.

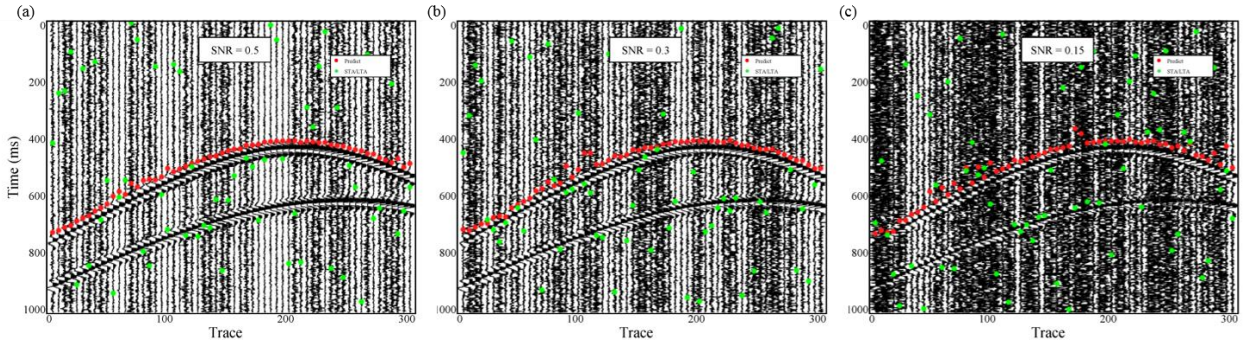


Figure 2.19. The comparisons between the results by using our proposed method (red dots) and STA/LTA (green dots) on three different noisy synthetic example. The average SNR is 0.5 in (a), 0.3 in (b) and 0.15 in (c).

To determine the minimum SNR that our proposed method can handle, we test three extremely noisy synthetic examples. Figures 2.19a, 2.19b and 2.19c show these three synthetic examples with the average SNR of 0.5, 0.3 and 0.15, respectively. The first arrival picking by using our proposed method and STA/LTA are indicated by red dots and green dots, respectively.

Note that the picking result by using our proposed method shown in Figure 2.19c is still acceptable. Thus, the minimum SNR for our proposed method is lower than 0.15, which illustrate the robustness of our method to noise.

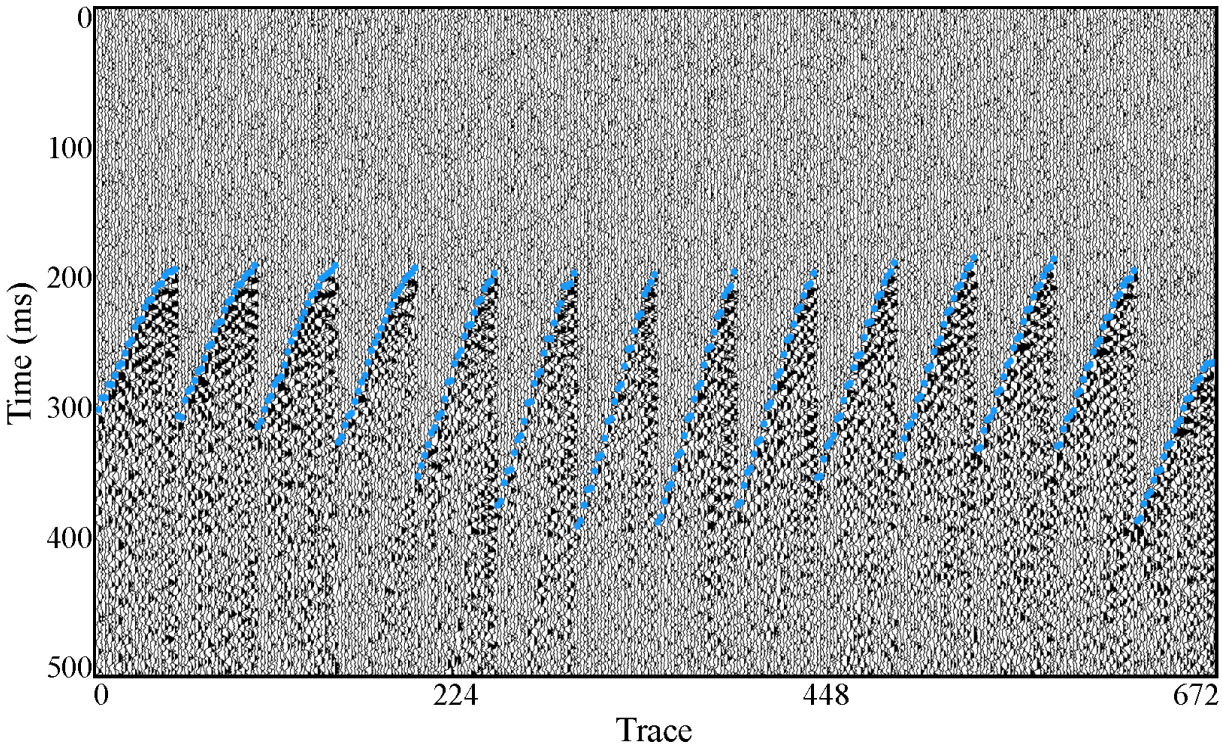


Figure 2.20. Manually interpreted first arrivals (blue dots) of a field data example.

Predict the first arrival of micro-seismic events for field data

To demonstrate the effectiveness and robustness of our proposed method on the field data, we use the real micro-seismic gathers are shown in Figure 2.20. We have 14 gathers (672 traces) in total and the manually interpreted first arrivals are indicated by blue dots in Figure 2.20. To simulate the procedure of manual picking, we choose the first two gathers (96 traces) as the training traces that are shown in Figure 2.21a. Figures 2.21b and 2.21c show the corresponding training data and training label, respectively. We then repeat the procedure of the synthetic example and apply the trained neural network into the rest 12 gathers (576 traces) to predict the first arrivals. Figure 2.22a shows the predicted label by using our proposed method.

Figure 2.22b shows the first arrival picking by using our proposed method (green dots) and STA/LTA (red dots). Our method has a good match with the ground truth (blue dots), while STA/LTA failed to pick the correct first arrivals. Figure 2.22c shows the difference with the ground truth by using our proposed method (blue dots) and STA/LTA (red dots), respectively. The absolute average differences are 1.35*ms* and 136.37*ms* that corresponding to our proposed method and STA/LTA. Figure 2.23 shows the comparisons between the predicted, STA/LTA and ground truth first arrival for three representative micro-seismic events. Our proposed method matches well with the ground truth and STA/LTA does not match.

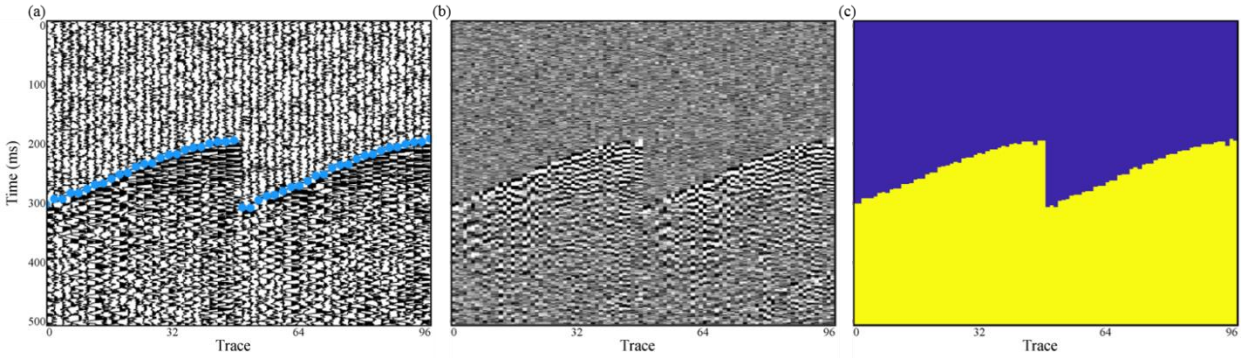


Figure 2.21. Training data and training label for field data. (a) The first two gathers are selected as the training traces. (b) The converted gray image from the training traces. (c) The training labels according to the manual interpretation at the training traces.

For the time cost comparison, the proposed method takes 53.3s while the manual interpretation takes about 300s. The computation is done on a workstation equipped with a 4 core Intel Xeon 4.20 GHz CPU of 64GB memory and an NVIDIA Quadro P5000 GPU of 16 GB memory. Thus, our proposed method can save massive time for the first arrival picking of micro-seismic events.

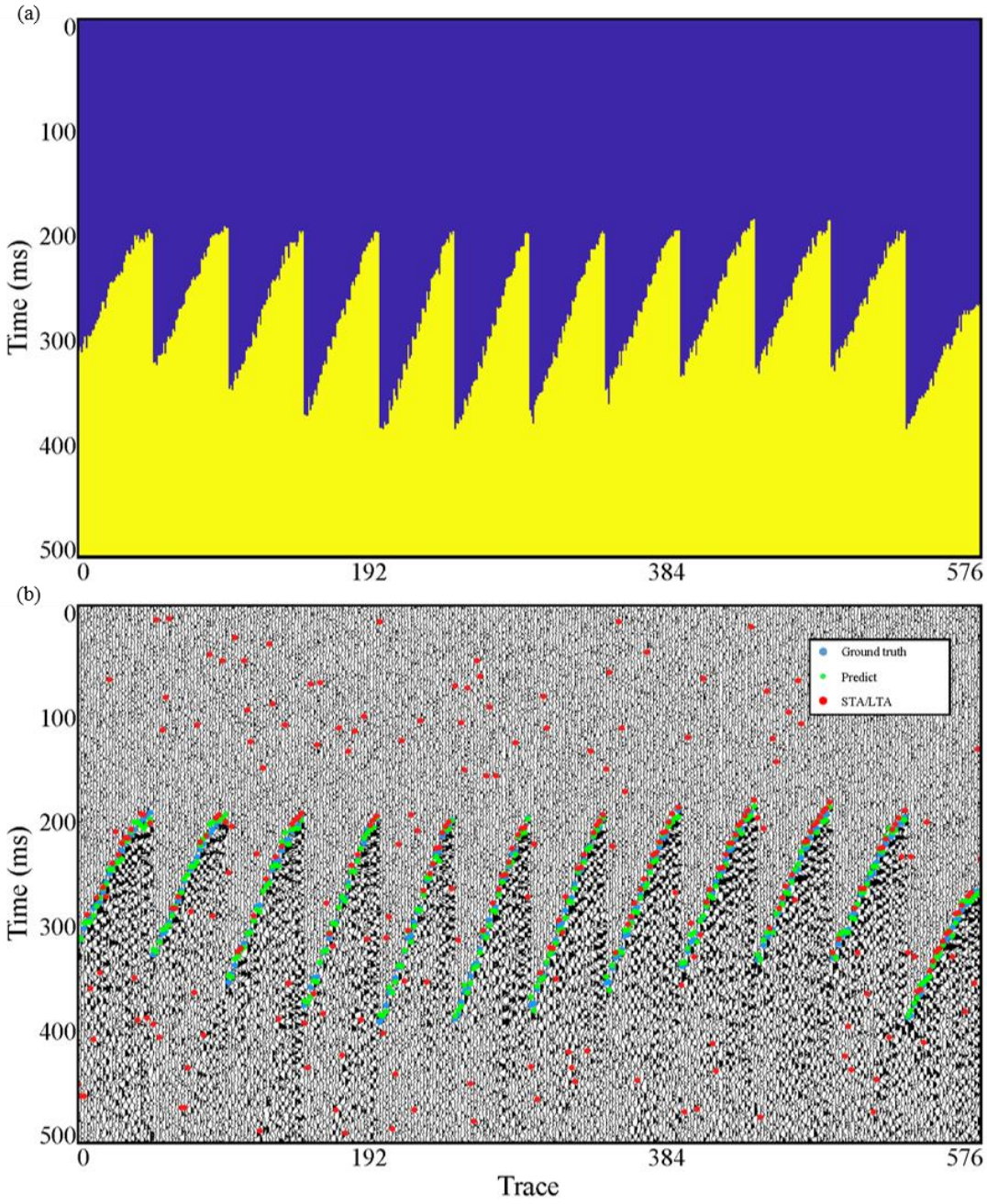


Figure 2.22. The results of the field data example. (a) The predicted label by using the proposed method (b) The results by using our proposed method (green dots) and STA/LTA (red dots) and the manually picked first arrival is indicated by blue dots. (c) The difference between the ground truth and predicted result (blue dots), the difference between the ground truth and the result by using STA/LTA (red dots).

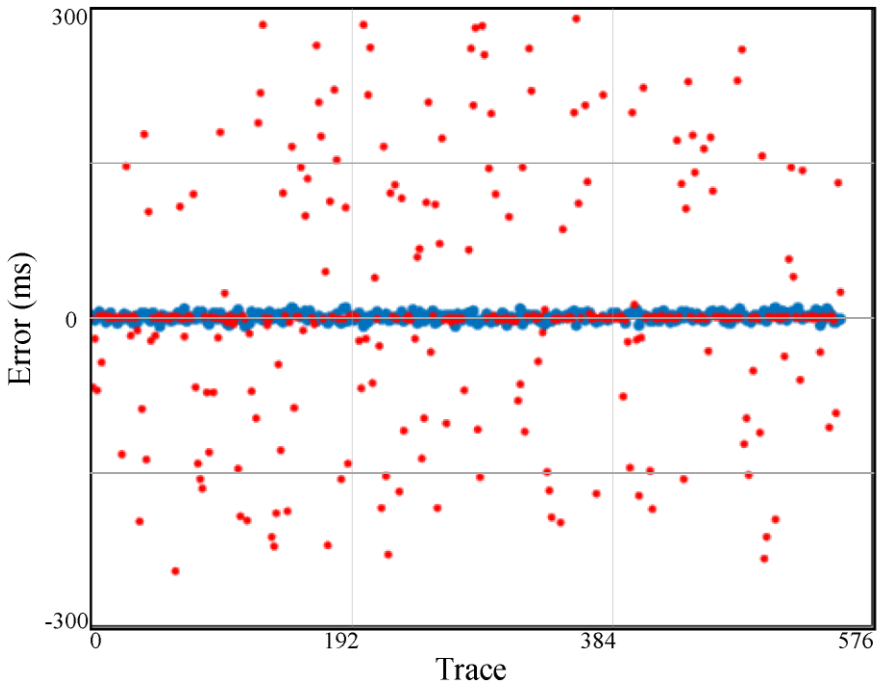


Figure 2.23. The difference between the ground truth and predicted outcome (blue dots), the difference between the ground truth and the result using STA/LTA (red dots) on the field data.

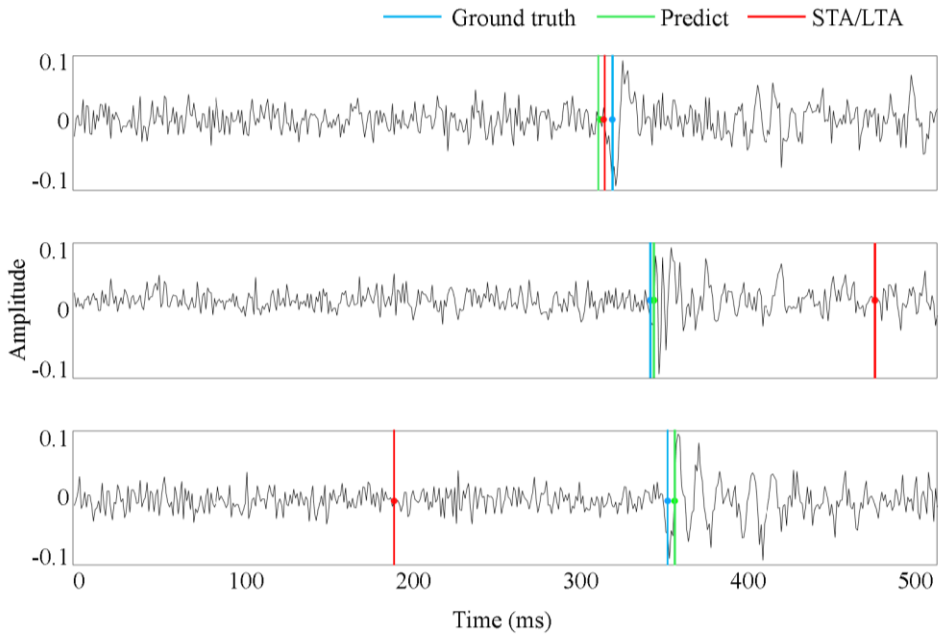


Figure 2.24. Comparison between the predicted and ground truth first arrival for three representative micro-seismic events of real data test.

DISCUSSION

The synthetic and real applications demonstrate that our proposed method has a good performance on first arrival picking of micro-seismic events. To successfully apply our proposed method into the task of micro-seismic first arrival picking, several key factors should be taken into account. The first key factor is the number of layers for building the deep convolutional neural network architecture. The layer number depends on the size of input data and the data size would be reduced by half at each layer. We obtained the best performance when the ratio between the size of the last layer and label number is four. We suggest that the size of the last layer should be between two to six times larger than the label number. The second key factor is the number of training data. In our case, we test 1 gather (7.14%), 2 gathers (14.3%), 3 gathers (21.4%), 4 gathers (28.6%) and 5 gathers (35.7%). We obtain the same level of the predicted result when the number of training gathers is larger than 2. However, our proposed method fails to predict the correct first arrival if the number of training gathers is smaller than 2. Thus, we suggest that the amount of training data should be large than 15% of the whole dataset. In addition, we can select part of the dataset and test to use a different amount of training data prior to applying the proposed method into the whole dataset, then decide how much of the training data is good enough to obtain the correct result. The third key factor is the way of manual interpretation. Because the real data are various from trace to trace. We conclude that either the first peak or first trough can be the first arrival that ensures the neural network is learning all the different features of the first arrival. The fourth key factor is choosing the proper shape of the input data. In our field data example, we test converting the training data from 1D input data to 2D images. However, using 2D images as the input data failed to pick the correct first arrival. It is possible that the number of training data has been reduced and it is easy to cause the problem

of overfitting. Alternatively, the feature of horizontal continues in micro-seismic is not as good as in normal seismic data and it is very hard for CNN to learn the bad continue feature. Thus, we suggest choosing 1D input data for the task of micro-seismic first arrival picking.

CONCLUSIONS

We propose a novel micro-seismic first arrival picking method that is based on the pixel-wise convolutional image segmentation method. The first step is building the training data by randomly selecting part of the micro-seismic traces for manual interpretation. The second step is building the training labels by labeling the selected micro-seismic traces into two parts according to the time index of the manual picking. The third step is building an encoder-decoder convolutional neural network hierarchy by learning the segmented training data and labels. The final step is obtaining the predicted first arrivals of micro-seismic events by applying the trained network hierarchy into the testing data. We modified the size of convolution kernels from constant to downward trend and obtained higher accuracy in the predicted result. Both synthetic and real data applications illustrate that our proposed method is superior to the traditional method of STA/LTA and can save huge labor work than manual interpretation.

REFERENCES

Badrinarayanan, V., A. Kendall, and R. Cipolla, 2015, SegNet: A deep convolutional encoder-decoder architecture for image segmentation, arXiv preprint.

Baer, M., and U. Kradolfer, 1987, An automatic phase picker for local and teleseismic events, Bulletin of Seismology Society of American, vol. 77, pp. 1598-1612.

Chen. Y., 2017, Automatic microseismic event picking via unsupervised machine learning, Geophysical Journal International, vol. 212, pp. 88-102.

Chen, Z., and R. Stewart, 2005, Multi-window algorithm for detecting seismic first arrivals, CREWES.

Girshick, R., J. Donahue, T. Darrell, and J. Malik, 2014, Rich feature hierarchies for accurate object detection and semantic segmentation, IEEE Conference on Computer Vision Pattern Recognition, pp. 580-587.

He, K., X. Zhang, S. Ren, and J. Sun, 2015, Deep residual learning for image recognition, IEEE Conference on Computer Vision Pattern Recognition, pp. 770-778.

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

Kingma, D. P., and J. L. Ba, 2015, Adam: A method for stochastic optimization, International Conference on Learning Representations.

Krizhevsky, A., I. Sutskever, and G. Hinton, 2012, Imagenet classification with deep convolutional neural network, Advances in Neural Information Processing Systems, vol. 27, pp. 1090-1098.

LeCun, Y., L. Bottou, Y. Bengio, and P. Haffner, 1998, Gradient-based learning applied to document recognition, Proceeding of IEEE, vol. 86, no. 11, pp. 2278-2324.

Li. F, J. Rich, and K. Marfurt, 2014, Automatic event detection on noisy microseismogram, SEG expanded abstract, pp. 2363-2367.

Li. F, and W. Song, 2017, Automatic arrival identification system for real-time microseismic event location, SEG expanded abstract, pp. 2934-2939.

Long, J., E. Shelhamer, and T. Darrell, 2015, Fully convolutional networks for semantic segmentation, IEEE Conference on Computer Vision Pattern Recognition, pp. 3431-3440.

Saragiotis, C. D., L. J. Hadjilenontiadis, and S. M. Panas, 1999, A high-order statistics-based phase identification of three-component seismograms in a redundant wavelet transform domain, Proceeding of IEEE Signal Processing Workshop on High-Order Statistics, pp. 396-399.

Song, F., H. S. Kuleli, M. N. Toksoz, E. Ay, and H. Zhang, 2010, An improved method for hydrofracture-induced microseismic event detection and phase picking, Geophysics, vol. 75, no. 6, pp. A47-A52.

Simonyan, K., and A. Zisserman, 2015, Very deep convolutional networks for large-scale image recognition, International Conference on Learning Representations.

Szegedy, C., W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, 2015, Going deeper with convolutions, arXiv preprint.

Wong, J., L. Han, J. Bancroft, and R. Stewart, 2009, Automatic time-picking of first arrivals on noisy microseismic data, CSEG Conference Abstract, vol. 1, pp. 1-4.

Zhang, H., C. Thurber, and C. Rowe, 2003, Automatic p-wave arrival detection and picking with multiscale wavelet analysis for single-component recordings, Bulletin of Seismology Society of American, vol. 93, no. 5, pp. 1904-1912.

CHAPTER 3
WHITE NOISE ATTENUATION OF SEISMIC DATA BY INTEGRATING VARIATIONAL MODE DECOMPOSITION AND CONVOLUTIONAL NEURAL NETWORK

Hao Wu[1], Bo Zhang[1], Tengfei Lin[2], Fangyu Li[3], Naihao Liu[4]
[1]University of Alabama, Department of Geological Science
[2] Department of Middle East E&P, Research Institute of Petroleum Exploration & Development, CNPC,
[3]University of Georgia, College of Engineering.
[4]Xi'an Jiaotong University, School of Electronic and Information Engineering.
This paper was submitted to SEG journal Geophysics in 2019

ABSTRACT

Seismic noise attenuation is an important step in seismic data processing. Most of the noise attenuation algorithms are based on the analysis of time-frequency features of the seismic data and noise. We propose to attenuate the white noise of seismic data using convolutional neural network (CNN). Traditional CNN based noise attenuation algorithms need prior information (the "clean" seismic data or the noise contained in the seismic) in the training process. However, it is tough to obtain such prior information in practice. We assume that enough user-generated white noise realizations can simulate the white noise contained in the seismic data. We then propose to attenuate the seismic white noise using the modified denoising convolutional neural network (MDnCNN). The MDnCNN does not need "clean" seismic data nor the noise contained in the seismic data in the training procedure. To accurately and efficiently learn the features of seismic data and band-limited noise at different frequency bandwidths, we propose to decompose the seismic data into several components before the training procedure of the MDnCNN. There are four main steps in the proposed workflow. The first step is decomposing the seismic volume into different intrinsic mode function and a residual component by using variational mode decomposition. The second step is adding user-generated

white noise to each decomposed component. The third step is building the neural network hierarchy to learn the feature of additive white noise. The last step is obtaining denoised seismic data by applying the well-trained network to the original seismic data. We use both synthetic and field data examples to illustrate the robustness and superiority of the proposed method.

INTRODUCTION

Seismic noise attenuation is a crucial step to enhance the quality of seismic data. Seismic denoising not only lowers the effect of subjectivity in seismic interpretation but also improve the reliability of seismic inversion. In recent decades, numerous seismic denoising approaches have been developed and widely applied in practice. Methods for seismic denoising can be generally classified into four categories: The first category is based on building a prediction filter to remove the noise of seismic data. The commonly used algorithms within the first category includes *f-x* predictive filtering (Canales, 1984), *t-x* predictive filtering (Abma et al., 1995), the forward-backward prediction approach (Wang, 1999), the polynomial fitting-based approach (Liu et al. 2011), and non-stationary predictive filtering (Liu et al., 2012). The second category projects the seismic data to a transformed domain and rejects the noise by applying a bandpass filter to the transformed data. Finally, obtain the denoised seismic data by projecting the filtered data back to the time domain. The commonly used algorithms within the second category include Fourier transform (Chen et al., 2014), curvelet transform (Herrman et al., 2008), seislet transform (Fomel et al., 2010), shearlet transform (Kong et al., 2015), Radon transform (Trad et al., 2002; Xue et al., 2016), wavelet transform (Donoho et al., 1994), and dictionary learning based sparse transform (Elad et al., 2006). The third category decomposes the seismic traces into a set of components. Then, examine the time or frequency features of each decomposed component. Finally, obtain the "clean" seismic traces by rejecting the components which are regarded as "noise". The commonly used algorithms within the third category include empirical mode decomposition (EMD) (Huang et al., 1998; Bekara et al., 2009), variational mode decomposition (VMD) (Dragomiretskiy et al., 2014; Li et al., 2017), and singular value decomposition (SVD) based approaches (Bekara et al., 2007). Yuan et al. (2018) proposed a novel inversion-based

denoising method. The method has the advantages of preserving 3-D spatial edges and low-frequency signals. The fourth category is based on the rank-reduction reconstruction of seismic data. The commonly used algorithms within the fourth category include Cadzow filtering (Trickett, 2008), and singular spectrum analysis (Vautard et al., 1992; Oropeza et al., 2011).

Deep learning is a subset of machine learning that is based on learning data representation. Convolutional neural network (CNN) (LeCun et al., 1998) is one of the most popular and widely used deep learning algorithms. The CNN based algorithms already achieve great success in the field of computer vision. CNN is extremely efficient in learning the features of the images and labeling the objectives in the images. Numerous CNN based algorithms also have been proposed to address the problem of image denoising. Jain et al. (2009) successfully applied CNN to images denoising. Burger et al. (2012) denoise the images using the multiple layer perceptron (MLP). Other popular CNN based image denoising methods include stacked sparse denoising autoencoder (Xie et al., 2012) and trainable nonlinear reduction diffusion (TNRD) model (Chen et al., 2015). Zhang et al. (2017) proposed denoise convolutional neural network (DnCNN) to learn the feature of noise contained in the images. The main disadvantage of current CNN-based denoising methods is that these methods need "clean" data and the corresponding noisy data in the training process. Unfortunately, it is unfeasible to obtain "clean" seismic data for training in practice.

We organized this paper as follow: we first describe our modification for DnCNN. We then introduce the workflow of noise attenuation. We finally use both synthetic and field data examples to illustrate the robustness and superior performance of VMD- MDnCNN over *f-x* deconvolution.

THEORY

There are many successful applications of images denoising by using CNN based algorithms (Xie et al., 2012; Zhang et al., 2017). The main advantage of CNN based denoising methods is that CNN with multiple hidden layers (deep architecture) can recognize various features of the input data and classify the recognized features into corresponding categories.

Objective function

One 3D seismic volume can be reshaped to many 2D seismic sections through the inline and crossline directions. Each seismic section can be treated as a 2D image. The image which needs to be denoised can be defined as $\boldsymbol{y} = \boldsymbol{x} + \boldsymbol{n}$, where $\boldsymbol{y}$ is the noisy image, $\boldsymbol{x}$ is the corresponding clean image, and $\boldsymbol{n}$ is the additive noise. The goal of image denoising is building a model to recover the clean image $\boldsymbol{x}$ from the corresponding noisy image $\boldsymbol{y}$. According to the objective function, the image denoising methods using CNN can be classified into two categories.

The first category is modeling the "clean" image (Jain et al., 2009; Xie et al., 2012) by minimizing the following objective function $\boldsymbol{J}(\theta)$:

$$\boldsymbol{J}(\boldsymbol{\theta}) = \arg\min_{\theta} \frac{1}{M} \sum_{i=1}^{M} \left\| \boldsymbol{x}_i - R_\theta(\boldsymbol{y}_i) \right\|^2 \tag{1}$$

where $\{(\boldsymbol{y}_i, \boldsymbol{x}_i)\}_{i=1}^{M}$ represents M noisy-clean image pairs, R_θ denotes the entire convolutional neural network with all the trainable parameters (convolution filter and bias) $\boldsymbol{\theta}$, $R_\theta(\boldsymbol{y}_i)$ is the predicted clean image by using the trained convolutional neural network R_θ.

The second category is modeling the "noise" by applying the residual learning formulation (Zhang et al., 2017):

$$\boldsymbol{J}(\boldsymbol{\theta}) = \arg\min_{\theta} \frac{1}{M} \sum_{i=1}^{M} \left\| R_\theta(\boldsymbol{y}_i) - (\boldsymbol{y}_i - \boldsymbol{x}_i) \right\|^2 \tag{2}$$

where $R_\theta(\boldsymbol{y}_i)$ is the predicted noise by using the trained convolutional neural network R_θ.

In geophysics field, $\boldsymbol{x}$ and $\boldsymbol{y}$ can be regarded as the noise-free seismic images and noise contaminated seismic images, respectively. Both equations 1 and 2 require clean and the corresponding noisy data in the training process. Unfortunately, it is unfeasible to obtain purely clean seismic data in practice. However, we assume that the white noise contained in the seismic data can be simulated by enough user-generated white noise realizations $\boldsymbol{n}'$ (Wu and Huang, 2009). The seismic image with additive white noise can be expressed as

$$\boldsymbol{y}' = \boldsymbol{x} + \boldsymbol{n} + \boldsymbol{n}' \tag{3}$$

The new objective function of MDnCNN is given by:

$$\boldsymbol{J}(\boldsymbol{\theta}) = \arg\min_{\theta} \frac{1}{M} \sum_{i=1}^{M} \left\| R_\theta(\boldsymbol{y}_i) - (\boldsymbol{y}_i' - \boldsymbol{y}_i) \right\|^2 \tag{4}$$

$$\boldsymbol{J}(\boldsymbol{\theta}) = \arg\min_{\theta} \frac{1}{M} \sum_{i=1}^{M} \left\| R_\theta(\boldsymbol{y}_i) - \boldsymbol{n}_i' \right\|^2 \tag{5}$$

According to the statistical properties of white noise, the distribution of the original white noise $\boldsymbol{n}$ and the additive white noise $\boldsymbol{n}'$ are given by (Wu and Huang, 2004):

$$\boldsymbol{n}_i \sim \chi\left(\mu_1, {\sigma_1}^2\right), \boldsymbol{n}_i' \sim \chi\left(\mu_2, {\sigma_2}^2\right), \tag{6}$$

where χ denotes the normal distribution, μ_1 is the expectation and σ_1 is the standard deviation of the noise contained in the seismic data, μ_2 is the expected value and σ_2 is the standard deviation of the additive white noise. Equation 6 illustrates that we can simulate the noise contained in the seismic data if we have enough trials. To accurately simulate the noise contained in the seismic data, the noise level of additive white noise should be close to the noise level of the original white noise. We employ the peak signal noise ratio (PSNR) method to compute the signal to noise ratio (SNR) for the input seismic data:

$$SNR \approx PSNR = \frac{\max\left(\boldsymbol{y}^2\right)}{MSE}, \tag{7}$$

$$MSE = \frac{1}{M}\sum_{i}^{M}(\boldsymbol{y}_i - \hat{\boldsymbol{y}})^2, \tag{8}$$

where $\hat{\boldsymbol{y}}$ denotes the mean value of the original seismic data and MSE is the mean squared error (Li et al., 2017).

After adding enough additive white noise realizations, the expectation μ_2 and standard deviation $\sigma_2$of certain realization of simulated noise should approximately equal to the expectation $\mu_1$and standard deviation σ_1 of noise contained in the seismic data (Wu and Huang, 2004):

$$\mu_2 \approx \mu_1,\ \sigma_2 \approx \sigma_1 \tag{9}$$

Then, we obtain:

$$\boldsymbol{J}(\boldsymbol{\theta}) = \arg\min_{\theta} \frac{1}{N}\frac{1}{M}\sum_{i=1}^{M}\sum_{k=1}^{N}\left\|R_\theta(\boldsymbol{y}_i) - \boldsymbol{n}'_{ik}\right\|^2 \approx \arg\min_{\theta} \frac{1}{M}\sum_{i=1}^{M}\left\|R_\theta(\boldsymbol{y}_i) - \boldsymbol{n}_i\right\|^2, \tag{10}$$

where N is the number of additive white noise realizations, $\boldsymbol{n}'_{ik}$ represent the i^{th} additive white noise image of the k^{th} additive white noise realizations. Equation 10 indicates that the proposed method does not require clean seismic data in the training process.

Architecture

The architecture of the proposed neural network is a sequence of nonlinear processing layers followed by a sigmoid classifier layer based on the architecture of MDnCNN (Figure 3.1). The input of the network is the original seismic data $\boldsymbol{y}$ and the seismic data with additive white noise $\boldsymbol{y}' = \boldsymbol{y} + \boldsymbol{n}'$. The network contains 17 layers in total. The first layer contains 64 convolution filters of size 3×3 and 64 rectified linear units (ReLU) activation operator. The objective of the convolution filter is to generate feature maps of the input seismic data. The objective of the ReLU is to activate the main features contained in the feature map. Different

from the first layer, a batch-normalization (BN) (Ioffe and Szegedy, 2015) is added between the convolution filter and ReLU for the following 2~16 layers. Batch-normalization is a re-parametrization aimed to stabilize the parameters updating and improve the learning process. The last layer only contains 64 convolution filters of size 3×3 to reconstruct the output. Then the built neural network transforms the seismic noise attenuation procedure into an optimization problem by solving a sequence of nonlinear functions. A gradient-based optimization algorithm of adaptive moment estimation (Adam) (Kingma et al., 2015) is employed to minimize the proposed objective function through iterative updating the parameters of the network.

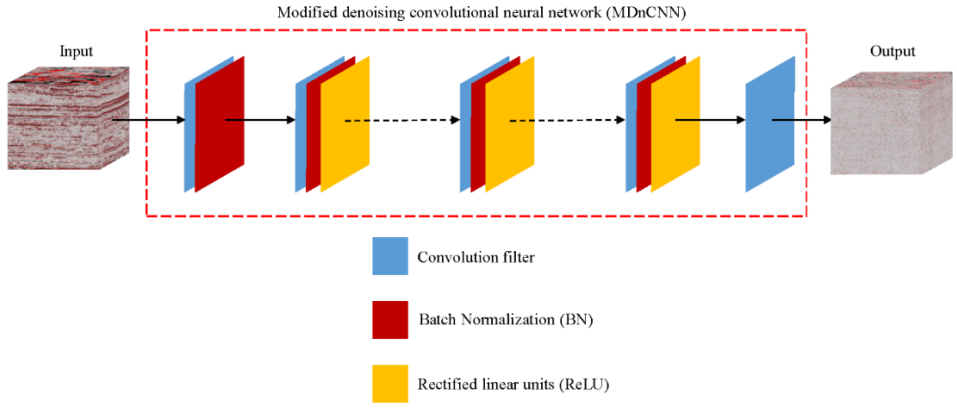


Figure 3.1. The overall architecture of MDnCNN. It contains 17 layers and each layer includes convolution filter (Conv), Batch Normalization (BN) and Rectified linear units (ReLU).

VMD-MDnCNN

Figure 3.2 shows a real seismic section of F3-block in the North Sea, Netherlands. Figures 3.3a and 3.3b show the denoised result and the rejected noise by using the MDnCNN. Note the noise indicated by the yellow arrows in Figure 3.3a and rejected visible reflections indicated by the red arrows in Figure 3.3b. The main reason for this phenomenon is that that seismic data are band-limited and the noise to signal ratio (SNR) is varying with bandwidths. If we only use the MDnCNN to learn the feature of original white noise at full bandwidth, the

MDnCNN cannot attenuate all the original white noise and will remove some visible reflections. To better learn the noise feature, we propose first to decompose the seismic data into different components and then apply the MDnCNN to each decomposed component. In this work, we use the variational mode decomposition (VMD) to decompose the seismic traces. We named the proposed seismic noise attenuation procedure as VMD- MDnCNN.

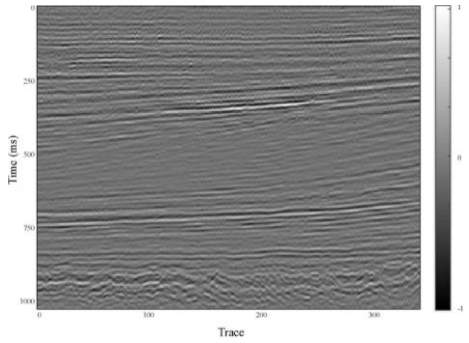


Figure 3.2. One section example of F3-block seismic survey.

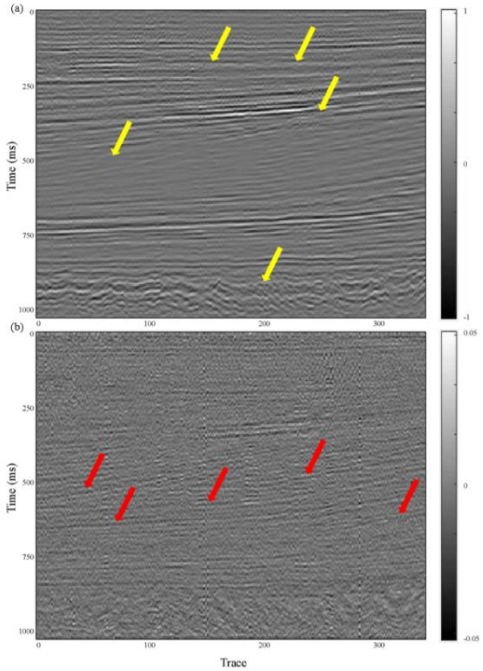


Figure 3.3. Illustration of the test on F3-block using MDnCNN. (a) the denoised result. (b) the rejected noise.

VMD is an adaptive and non-recursive signal decomposition method (Dragomiretskiy and Zosso, 2013). VMD decompose a signal into a series of intrinsic mode function (IMF) where IMFs have the sparsity properties in the frequency domain. The frequency spectrum of every IMF is computed around the center frequency ω_i and the sparsity of every IMF is constrained by its bandwidth in the frequency domain. In other words, VMD decomposes the signal into different IMFs and the frequency spectrum of each component tuned around the center frequency ω_i. We obtain each IMF by recursively solving the following optimization problem:

$$\min_{\{u_i\},\{\omega_i\}} \left\{ \sum_i \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * u_i(t) \right] e^{-j\omega_i t} \right\|_2^2 \right\}, \qquad (11)$$
$$\sum_i^L u_k = s(t)$$

where $u_i$and $\omega_i$are modes and their center frequencies, respectively, $\delta(t)$is a Dirac impulse, $s(t)$is the signal to be decomposed, the constraint condition is that the summation over all modes should be the input signal, the term $\left(\delta(t) + \frac{j}{\pi t}\right) * u_i(t)$is the Hilbert transform of u_i, and the parameters L is the user defined decomposed number.

Then, the denoising objective function, $\boldsymbol{J}(\boldsymbol{\theta})^{(j)}$, for the j^{th} decomposed seismic component is given as:

$$\boldsymbol{J}(\boldsymbol{\theta})^{(j)} = \arg\min_{\theta} \frac{1}{N} \frac{1}{M} \sum_{i=1}^{M} \sum_{k=1}^{N} \left\| R_\theta^{(j)} \left(\boldsymbol{y}_i^{(j)} \right) - \left(\boldsymbol{y}_{ik}^{(j)} - \boldsymbol{y}_i^{(j)} \right) \right\|^2, \qquad (12)$$

where $\boldsymbol{y}^{(j)}$ and $\boldsymbol{y'}^{(j)}$ represent the j^{th} decomposed component and the components with additive white noise, respectively.

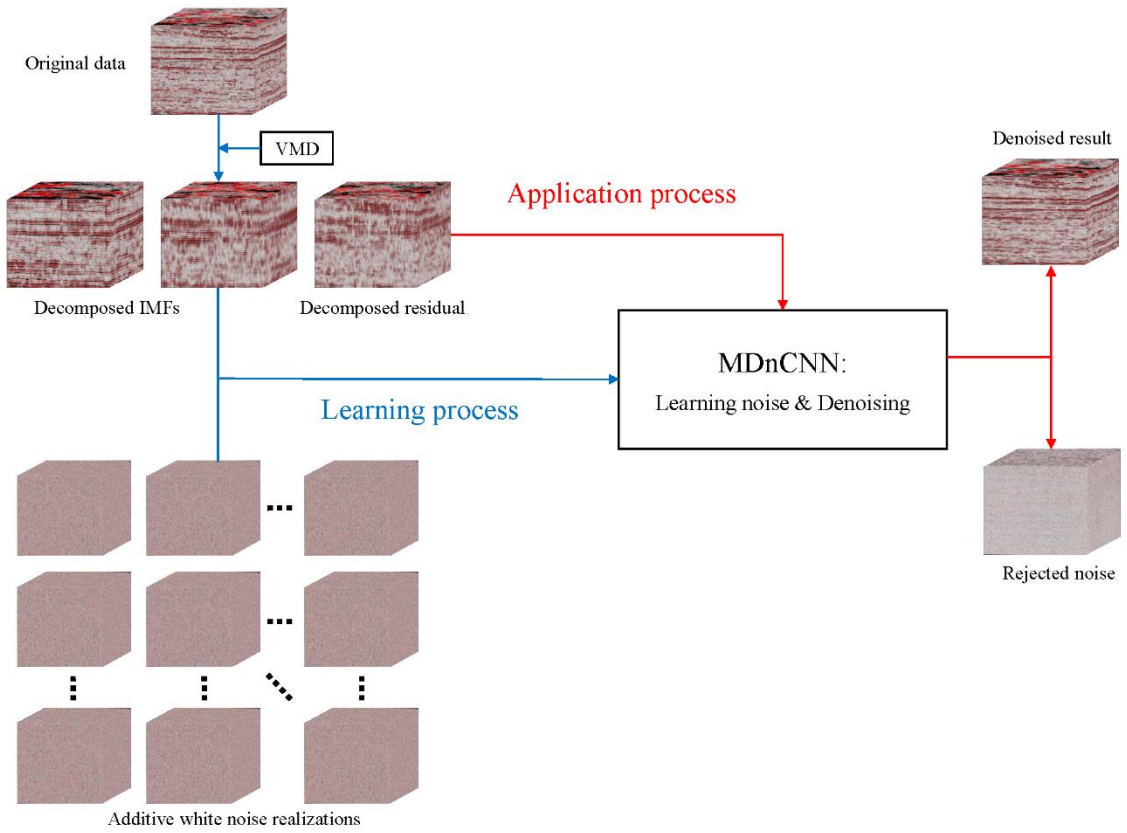


Figure 3.4. The workflow of VMD-MDnCNN for seismic noise attenuation.

Figure 3.4 shows the proposed denoising workflow using VMD-MDnCNN. We first decompose the seismic data (the 3D seismic data have been reshaped to 2D seismic sections in both the inline and crossline directions) into several decomposed components and compute the SNR for each component. We produce a "noisier" IMF by adding additive white noise to each decomposed IMF and the residual component. The energy of the additive white noise approximately equals to the energy of the white noise estimated within each decomposed component. We next learn the feature of white noise by minimizing the difference between additive noise and learned white noise from the "noisier" decomposed components (Equation 12). We produce the denoised components by subtracting the learned noises from the corresponding decomposed components. We finally obtain the denoised seismic data by

integrating the ensemble of denoised components. Figures 3.5a and 3.5b show the denoised results and rejected noise by using VMD-MDnCNN. Note that our method successfully rejects the noise indicated by the yellow arrows in Figure 3.3a and preserves the seismic reflections indicated by the red arrows in Figure 3.3b.

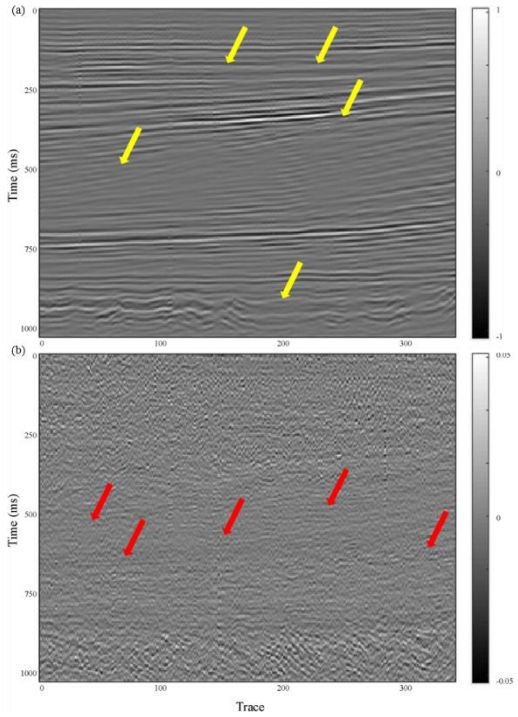


Figure 3.5. Illustration of the test on F3-block using VMD-MDnCNN. (a) the denoised result. (b) the rejected noise.

SYNTHETIC EXAMPLE

To demonstrate the performance of VMD-MDnCNN, we first test our method using synthetic seismic data (Figure 3.6a) generated using the Marmousi model. We use zero-phase Ricker wavelet to generate our synthetic seismic data. The dominant frequency of the Ricker wavelet is 30 Hz. The synthetic seismic data contains 128 traces. Each trace has 128 time samples and the time sample interval is 4ms. Figure 3.6b shows the noisy synthetic seismic data. The additive noise is Gaussian noise and the SNR is 2. To ensure that the additive noise has the

same frequency bandwidth with the seismic data, we applied a band-pass Butterworth filter (5-10-95-100Hz) to the Gaussian noise before we add the noise to the noise-free synthetic seismic data. Figures 3.7a, 3.7b and 3.7c show the denoised seismic data using *f-x* deconvolution, MDnCNN and VMD-MDnCNN, respectively. Figures 3.8a, 3.8b and 3.8c show the rejected noise using *f-x* deconvolution, MDnCNN and VMD-MDnCNN, respectively. Note that the VMD-MDnCNN not only rejects the white noise (the yellow arrows in Figures 3.7a, 3.7b and 3.7c) but also preserves the visible reflections rejected by *f-x* deconvolution and MDnCNN (the red arrows in Figure 3.8a, 3.8b and 3.8c).

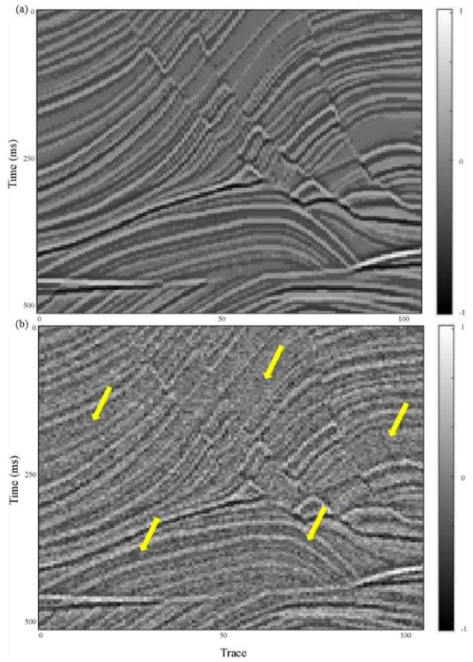


Figure 3.6. We add gaussian noise with the SNR is equal to 2 into the noise-free synthetic data to generate the noisy synthetic data. (a) The noise-free synthetic data. (b) The noisy synthetic data.

In this work, we found that that there is no obvious difference between the denoised results both for the synthetic and real seismic data if the IMFs number is equal or greater than 2. The computation cost

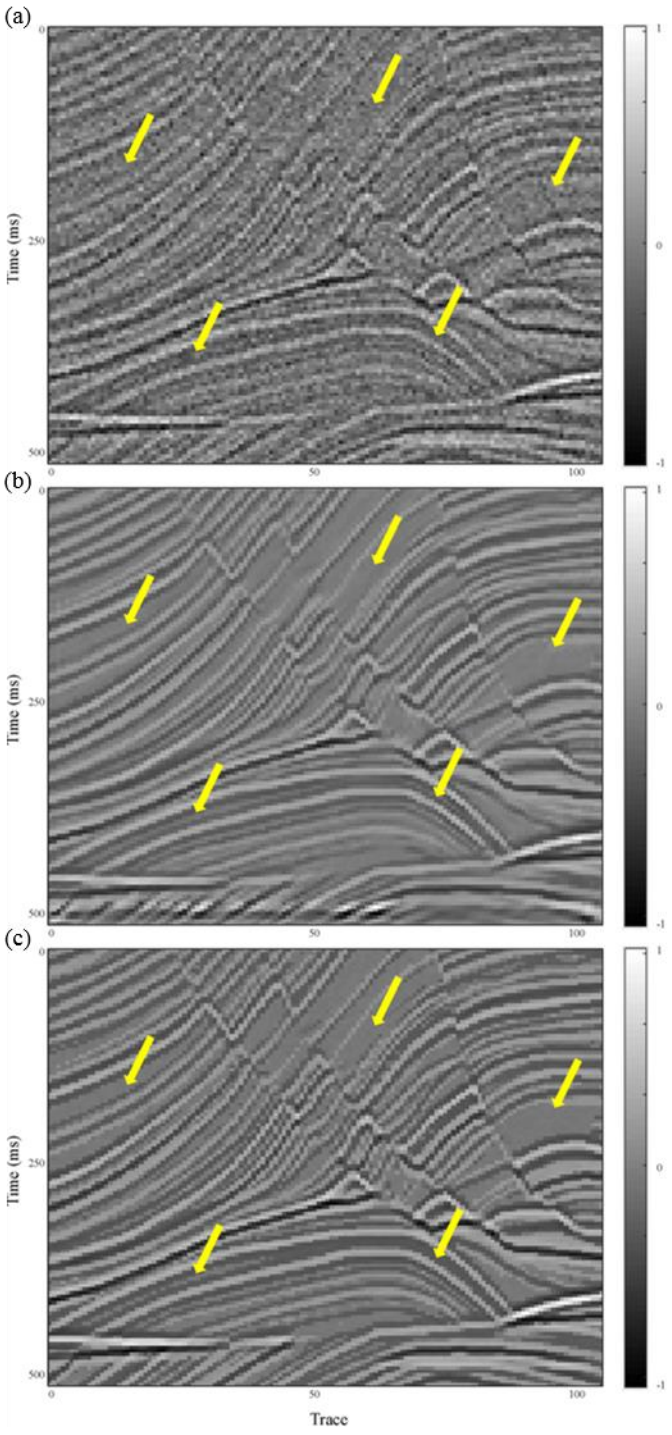


Figure 3.7. Illustration of the denoised result on synthetic example. (a) *f-x* deconvolution. (b) MDnCNN. (c) VMD-MDnCNN.

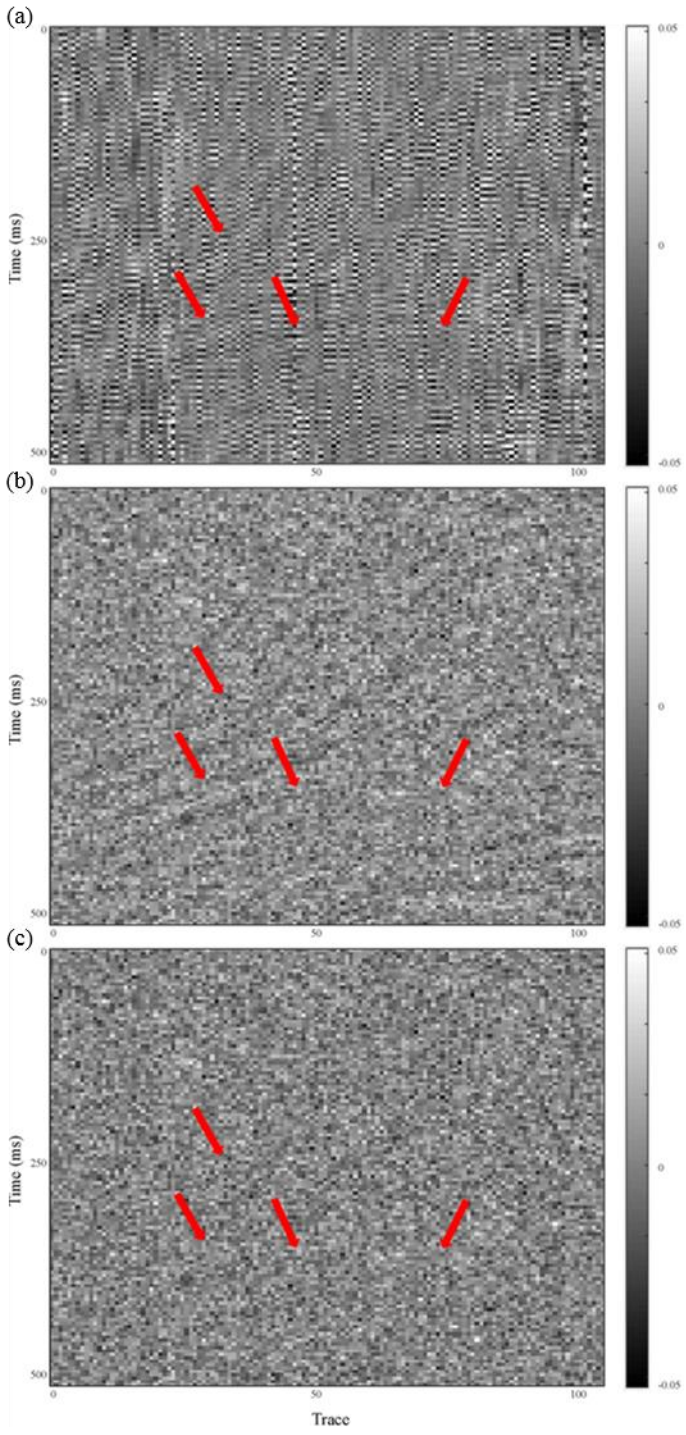


Figure 3.8. Rejected noise through different denoising methods on synthetic example. (a) *f-x* deconvolution. (b) MDnCNN. (c) VMD-MDnCNN.

increases with the increasing of IMFs number. We set the IMFs number as 2 in our synthetic testing. The moderate bandwidth constraint and the tolerance of convergence criterion are 100 and 0.01 for VMD decomposition, respectively. The pre-set center frequencies for these two IMFs are 25 and 45Hz, respectively. The input for our VMD-MDnCNN is the first decomposed IMF plus additive noise, the second decomposed IMF plus additive noise, and the residual component plus additive noise. Figure 3.9 shows the training and validation loss varying with optimization epochs. Considering that the data size used in deep learning is enormous, we usually divide the learning data set into several small subsets (batch). The optimization procedure is implemented batch by batch and one epoch means one optimization pass of the full batches. We obtain the training and validation loss by applying the objective function shown in Equation 12 to the training and validation data set, respectively. To overcome the overfitting problem in the training procedure, the training and validation seismic traces are randomly selected during each optimization epoch. The percentage of the training and validation seismic traces in this paper is 70% and 30% both for the synthetic and real seismic data, respectively. A specific seismic trace may belong to training seismic traces set in current optimization epoch but may belong to validation seismic traces set in the next optimization epoch. Figure 3.9 illustrates that we obtain a stable neural network hierarchy after 50 epochs in the training procedure.

Figure 3.10 shows the average amplitude spectrum of the original seismic (black), the denoised result using *f-x* deconvolution (red), the denoised result using MDnCNN (blue) and the denoised result using VMD-MDnCNN (green), respectively. Note that the amplitude spectrum of the denoised result using VMD-MDnCNN has a perfect match with that of original seismic data. Unfortunately, the denoised result using *f-x* deconvolution and MDnCNN lost certain middle and high-frequency content when compared that of original seismic data.

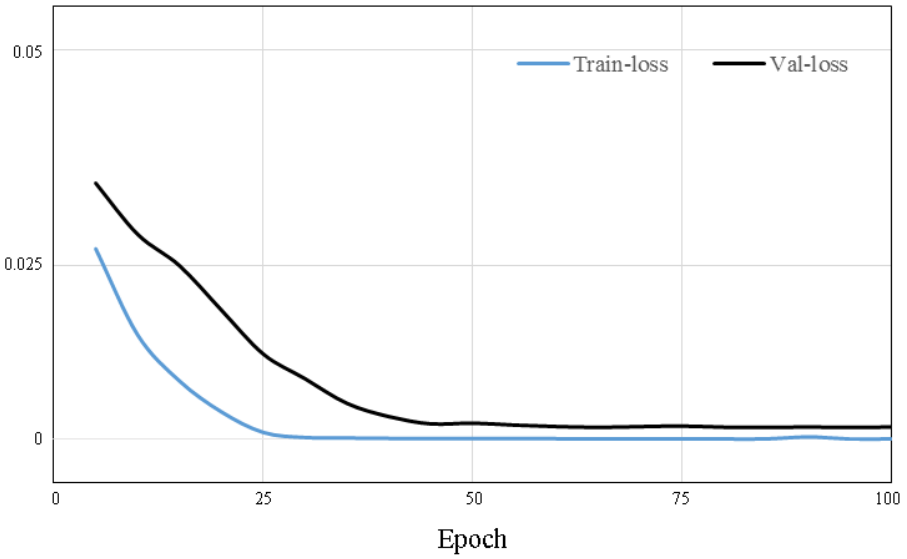


Figure 3.9. Train loss and valid loss varying with different epoch for synthetic example.

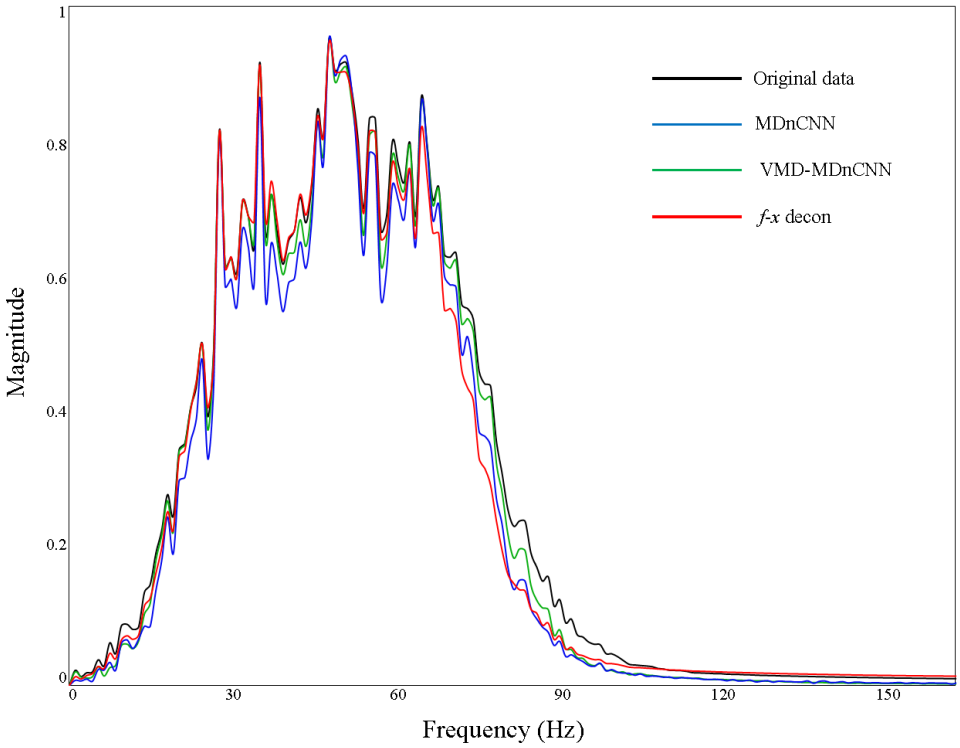


Figure 3.10. The frequency spectrum of the original data (black curve), MDnCNN result (blue curve), VMD-MDnCNN result (green curve) and *f-x* deconvolution result (red curve) for the synthetic test.

FIELD DATA EXAMPLE

We further apply VMD-MDnCNN to a public seismic survey (Penobscot) to illustrate the effectiveness of our proposed method. The Penobscot seismic survey was acquired over the Scotian shelf, oversea Canada. The seismic survey contains 601 inlines and 482 crosslines. The time increment of the seismic survey is 4ms. We observe both residual noise and possible migration artifacts indicated by the yellow arrows in Figure 3.11.

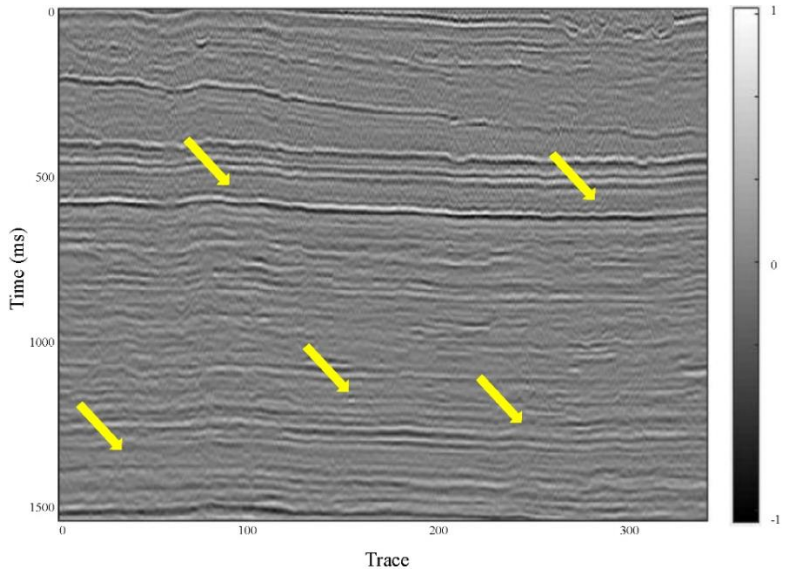


Figure 3.11. One representative inline section example of Penobscot.

We first generate a large number of band-limited additive white noise realizations. The bandwidth of the additive white noise obtained according to the frequency spectrum of the seismic data. We then use VMD to decompose the original seismic data into two IMFs and a residual volume. Based on the testing of real data application, we found that we can successfully simulate the noise contained in the seismic data if the number of white noises realizations is above 2000. In this work, we choose 2000 as the number of white noise realizations. The simulated noise is then added to the decomposed two IMFs and a residual component. Figure 3.12 illustrates that we obtain a stable neural network after 60 epochs in the training procedure.

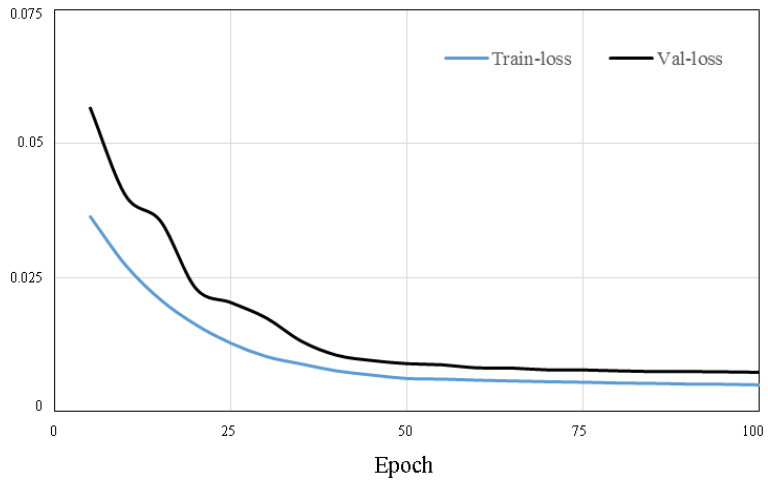


Figure 3.12. Train loss and valid loss varying with different epoch for the seismic survey of Penobscot.

Figures 3.13a, 3.13b and 3.13c show the denoised results using *f-x* deconvolution, MDnCNN and VMD-MDnCNN, respectively. Figures 3.14a, 3.14b and 3.14c show the difference between the original seismic data and denoised results using *f-x* deconvolution, MDnCNN and VMD-MDnCNN, respectively. Note that the VMD-MDnCNN not only reject the white noise and migration artifact indicated by yellow arrows in Figure 3.13, but also preserves the visible reflections rejected by the *f-x* deconvolution and MDnCNN indicated by red arrows in Figure 3.14. Figures 3.15, 3.16a, 3.16b and 3.16c show the 3D cube of original seismic data, denoised seismic data using *f-x* deconvolution, MDnCNN and VMD-MDnCNN, respectively Figures 3.17a, 3.17b and 3.17c show the 3D rejected noise using *f-x* deconvolution, MDnCNN and VMD-MDnCNN, respectively. Note that the VMD-MDnCNN have successfully rejected most of the white noise and migration artifacts indicated by yellow arrows in Figure 3.15. However, the denoised results using *f-x* deconvolution and the MDnCNN still contain white noise and migration artifacts indicated by yellow arrows in Figure 3.16. Figures 3.17a, 3.17b and 13.7c indicates that our

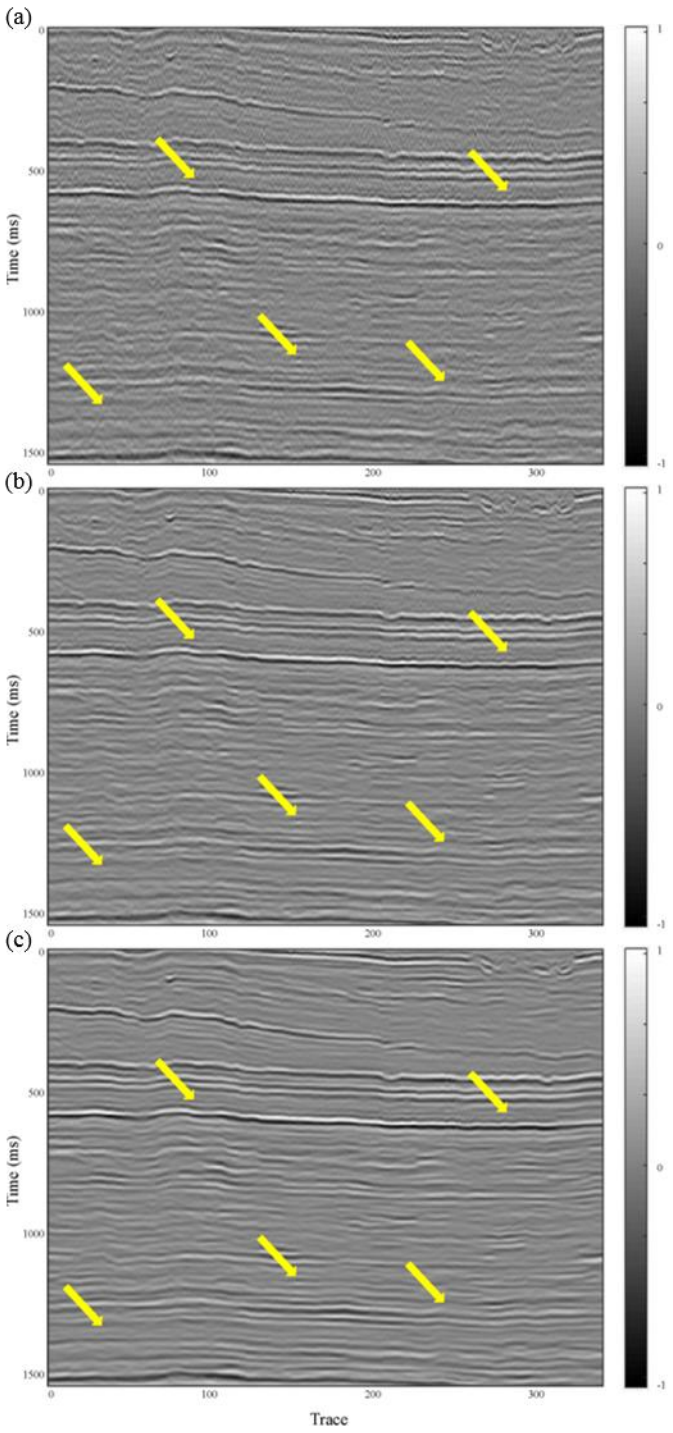


Figure 3.13. The denoised results on the seismic survey of Penobscot. (a) *f-x* deconvolution. (b) MDnCNN. (c) VMD-MDnCNN.

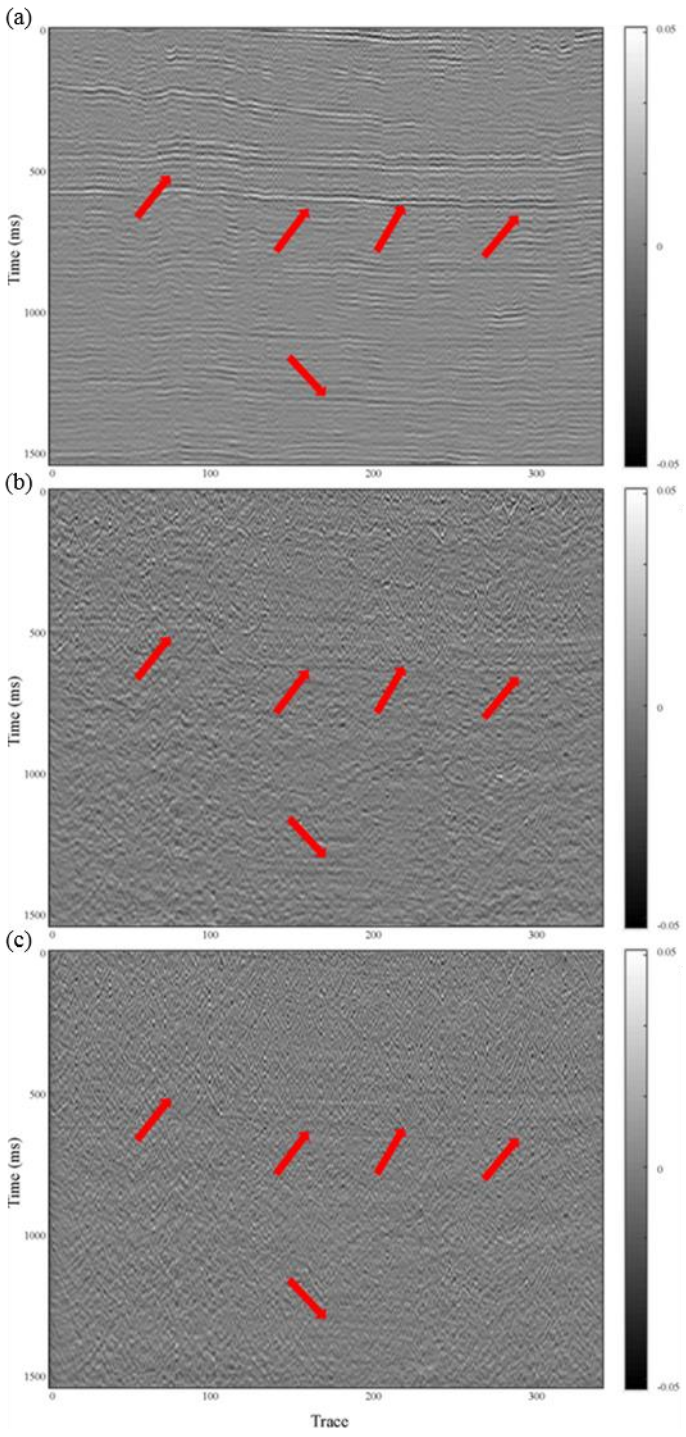


Figure 3.14. Rejected noises through different denoising methods on the seismic survey of Penobscot. (a) *f-x* deconvolution. (b) MDnCNN. (c) VMD-MDnCNN.

proposed method not only attenuate the white noise but also preserve most of the useful seismic amplitude indicated by red arrows. Here the SNR in MDnCNN is 26.1, the SNR in VMD-MDnCNN is 28.3 and the SNR in original data is 12.7. Thus, the contribution of MDnCNN for denoising is about 79.5%.

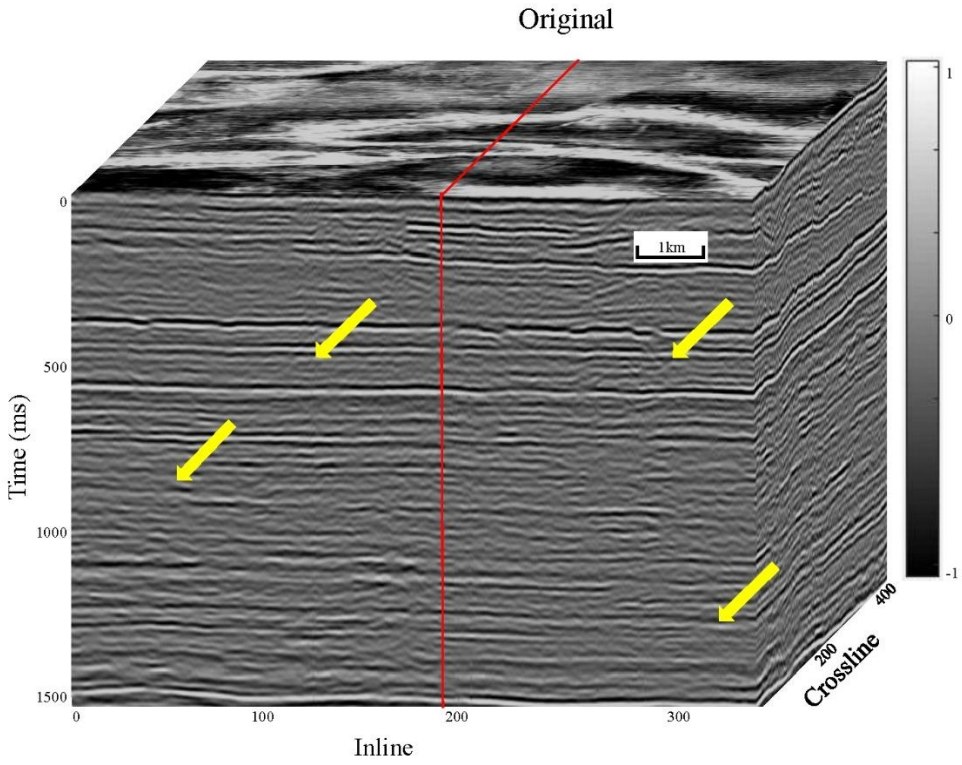


Figure 3.15. The original 3D seismic data of Penobscot.

Again, we further compare the spectrum of the original and denoised seismic data to show the effectiveness of our method. Figure 3.18 shows the amplitude spectrum of the original seismic (black), the denoised result using *f-x* deconvolution (red), MDnCNN (blue) and the VMD-MDnCNN (green), respectively. Note that the average amplitude spectrum of denoised data using VMD-MDnCNN has a very good match with that of original seismic data. Unfortunately, the denoised result using *f-x* deconvolution and MDnCNN lost certain middle and high-frequency content when compared that of original seismic data.

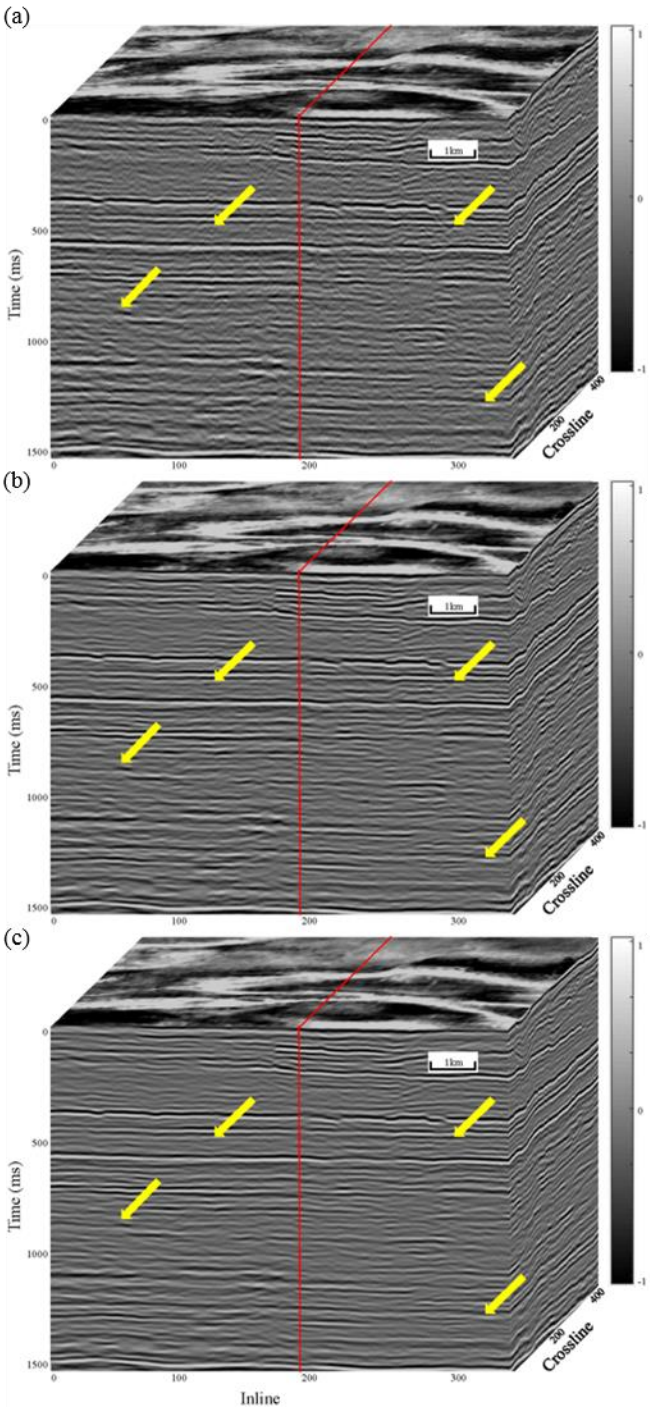


Figure 3.16. Illustration of the 3D volume denoised result on the seismic survey of Penobscot. (a) *f-x* deconvolution. (b) MDnCNN. (c) VMD-MDnCNN.

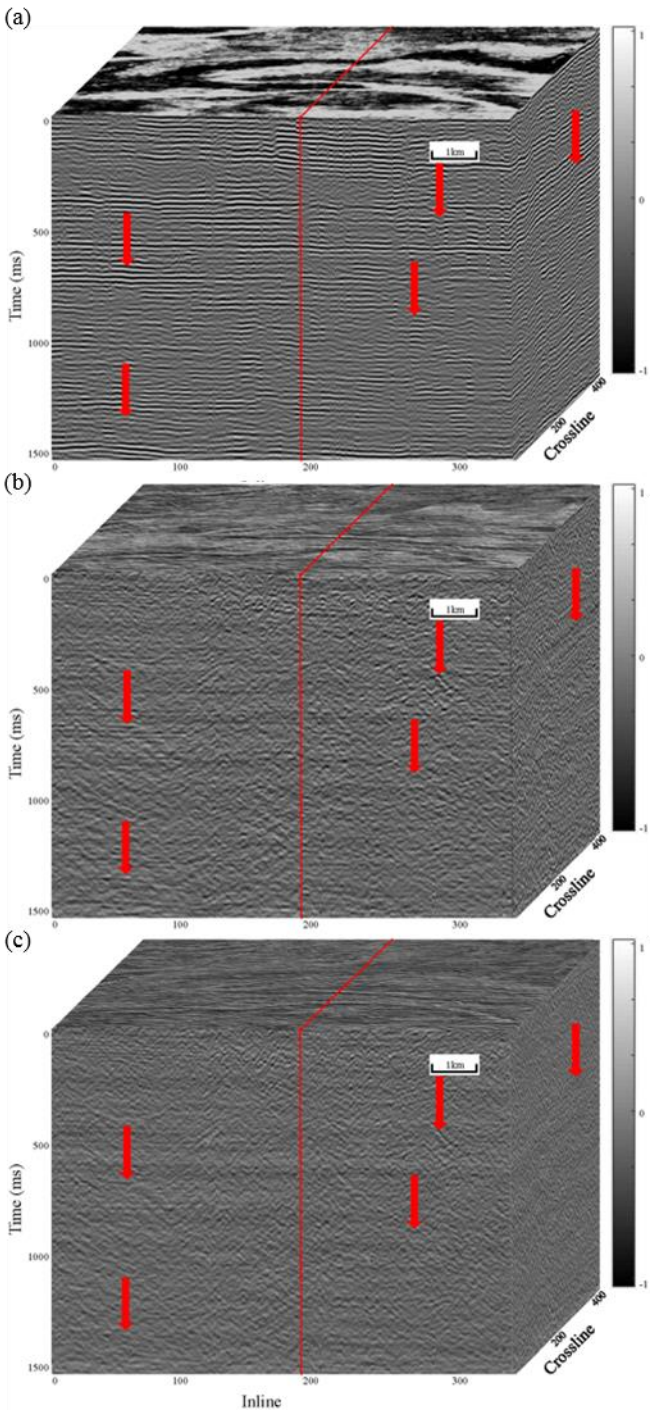


Figure 3.17. Rejected noises on 3D volume through different denoising methods on the seismic survey of Penobscot. (a) *f-x* deconvolution. (b) MDnCNN. (c) VMD-MDnCNN.

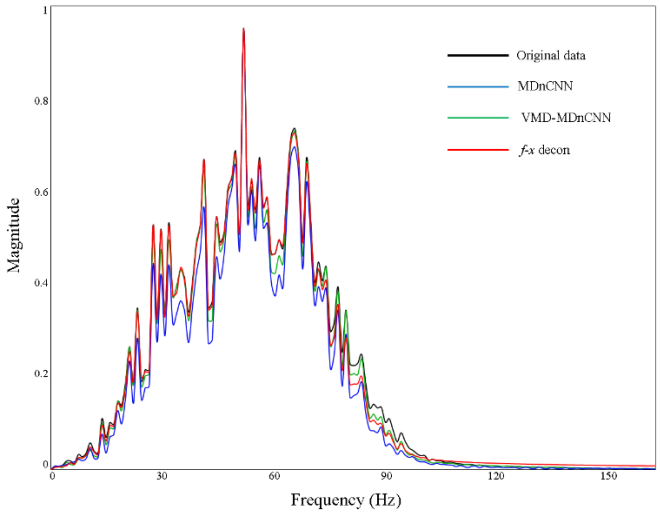


Figure 3.18. The frequency spectrum of the original data (black curve), MDnCNN result (blue curve), VMD-MDnCNN result (green curve) and *f-x* deconvolution result (red curve) for the seismic survey of Penobscot.

CONCLUSIONS

We propose a novel seismic noise attenuation method (VMD-MDnCNN) by integrating our MDnCNN with VMD. Current CNN based denoising methods either require the label of clean seismic data or the label of noise contained in the seismic data. Our method does not require clean seismic label nor purely noise label. The applications demonstrate that the white noise contained in the seismic can be simulated by enough user-generated white noise realizations. In addition, the applications demonstrate that the MDnCNN can obtain a more accurate estimation of the noise feature from the decomposed bandlimited seismic data. Both synthetic and real seismic data applications illustrate that our method is superior to the traditional denoising method of *f-x* deconvolution. The applications also demonstrate that our method not only effectively reject the white noise but also the migration artifacts contained in the seismic data.

REFERENCES

Abma, R., and J. Claerbout, 1995, Lateral prediction for noise attenuation by t-x and F-X techniques, Geophysics, 60, no. 6, pp. 1887-1896.

Bekara, M., and M. van der Baan, 2007, Local singular value decomposition for signal enhancement of seismic data, Geophysics, 72, no. 2, V59-V65.

Bekara, M., and M. van der Baan, 2009, Random and coherent noise attenuation by empirical mode decomposition, Geophysics, vol. 74, no. 5, V89-V98.

Burger, H., C. Schuler, and S. Harmeling, 2012, Image denoising with multi-layer perceptrons, part 1: comparison with existing algorithms and with bounds, Journal of Machine Learning Research.

Canales, L., 1984, Random noise reduction: 54th Annual International Meeting, SEG, Expanded Abstract, 525-527.

Chen, Y., J. Ma, 2014, Random noise attenuation by f-x empirical-mode decomposition predictive filtering, Geophysics, vol. 79, no. 3, V81-V91.

Chen, Y., W. Yu, and T. Pock, 2015, On learning optimized reaction diffusion processes for effective image restoration, Process of IEEE Conference, Computer Vision Pattern Recognition, 5261-5269.

Donoho, D. L., and I. M. Johnstone, 1994, Ideal spatial adaptation by wavelet shrinkage, Biometrika, vol. 81, 425-455.

Dragomiretskiy, K., and D. Zosso, 2014, Variational mode decomposition, IEEE Transactions on Signal Processing, vol. 62, 531-544.

Elad, M., and M. Aharon, 2006, Image denoising via sparse and redundant representations over learned dictionaries, IEEE Transaction on Image Processing, vol. 15, no. 12, 3736-3745.

Fomel, S., and Y. Liu, 2010, Seislet transform and seislet frame, Geophysics, vol. 75, V25-V38.

Hermann, F. J., U. Boniger, and D. J. Verschuur, 2007, Non-linear primary multiple separation with directional curvelet frames, Geophysical Journal International, vol. 170, 781-799.

Huang, N., Z. Shen, S., Long, M. Wu, H. Shih, Q. Zheng, and H. Liu, 1998, The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis, Proceedings of the Royal Society of London A: Mathmatical, Physical and Engineering Sciences, vol. 454, 903-955.

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

Jain, V., and S. Seung, 2009, Natural image denoising with convolutional networks, Advances in Neural Information Processing Systems, 769-776.

Kingma, D. P., and J. L. Ba, 2015, Adam: A method for stochastic optimization, International Conference on Learning Representations.

Kong, D., and Z. Peng, 2015, Seismic random noise attenuation using shearlet and total generalized variation, Journal of Geophysics and Engineering, vol. 12, no. 6, 1024-1035.

LeCun, Y., L. Bottou, Y. Bengio, and P. Haffner, 1998, Gradient-based learning applied to document recognition, Proceeding of IEEE, vol. 86, no. 11, 2278-2324.

Li, F., B. Zhang, S. Verma, and K. Marfurt, 2017, Seismic signal denoising using thresholded variational mode decomposition, Exploration Geophysics, vol. 49, no. 4, 450-461.

Liu, G., X. Chen, J. Du, and J. Song, 2011, Seismic noise attenuation using nonstationary polynomial fitting, Applied Geophysics, vol. 8, 18-26.

Liu, Y, N. Liu, C. Liu, 2012, Adaptive prediction filtering in t-x-y domain for random noise attenuation using regularized nonstationary autoregression, Geophysics, vol. 80, no. 1, V13-V21.

Oropeza, V., and M. Sacchi, 2011, Simultaneous seismic data denoising and reconstruction via multichannel singular spectrum analysis, Geophysics, vol. 76, no. 3, V25-V32.

Trad, D., T. Ulrych, and M. Sacchi, 2002, Accurate interpolation with high-resolution time variant Radon transforms, Geophysics, vol. 67, no. 2, 644-656.

Trickett, S., 2008, F-xy cadzow noise suppression, CSPG CSEG CWLS Convention, 303-306.

Vautard, R., P. Yiou, and M. Ghil, 1992, Singular-spectrum analysis: A toolkit for short, noisy chaotic signals, Physica D: Nonlinear Phenomena, vol. 58, no.1, 95-126.

Versteeg, R., 1994, The Marmousi experience: Velocity model determination on a synthetic complex data set, The Leading Edge, vol. 13, 927-936.

Wang, Y., 1999, Random noise attenuation using forward-backward linear prediction, Journal of Seismic Exploration, vol. 8, 133-142.

Wu, Z., and N. Huang, 2004, A study of the characteristics of white noise using the empirical mode decomposition method, Proceedings of The Royal Society, vol. 460, 1597-1611.

Wu, Z., and N. Huang, 2009, Ensemble empirical mode decomposition: a noise-assisted data analysis method, Advances in Adaptive Data Analysis, vol. 1, 1-41.

Xie, J., L. Xu, and E. Chen, 2012, Image denoising and inpainting with deep neural networks, Advances in Neural Information Processing Systems, 350-358.

Xue, Y., J. Yang, J. Ma, and Y. Chen, 2016, Amplitude-preserving nonlinear adaptive multiple attenuation using the high-order sparse radon transform, Journal Geophysics and Engineering, vol. 13, 207-219.

Yuan, S., S. Wang, C. Luo, and T. Wang, 2018, Inversion-based 3-D seismic denoising for exploring spatial edges and spatio-temporal signal redundancy. IEEE Geoscience and Remote Sensing Letters, vol. 15, no. 11, 1682–1686.

Zhang, K., W. Zuo, Y. Chen, D. Meng, and L. Zhang, 2017, Beyond a Gaussian denoiser: residual learning of deep CNN for image denoising, IEEE Transaction on Image Processing, vol. 26, no. 7, 3142-3155

CHAPTER 4
SEMI-AUTOMATED SEISMIC HORIZON INTERPRETATION USING ENCODER-DECODER CONVOLUTIONAL NEURAL NETWORK

Hao Wu[1], Bo Zhang[1], Tengfei Lin[2], Danping Cao[3] and Yihuai Lou[1]
[1]The University of Alabama, Department of Geological Science
[2] Department of Middle East E&P, Research Institute of Petroleum Exploration & Development, CNPC,
[3]China University of Petroleum (East China), School of Geoscience
This paper was submitted to SEG journal Geophysics in 2019

ABSTRACT

Seismic horizon is one of the critical inputs for the structure and stratigraphy modeling of reservoirs. It is extremely hard to automatically obtain an accurate horizon interpretation for the seismic data where the lateral continuity of reflections are interrupted by the faults and unconformities. The process of seismic horizon interpretation segment the seismic traces into different parts and we consider the horizon interpretation as an image segmentation problem by treating each seismic trace as a 1D image. We propose a novel semi-automated seismic horizon interpretation method using an encoder-decoder convolutional neural network (CNN). To efficiently learn the waveform pattern that bounded by two adjacent horizons, we use a variable size for the convolution filters which differs from the current CNN based image segmentation. We form the training data by extracting the seismic traces on a user-defined coarse grid. The interpreted horizons for the seismic traces on the coarse grid separate the seismic traces into several segments. Finally, we obtain the seismic horizons over the whole seismic survey by applying the trained network to the rest of seismic traces. Two field data examples demonstrate that our proposed method can accurately generate seismic horizons for each seismic trace if we only have interpreted horizons on a coarse grid of the seismic survey.

INTRODUCTION

Seismic horizon interpretation is an essential step for reservoir characterization. The seismic horizons can be treated as the stratigraphic boundaries that represent the depositional environments and geological features. Manual interpretation is the most familiar but time-consuming interpretation technique. In recent decades, seismic horizon interpretation has been automated to some extent by others. Zeng et al. (1998) proposed an interpolation method, which first manually picking several reference strata slices and then build a surface volume by interpolating the interpreted seismic surfaces. However, the interpolated surfaces may not follow the local discontinues, like the fault and unconformities. Stark (2003) computed the unwrapping instantaneous phase to generate a relative geological time (RGT) volume and produce multiple horizons simultaneously. The seismic horizons that generated by RGT volume can provide a quality measure along the unconformities, but it cannot follow the fault gap. Wu and Zhong (2012) improved this method by using the graph-cut phase unwrapping, which performs well at strong discontinues structure zone. Lomask et al. (2006) first calculated the local dips over the entire seismic volume and transferred them into the time shift and then apply the least-square method to track the seismic horizons automatically. Parks (2010) proposed a slope-based flattening method, which use the structure tensor to estimate the dip and track seismic horizons automatically. Fomel (2010) use the algorithm of predictive painting to build a 3D seismic surface volume. Wu and Hale (2015) proposed an automatic horizons tracking method by adding control points in the complex structure zone. Wu and Fomel (2018) suggested a novel automatically seismic horizon interpretation method with local slope and multi-grid correlation that can effectively deal with faults without any human effort or very few control points.

In recent years, deep learning has captured significant attention in geoscience, and some deep learning based methods have been successfully applied to the geological structure detection (Huang et al., 2017; Wu et al., 2018; Zhao et al., 2018; Di et al., 2018;). Deep learning is a subset of machine learning, and CNN (LeCun et al., 1998) is one of the most popular and widely used deep learning algorithms, which achieves great success in the field of computer vision and pattern recognition. The basic theory of CNN is learning the features of the images and classify the images by different labels. Many algorithms based on CNN have been proposed to address the problem of images classification (Krizhevsky et al., 2012; Szegedy et al., 2015; Simonyan and Zisserman, 2015), and the accuracy has beyond the human level. However, those methods can only recognize one dominant object in one single image. Girshick et al. (2014) proposed the algorithm of regions with CNN features (R-CNN), which can localize and detect multiple objects in one single image. This method uses a series of different scale rectangle regions to segment multiple objects, but it cannot provide accurate boundaries of the objects. A breakthrough came from the fully connected neural network (FCN) (Long et al., 2015) that enhance the capability of image segmentation from the regional level to the pixel level. He et al. (2015) developed the ResNet with skip-connection, which make the training of very deep convolutional neural network possible. Other pixel-level image semantic segmentation methods include SegNet (Badrinarayanan et al., 2015), U-Net (Ronneberger et al., 2015), E-Net (Paszke et al., 2016), and Mask R-CNN (He et al., 2017).

In this paper, we treat each seismic trace as a 1D image and proposed a novel semi-automated seismic horizon interpretation method using encoder-decoder CNN. The framework of our network is similar to SegNet (Badrinarayanan et al., 2015). However, the size of the convolution kernel varies with layers where the common CNN based segment algorithms employ

a fixed size of convolution kernel for all the layers. This paper is organized as follow: we first discuss the traditional seismic horizon interpretation techniques. We next introduce the details of our proposed method for seismic horizon interpretation. We finally present the testing results on two different field data and compare the results with the traditional seeded auto-tracking method.

CURRENT SEISMIC HORIZON INTERPRETATION TECHNIQUES

Current automatic and semi-automatic horizon interpretation algorithms are based on comparing the similarity between nearby seismic traces or the dip of seismic reflectors. Seed-based auto-tracking is the most popular semi-automated seismic horizon interpretation method that has been widely used in the industry and commercial software (Dorn, 1998). The first category is based on comparing the similarity $s(i,j)$ of waveforms between the seed trace $x(t)$ and a neighbor trace$y(t)$:

$$s(i,j)=\frac{\sum_{\tau=-w}^{w}\left(x(t_i-\tau)-\frac{1}{2w}\sum_{\tau=-w}^{w}x(t_i-\tau)\right)\left(y(t_i-\tau)-\frac{1}{2w}\sum_{\tau=-w}^{w}y(t_i-\tau)\right)}{\sqrt{\sum_{\tau=-w}^{w}\left(x(t_i-\tau)-\frac{1}{2w}\sum_{\tau=-w}^{w}x(t_i-\tau)\right)^2\sum_{\tau=-w}^{w}\left(y(t_i-\tau)-\frac{1}{2w}\sum_{\tau=-w}^{w}y(t_i-\tau)\right)^2}}, \quad (1)$$

where $x(t_i-\tau)$ denotes the amplitude of seismic at time index *i* with lag τ, $y(t_j-\tau)$ denotes the amplitude of synthetic at time index *j* with lag τ, $2w$ is the temporal length of the correlation window, *i* and *j* are time sample indices. The selected neighbor trace become "seed" trace if the similarity between the original seed trace and neighbor trace is greater than a pre-defined threshold. We repeat the process of similarity comparing and changing neighbor trace to "seed" trace until the seismic traces become "seed" traces.

Figure 4.1 shows two tests using seeded auto-tracking over the zone of faults (upper) and unconformities (lower). The red curves in Figure 4.1 indicate the ground truth of our target horizons. The white dots denote the seeds, and the green curves represent the corresponding interpreted horizons using seeded auto-tracking. Note that seeded auto-tracking achieved a good result at the smooth and flat areas. However, this method failed to detect the edge of unconformities and follow the path of fault indicated by blue arrows.

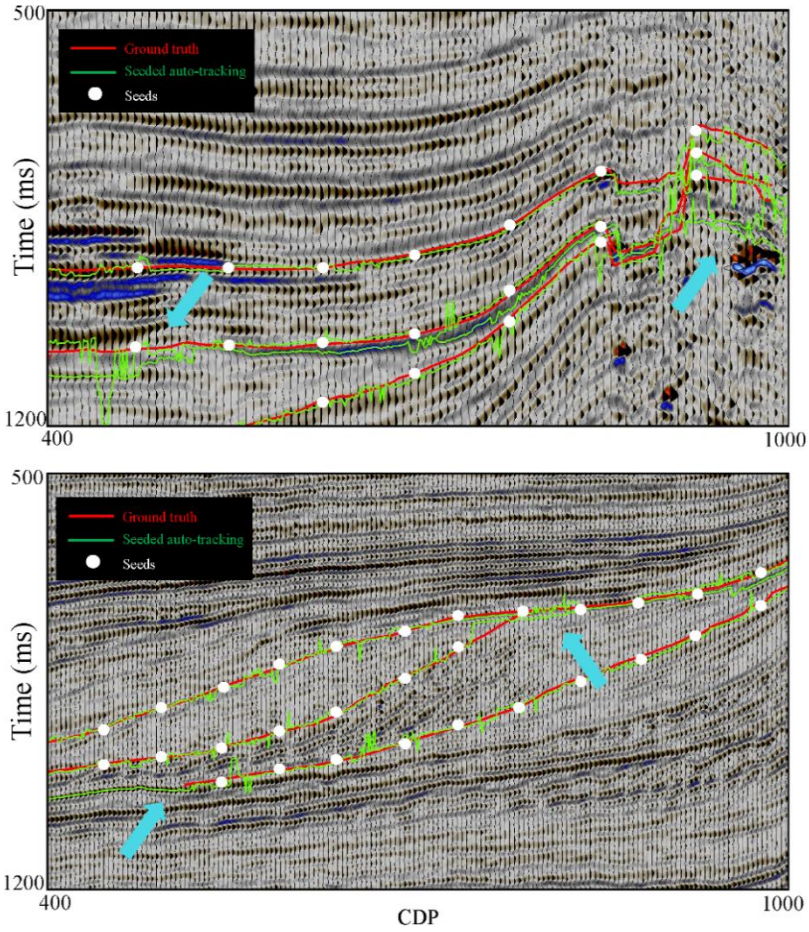


Figure 4.1. Illustrating the limitation of seeded auto-tracking for seismic horizon interpretation in complicated zones. The white dots denote the preset seeds. The red curves and green curves represent manually interpreted horizons and predicted horizons using seeded auto-tracking, respectively.

Wu and Hale (2015) treated the seed-based horizon interpretation as constrained optimized problem:

$$\text{Minimize } S(\boldsymbol{\beta}) = \sum_{i=1}^{m\times n} \left|\Delta \mathbf{t}_i^{(\beta)} - \Delta \mathbf{t}_i^{(Seis)}\right|^2 \tag{2a}$$

$$\text{Subject to } \beta_s = t_0(seed), \tag{2b}$$

where $t_0(seed)$ is the TWT of the seed; $\boldsymbol{\beta} = (\beta_1, \ \beta_2, \ \ldots, \ \beta_{m*n})$ is the two-way travel time (TWT) of the samples on the horizon in a vector format, *m* and *n* are the inline and crossline number, respectively; $\Delta \mathbf{t}_i^{(\beta)}$ and $\Delta \mathbf{t}_i^{(Seis)}$ are the dips computed from the tracked horizon patch and the dips computed from seismic reflection events, respectively

$$\Delta \mathbf{t}_i^{(\beta)} = \begin{bmatrix} \beta_{i+m} - \beta_i \\ \beta_{i+1} - \beta_i \end{bmatrix} \tag{3}$$

$$\Delta \mathbf{t}_i^{(Seis)} = \begin{bmatrix} p \\ q \end{bmatrix} . \tag{4}$$

The objective is to minimize the difference between the dips computed from tracked horizon patches and the dips computed from seismic events.

In this paper, we illustrate the superiority of our algorithm by comparing our results with the results produced from a commercial software where the algorithm is based on Equation 1. Comparing the results with the results produced using Wu and Hale's (2015) method is out of the scope of this paper considering that we do not have access to their code.

Training set preparation

To generate the training data and training label for our network, we first manually interpret the seismic horizons on a coarse grid. The seismic traces on the coarse grid are treated as training data and the rest of seismic traces are treated as testing data. Next, the interpreted seismic traces have been segmented into different parts according to the time index of seismic horizons. We then produce the training label by assigning different symbols to different segmented parts that bounded by two adjacent seismic horizons. For example, if we have four manually interpreted seismic horizons in total and these four horizons segment the seismic traces into five different segments, so we label these five different segments from top to bottom with 0, 1, 2, 3 and 4, respectively. Researchers can use any unique symbols to label the segmented parts. Figure 4.2 shows an example of training data and the corresponding training label.

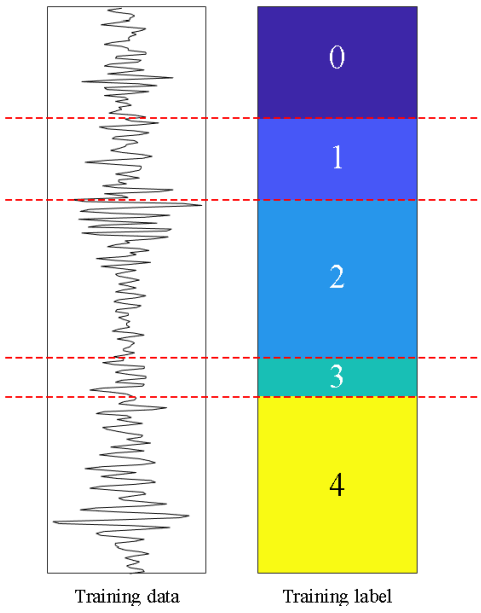


Figure 4.2. One example of a training trace and the corresponding label, the red dash lines represent different seismic surfaces.

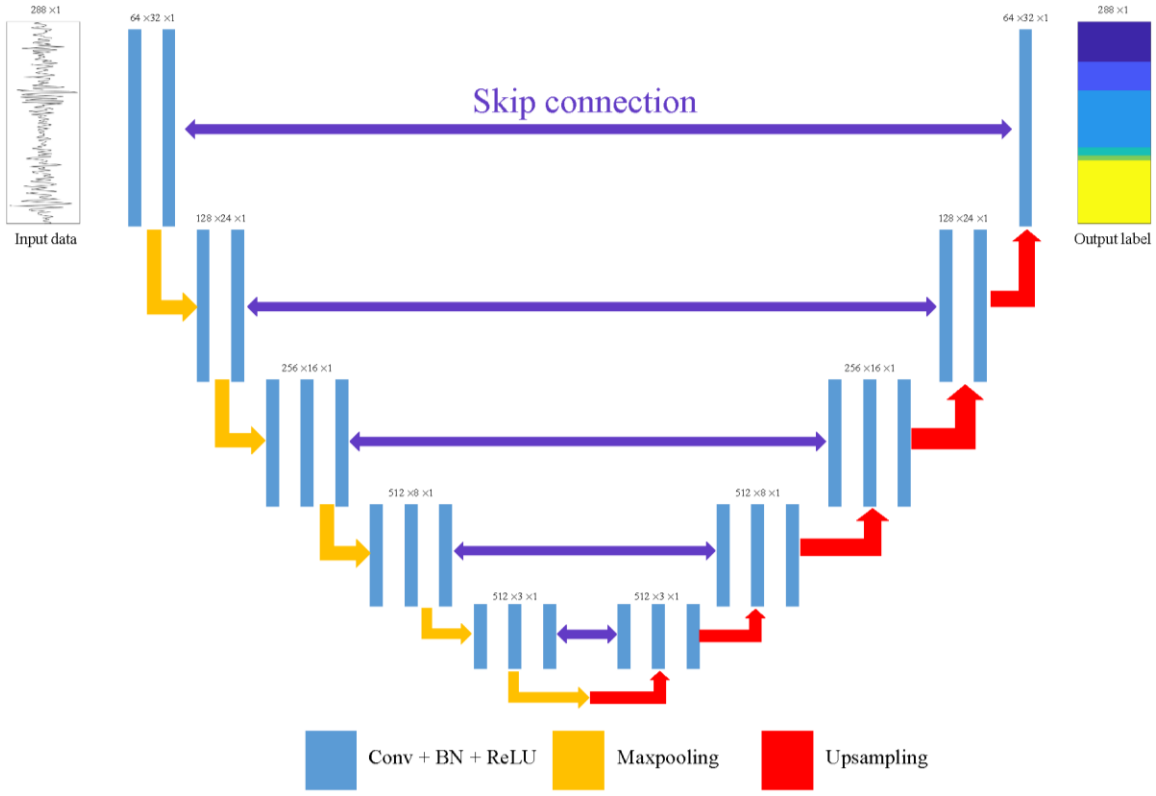


Figure 4.3. The framework of encoder-decoder CNN that used for seismic horizon interpretation.

CNN framework

Figure 4.3 shows the framework of the proposed network. This symmetrical framework contains two main parts: the encoder network and decoder network. A set of 1D seismic traces x_i are fed into the encoder as the input training data. The encoder network is composed of a series of repeated layers, and each layer contains convolution filter, batch-normalization regularizer (Ioffe and Szegedy, 2015), ReLU activation operator and a max-pooling operator. The objective of the convolution filter is to generate feature maps of the input seismic data. We choose to use downward trend size of convolution kernel in the encoder part, which are 32×1, 24×1, 16×1, 8×1 and 3×1, respectively. The reason for these downward kernel size is learning the feature of input traces in different spatial scale. The objective of the ReLU is to activate the main features

contained in the feature map. The maxpooling operator is used for downsampling and the size is 2×1 for all the convolution layers. The role of decoder network is reconstructing the image according to the feature maps and feed into the element-wise classification layer. The decoder also contains several layers that corresponding to the encoder network, and each layer is composed of upsampling operator, convolution filter, batch-normalization regularizer and ReLU activation operator. The upsampling operator is used for reconstruct the feature maps and the size is 2×1 for all the convolution layers. In the decoder, we choose to use the upward trend size of convolution kernel that the size are 3×1, 8×1, 16×1, 24×1 and 32×1, respectively. The last step of the network is an element-wise softmax classification layer with the size of 1×1 and the role of this layer is defining the weight of each pixel with the corresponding training label y_i .

Loss function and parameters

Next, the cross-entropy loss function is used in our network to calculate the loss between the ground truth and prediction, which is widely used in the image segmentation. The loss function $J(\theta)$ is given as follows:

$$h_\theta(x_i)=\frac{1}{1+e^{-\theta^T x_i}}, \tag{5}$$

$$J(\theta)=-\frac{1}{m}\sum_{i}^{m} y_i \log\left(h_\theta(x_i)\right)-\frac{1}{m}\sum_{i}^{m}(1-y_i)\log\left(1-h_\theta(x_i)\right), \tag{6}$$

where $h_\theta(x_i)$ is the prediction probabilities with all the trainable parameters θ of the input data x_i for the last classification layer, and y_i is the corresponding training label. Then, the built network transforms the procedure of seismic traces segmentation into an optimization problem by solving a sequence of nonlinear functions. A gradient-based optimization algorithm of

adaptive moment estimation (Adam) (Kingma and Ba, 2015) is employed to minimize the proposed objective function through iterative updating the parameters of the network.

In the training process, the size of input training data and training label is N×1, where N is the number of time samples of each seismic trace. To overcome the overfitting problem in the training procedure, the training and validation seismic traces are randomly selected during each optimization epoch. The percentage of the training and validation seismic traces in this paper is 70% and 30%, respectively. A certain seismic trace may belong to the training set in current optimization epoch but may belong to the validation set in the next optimization epoch.

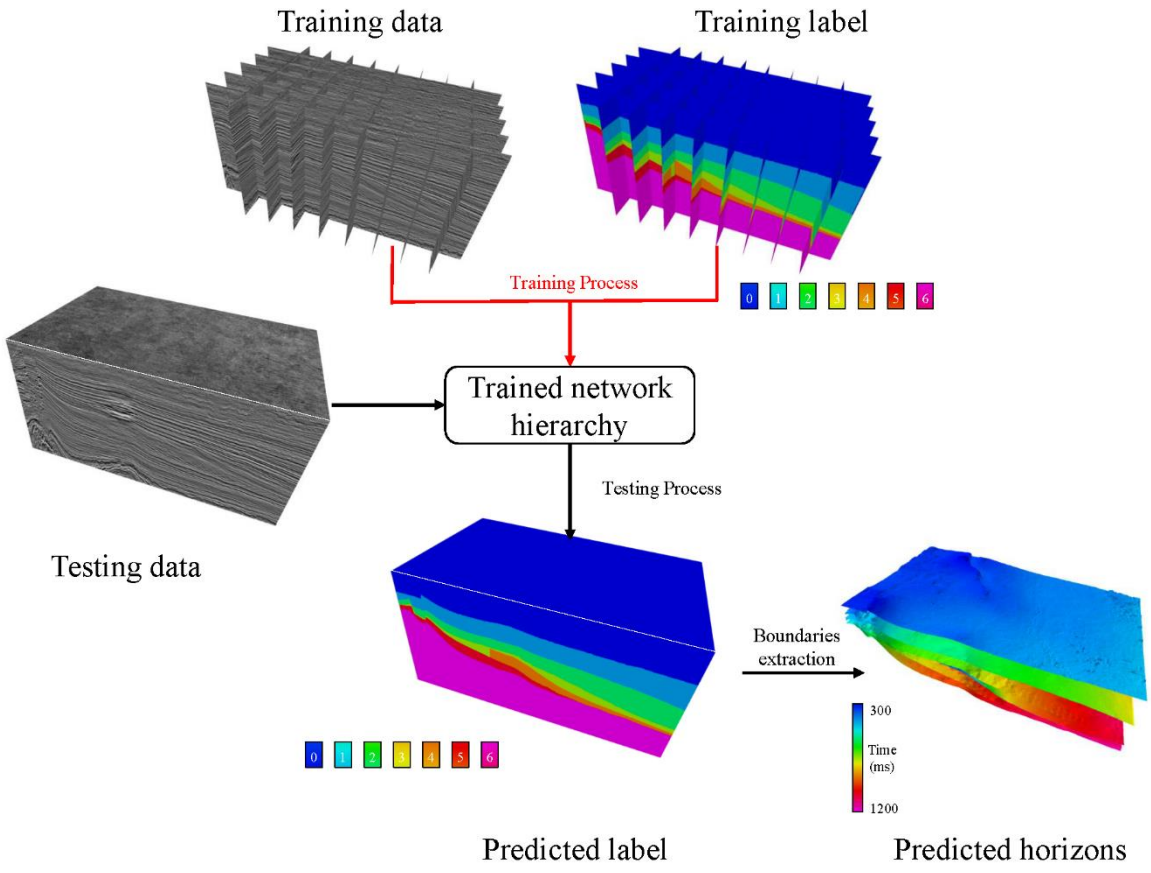


Figure 4.4. The workflow of our proposed method for semi-automated seismic horizon interpretation.

Once the training process is finished, the well-trained network hierarchy is applied to the test data. The trained network performs trace segmentation and label assignment trace by trace. Figure 4.4 illustrates the workflow of our proposed method for seismic horizon interpretation, which shows the examples of the training data, corresponding training label, testing data, and the corresponding predicted label using our proposed method. We finally extracted the boundaries of the predicted label as our final interpreted horizons.

APPLICATION

Data description

To demonstrate the effectiveness and superiority of our proposed method over the traditional method on seismic horizon interpretation. We test on two different field data examples and compare with the traditional method of seeded auto-tracking.

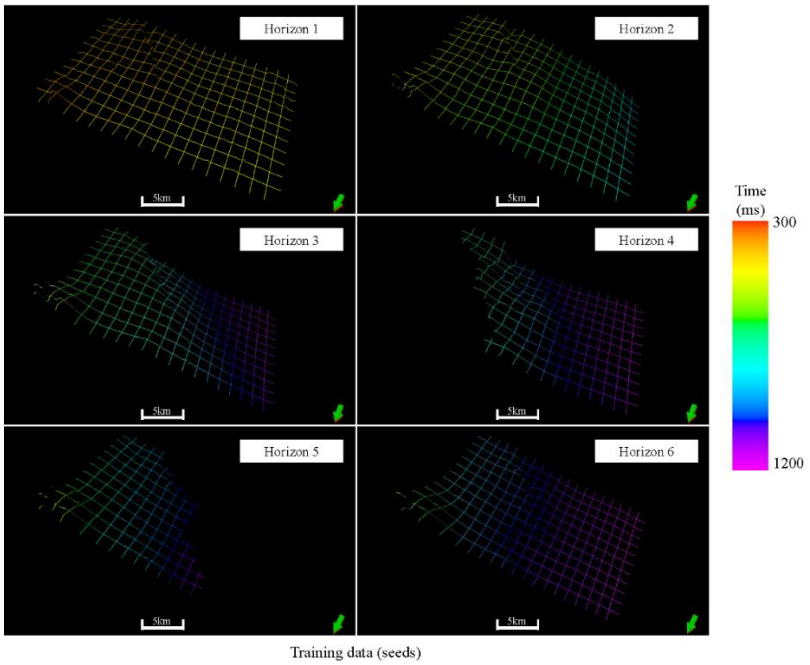


Figure 4.5. The manually interpreted seismic horizons on a 50×50 coarse grid for the first field data.

The first field data, F3-block, was a public data that acquired over the North Sea, Netherlands. The applied data has 601 inlines, 901 crosslines and 288 time samples in this seismic survey, the sample rate is 4ms. We first manually interpret six seismic horizons on a 50×50 coarse grid (Figure 4.5). We then use those manually interpreted seismic traces to build the training data and corresponding training labels for the following training process. We also manually interpret the seismic horizons on the rest of the seismic traces to produce the ground truth result (Figure 4.6), which is used to validate our predicted outcome.

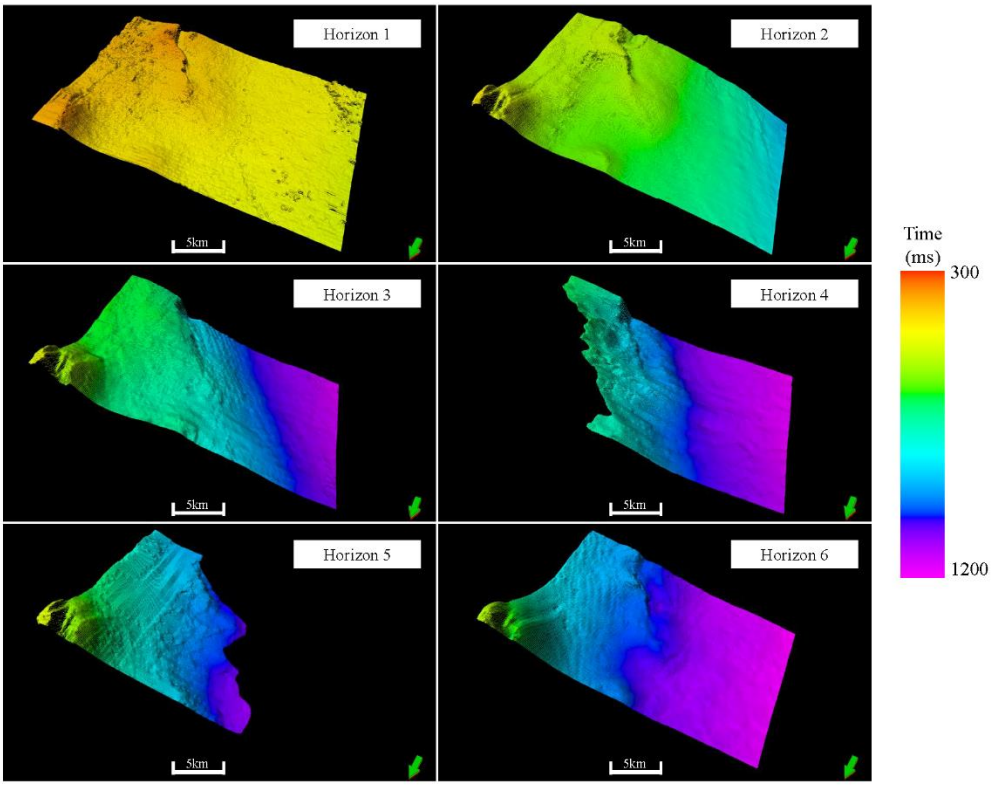


Figure 4.6. The manually interpreted seismic horizons over the whole seismic survey for the first field data. We treat that as the ground truth horizons to validate our predicted results.

The second field data, Penobscot, was a public data that acquired over Nova Scotia, Oversea Canada. The applied data has 500 inlines, 500 crosslines and 512 time samples in this

seismic survey, the sample rate is 4ms. Since the geology structure of this seismic survey is more smooth and flat than the first field data. We choose to use more sparse distributed seismic traces that we manually interpret four seismic horizons on a 100×100 coarse grid (Figure 4.7). We then use those manually interpreted seismic traces to build the training data and corresponding training labels for the following training process. We also manually interpret the seismic horizons on the rest of the seismic traces to produce the ground truth result (Figure 4.8), which is used to validate our predicted result.

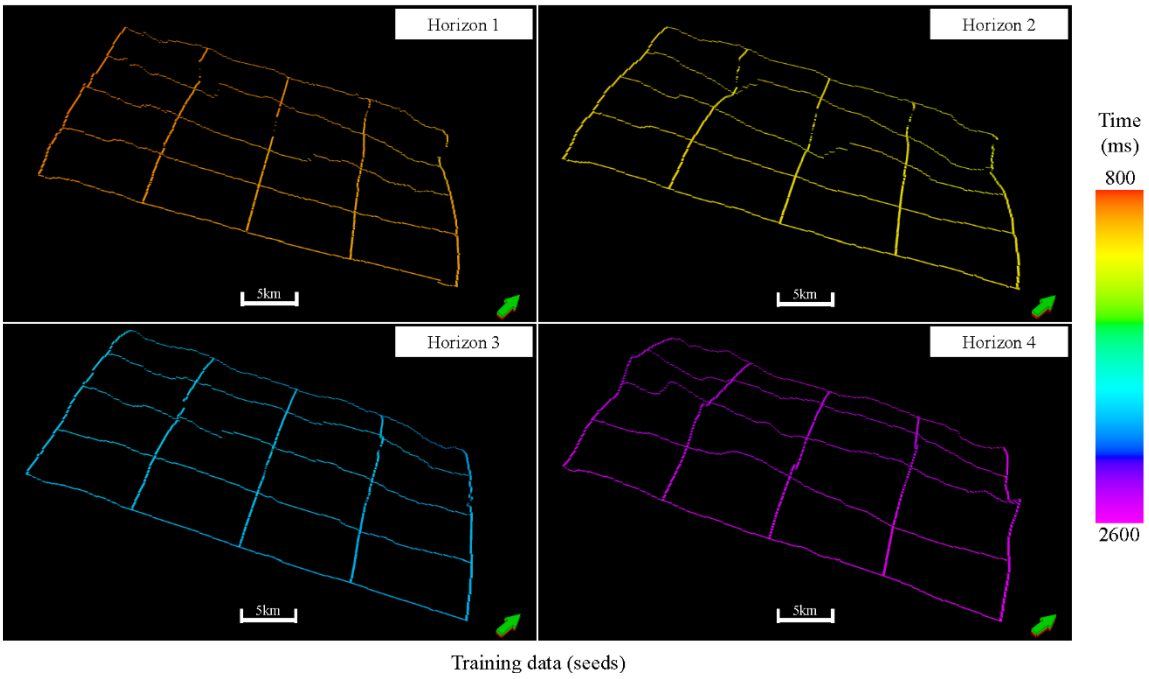


Figure 4.7. The manually interpreted seismic horizons on a 100×100 coarse grid for the second field data.

Training process

In the training process, we set the maximum epochs equal to 20 for both of these two datasets. Figure 4.9 shows the loss and accuracy of the training process for the first field data. Note that both the training and validation accuracies gradually increase to 95% while the training

and validation loss converges to 0.1 after ten epochs for the first field data. Figure 4.10 shows the loss and accuracy of the training process for the second field data. Note that both the training and validation accuracies gradually increase to 95% while the training and validation loss converged after ten epochs for the second field data.

Both of these two datasets are training on an NVIDIA Quadro P6000 GPU of 24 GB memory. The computation cost is about 45 minutes for the first field data and about 20 minutes for the second field data. The time cost for the manual interpretation on the coarse grid is about 60 minutes for the first field data and about 30 minutes for the second field data. The time cost for manually interpreting the whole seismic survey is about three days for the first field data and about two days for the second field data.

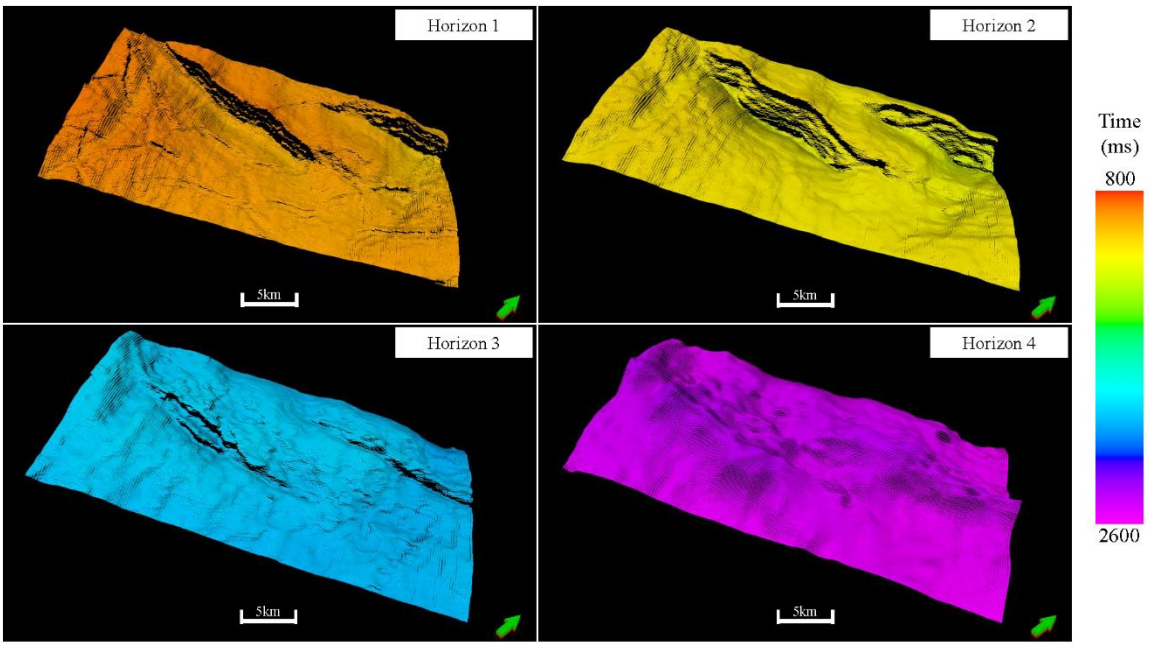


Figure 4.8. The manually interpreted seismic horizons over the whole seismic survey for the second field data. We treat that as the ground truth to validate our predicted results.

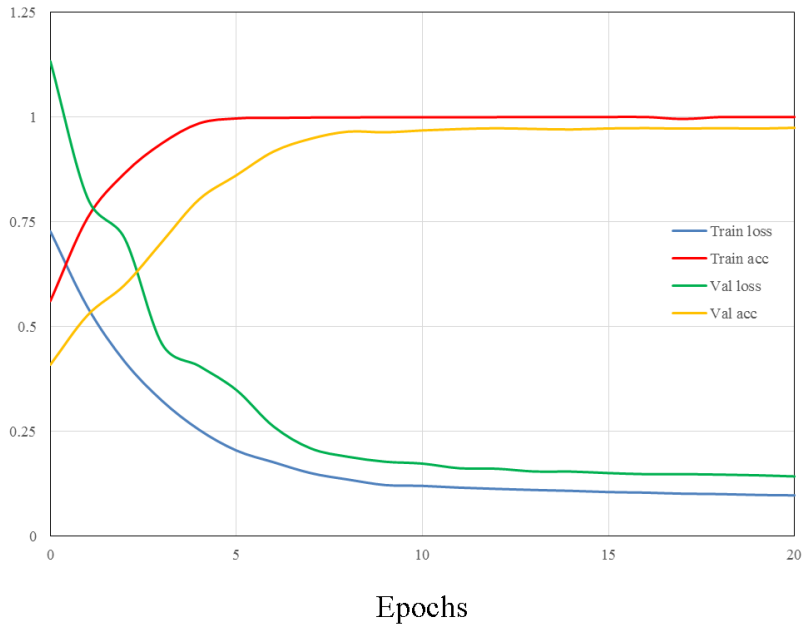


Figure 4.9. The loss and accuracy are varying with epochs of the first field data.

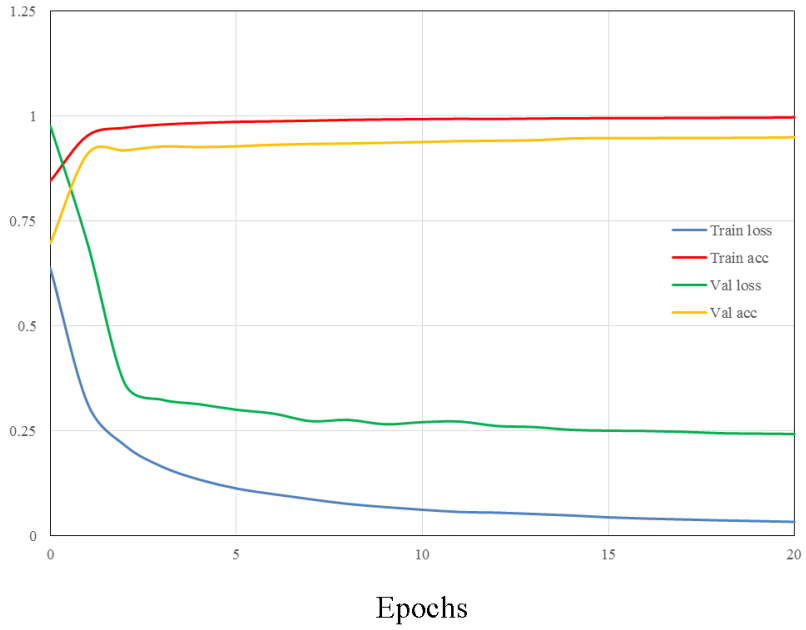


Figure 4.10. The loss and accuracy are varying with epochs of the second field data.

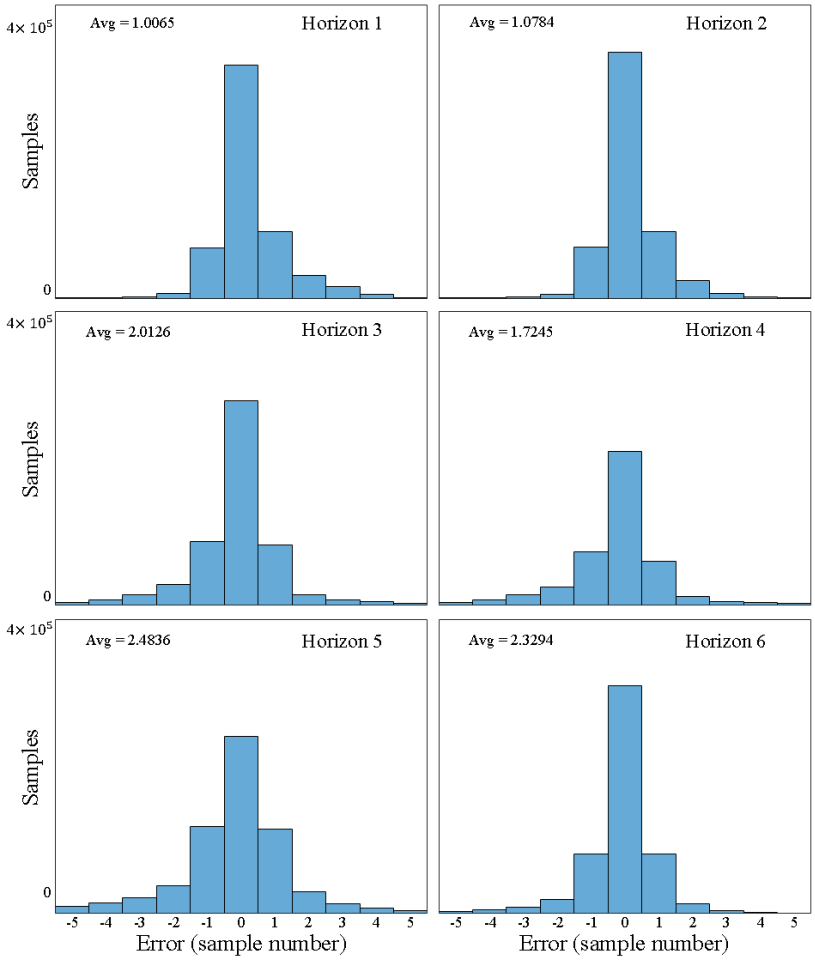


Figure 4.11. The error distribution and the average of absolute error for each horizon of the first field data.

Refining predicted horizons by spatial correlation constraint

We evaluate the prediction error using the average absolute difference between the ground truth and predicted result. The average absolute difference is defined as follow:

$$d = \frac{\sum_{i=1}^{N} |G_i - P_i|}{N} \tag{7}$$

where d denotes the absolute average difference, G_i denotes the ground truth of the horizon for the i^{th} trace, P_i denotes the predicted horizon for the i^{th} trace by using our proposed method, N

is the total trace number of the testing data. Figure 4.11 shows the error distribution and average of absolute error for each horizon. The *x*-axis is the error in sample and *y*-axis is the number of sample. Note that the average value of the six horizons are all smaller than 3 samples which demonstrate a very good prediction. Figure 4.11 shows that the error of 95% of prediction result is smaller than 1 sample. However we still have a small amount of seismic traces whose error is greater than 5 samples (20 ms). Figure 4.12 shows one representative inline seismic slice located at the boundary of our seismic survey. Note that the improper two-way travel time "jump" of predicted horizons indicated by the red rectangles in Figure 4.12. The yellow rectangle in Figure 4.12 indicates the reasonable two-way travel time "jump" caused the existing of fault. In this paper, we propose two-step filtering to differentiate the reasonable and improper two-way travel time jumps. We preserve the two-way travel jumps caused by faults and then smooth the two-way travel time for those improper jumps. The two-step filtering used in this paper is present in Algorithm 1:

Algorithm 1: The two-step filter to detect the spikes and faults

---

Input: predicted horizon $h(t)$ D with size (1, T)

Output: location of spike s

1: Initialize: $s \leftarrow [\,]$, $M \leftarrow 10$,

2: **for** i = 1 to T

3: **for** $j = i$ to i+10

4: **if** $\|h(j) - h(i)\| < M$

5: **continue**

6: **elseif** $\|h(j) - h(i)\| \geq M$

7: **for** $k = j$ to j+10

8: **if** $\|h(k) - h(i)\| \geq M$

9: **continue**

10: **elseif** $\|h(k) - h(i)\| < M$

11: $s \leftarrow [i{:}\, k; s]$

12: **end**

We then use the following equation to interpolate the two way travel time of the improper jumps. We designed a Gaussian window:

$$\varphi(\tau) = e^{-\frac{1}{2}\left(\frac{\tau-\omega}{\sigma\omega}\right)^2}, \sigma \leq 0.5.$$

We then smooth the spikes:

$$f(s) = \sum_{\tau=-\omega}^{\omega} h(s+\tau)\varphi(\tau),$$

where *s* denotes the indexes of the spikes. Figure 4.13 shows the final output horizons overlaid on the inline section slice shown in Figure 4.12. Note that we simultaneously mitigate the unreasonable jumps of two-way travel time for the same horizon and keep the reasonable jumps of two-way travel time across the faults.

Experiment result and comparison

We apply these two well-trained networks to the testing data of these two field data to predict the seismic horizons over the whole seismic surveys. The computation cost of testing for both two field data is under 1 minute. We also compare the results using our proposed method with the results using seeded auto-tracking. The time cost of seeded auto-tracking is about 16 minutes for the first field data and about 11 minutes for the second field data.

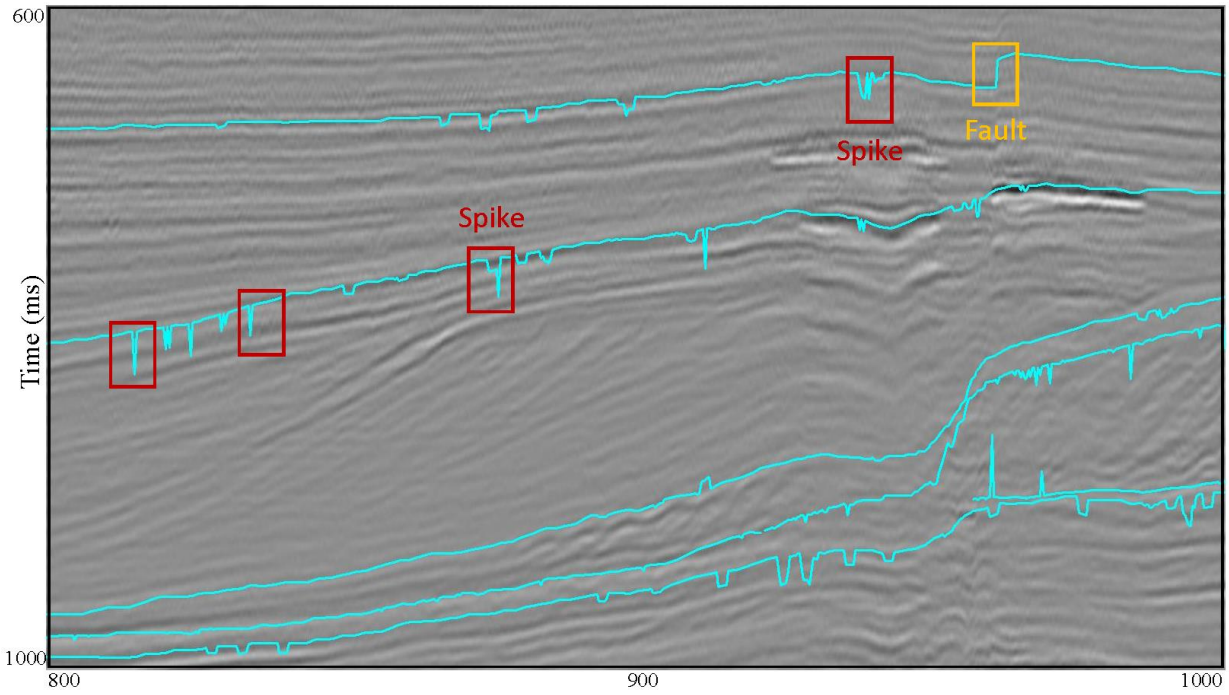


Figure 4.12. One representative inline seismic slice located at the boundary of our seismic survey.

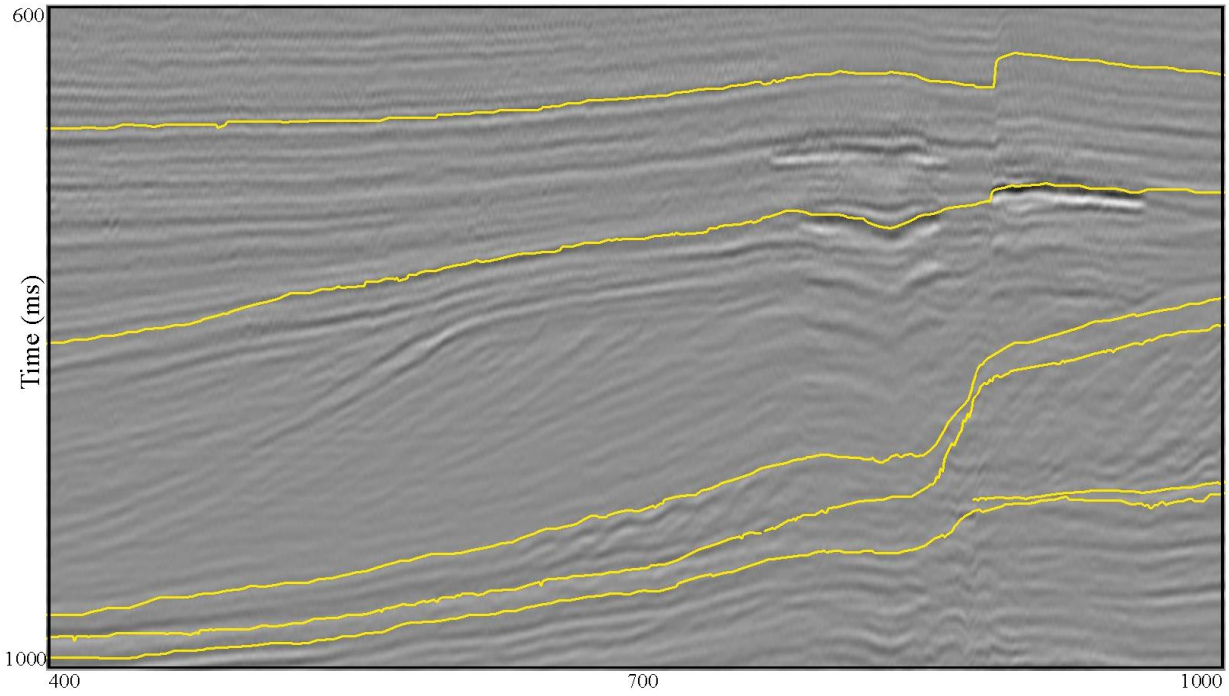
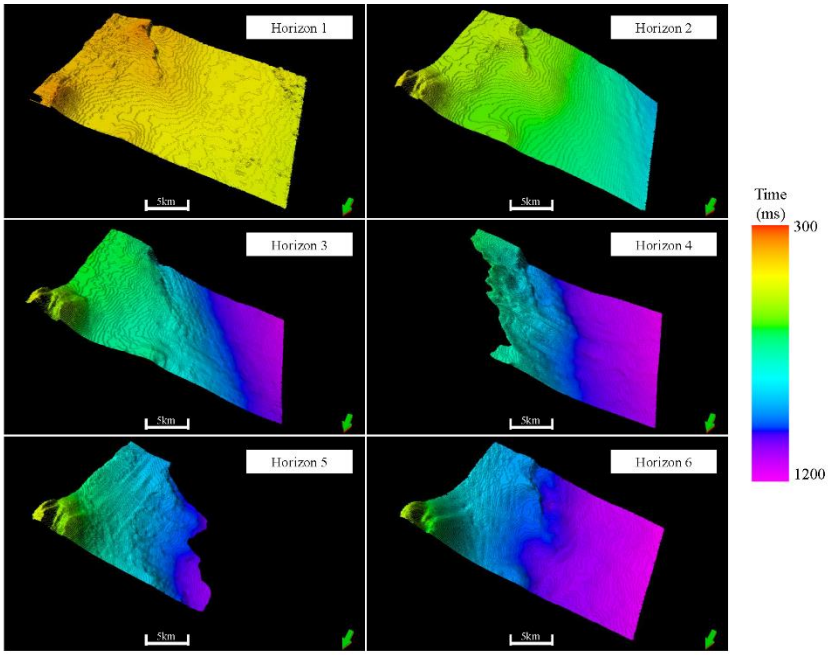
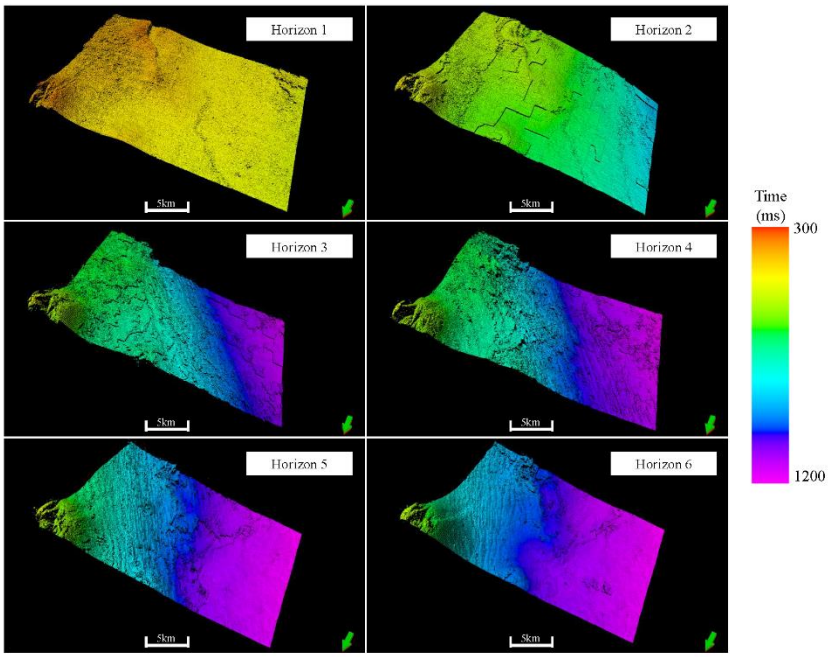


Figure 4.13. The final output horizons overlaid on the inline section slice shown in Figure 4.12.



CNN

Figure 4.14. The interpreted seismic horizons using our proposed method of the first field data.



Seeded auto-tracking

Figure 4.15. The interpreted seismic horizons using seeded auto-tracking of the first field data.

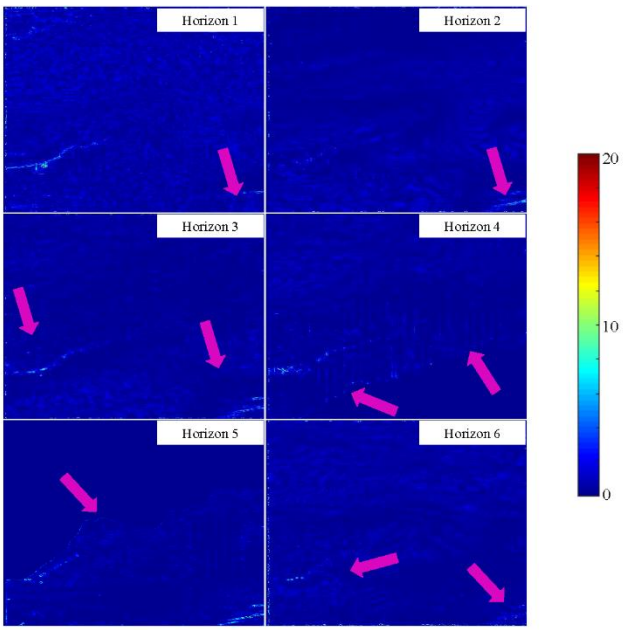


Figure 4.16. The absolute error maps between the ground truth and the results using our proposed method for the first field data.

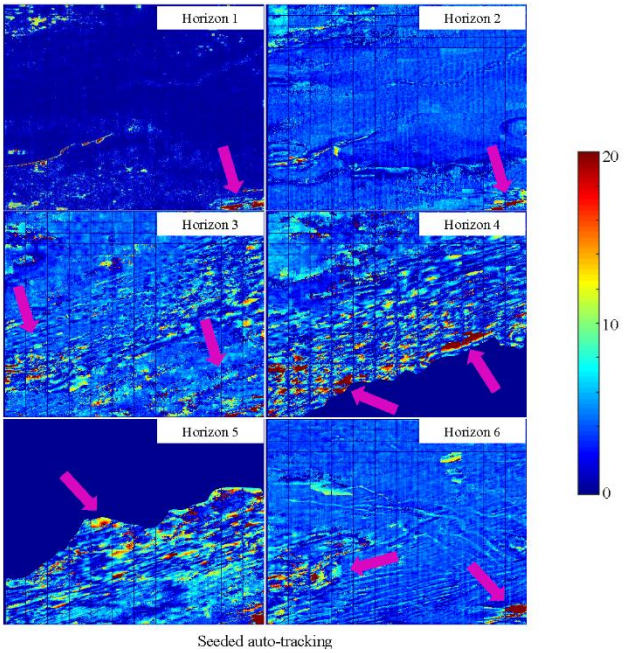
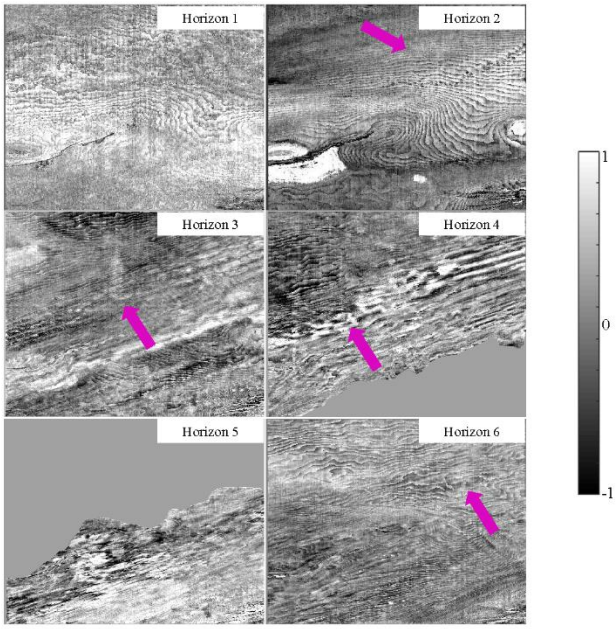
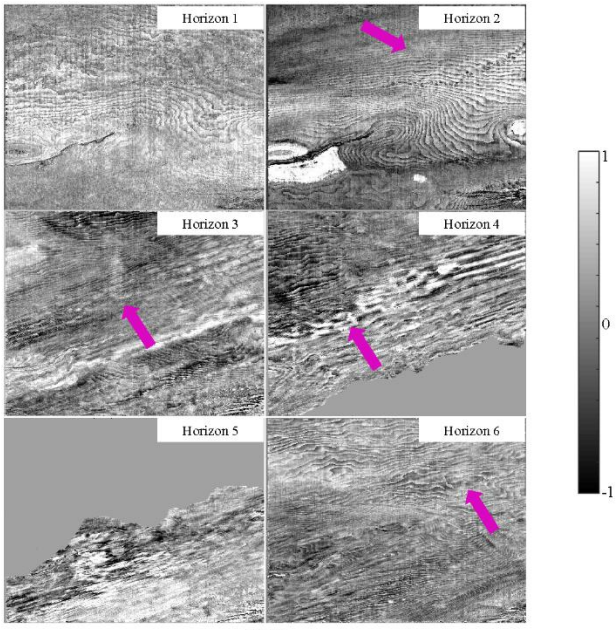


Figure 4.17. The absolute error maps between the ground truth and the results using seeded auto-tracking for the first field data.



Ground truth

Figure 4.18. The seismic amplitude maps on the horizon surfaces of ground truth of the first field data.



CNN

Figure 4.19. The seismic amplitude maps on the predicted horizon surfaces using our proposed method for the first field data.

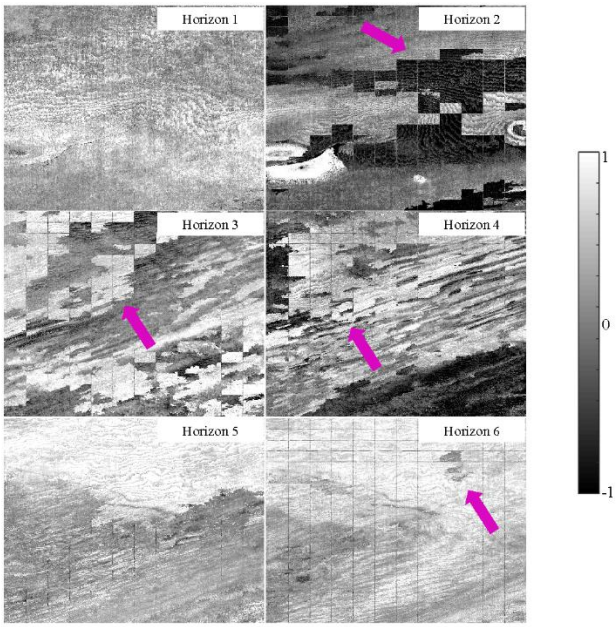


Figure 4.20. The seismic amplitude maps on the predicted horizon surfaces using seeded auto-tracking for the first field data.

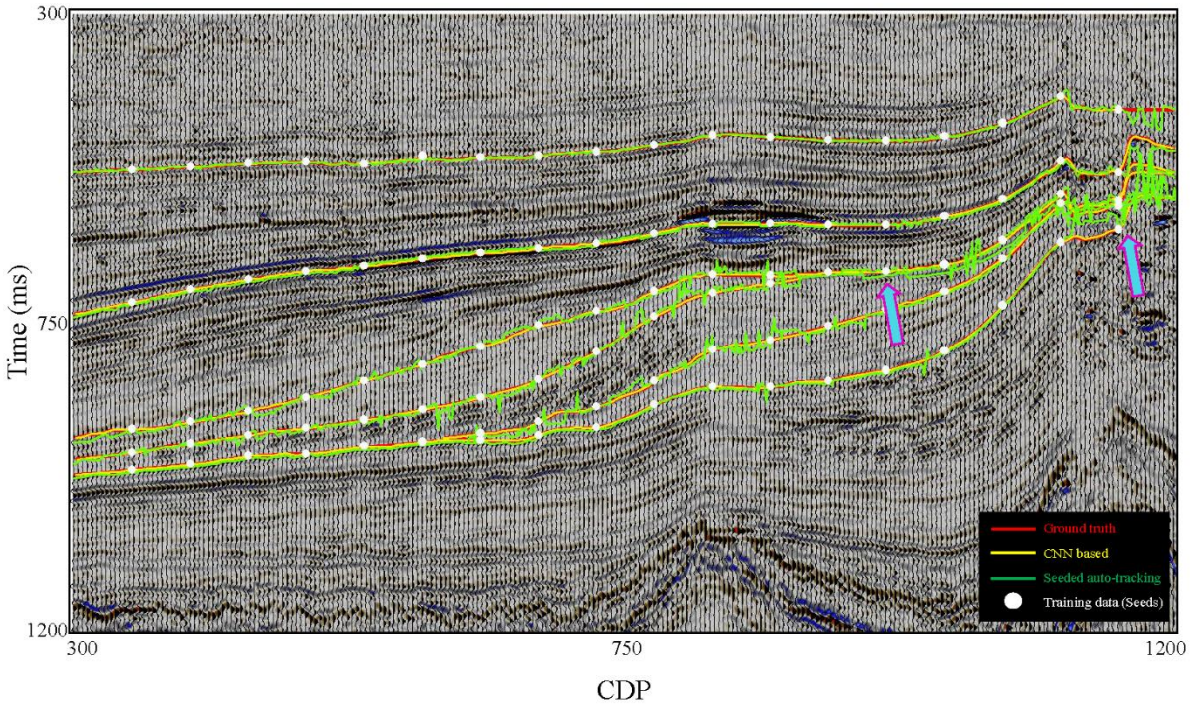


Figure 4.21. The comparison on an inline section of the first field data.

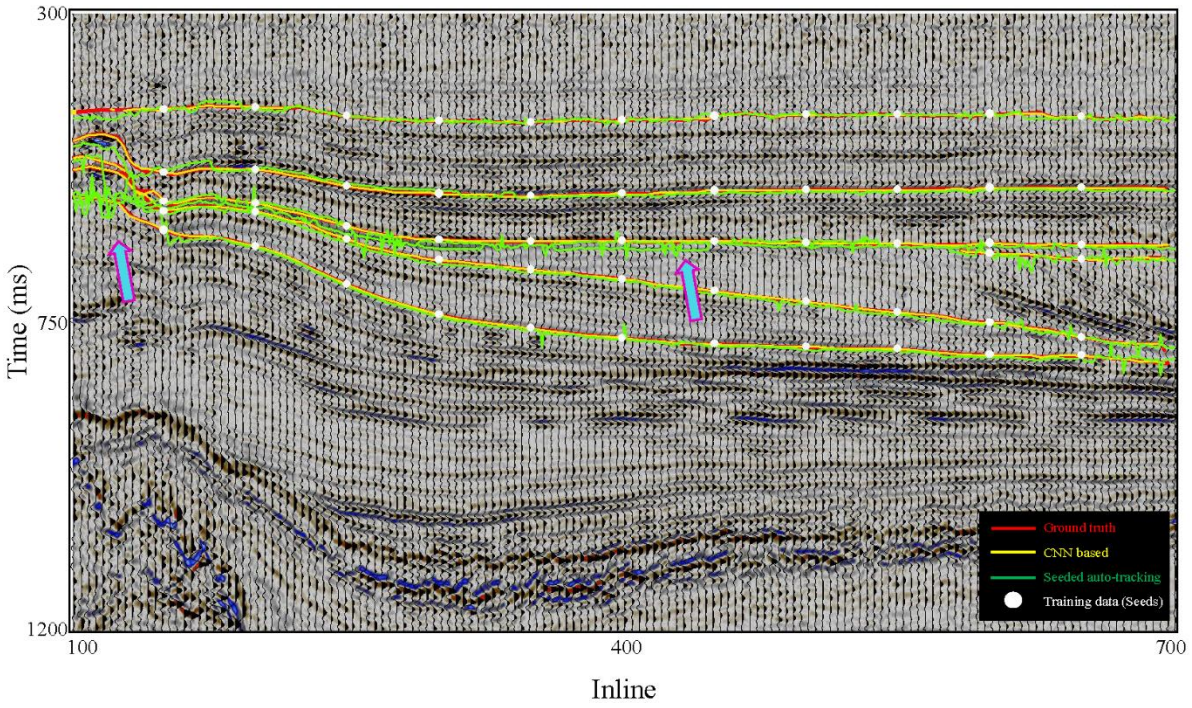


Figure 4.22. The comparison on a crossline section of the first field data.

For the first seismic survey, the predicted seismic horizons using our proposed method are shown in Figure 4.14. Figure 4.15 shows the interpolated seismic horizons using seeded auto-tracking. To evaluate the predicted results of seismic horizons, we calculate the absolute error of the time index between the ground truths with the predicted results. Figure 4.16 shows the absolute error maps between the ground truths with the predicted results using our proposed method that corresponds to the six horizons in the first field data. Figure 4.17 shows the absolute error maps between the ground truths with the predicted results using the seeded auto-tracking that corresponding to the six horizons in the first field data. Note that the predicted horizons using our proposed method have a very good match with the ground truths, while the predicted results using seeded auto-tracking have some obvious error with the ground truths that were indicated by purple arrows. We also plot the seismic amplitude on different horizon surfaces to evaluate whether the horizon laterally follows the consistent phase or not. Figures 4.18, 4.19 and

4.20 show the amplitude maps of seismic horizons using ground truth, our proposed method and seeded auto-tracking, respectively (the contour in the amplitude maps is caused by the onlap or downlap). Note that the amplitude maps of seismic horizons using our proposed method have a very good match with the amplitude maps of ground truths. However, the amplitude maps of seismic horizons using seeded auto-tracking have some inconsistent phase and mess area that indicated by purple arrows.

To further evaluate the performance of our proposed method, we show the examples of seismic sections from the direction of inline (Figure 4.21) and crossline (Figure 4.22), respectively. The interpretation results using our proposed method, seeded auto-tracking and manual picking are represented by yellow, green and red curves, respectively. The training data and training labels (seeds) are represented by white dots. Note that the method of seeded auto-tracking failed to detect the edge of the unconformities that indicated by the blue arrow. Also note that the seismic horizons obtained by using our proposed method have a very good match with the ground truth, while the seeded auto-tracking failed to follow the seismic horizons on the areas of faults and unconformities that indicated by blue arrows.

For the second seismic survey, the predicted seismic horizons using our proposed method are shown in Figure 4.23. Figure 4.24 shows the interpolated seismic horizons using seeded auto-tracking. Figure 4.25 shows the absolute error maps between the ground truths with the predicted results using our proposed method. Figure 4.26 shows the absolute error maps between the ground truths with the predicted results using the seeded auto-tracking. Note that the predicted horizons using our proposed method have a very good match with the ground truths, while the predicted results using seeded auto-tracking have some obvious error with the ground truths that were indicated by purple arrows. Figures 4.27, 4.28 and 4.29 show the amplitude maps of

seismic horizons using ground truth, our proposed method and seeded auto-tracking, respectively. Note that the amplitude maps of seismic horizons using our proposed method have a very good match with the amplitude maps of ground truths. However, the amplitude maps of seismic horizons using seeded auto-tracking have some inconsistent phase and mess area that indicated by purple arrows.

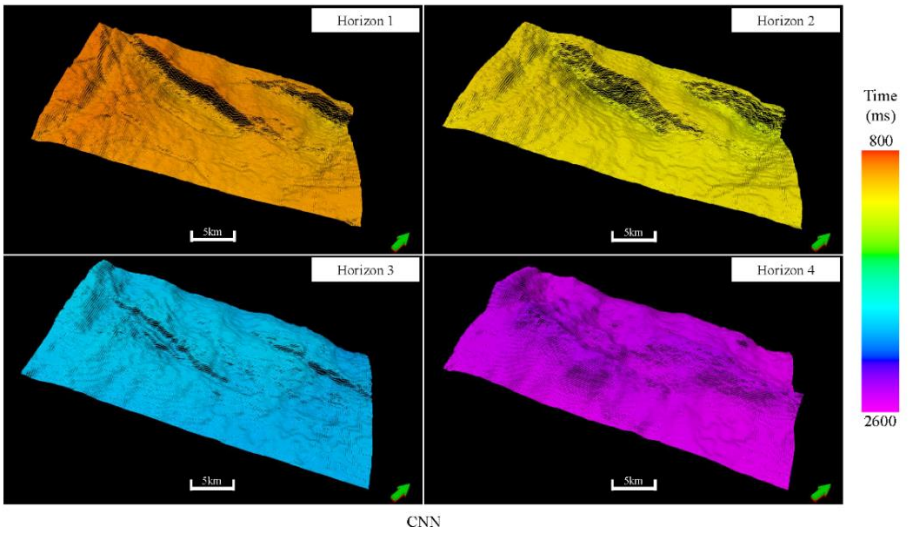


Figure 4.23. The interpreted seismic horizons using our proposed method of the second field data.

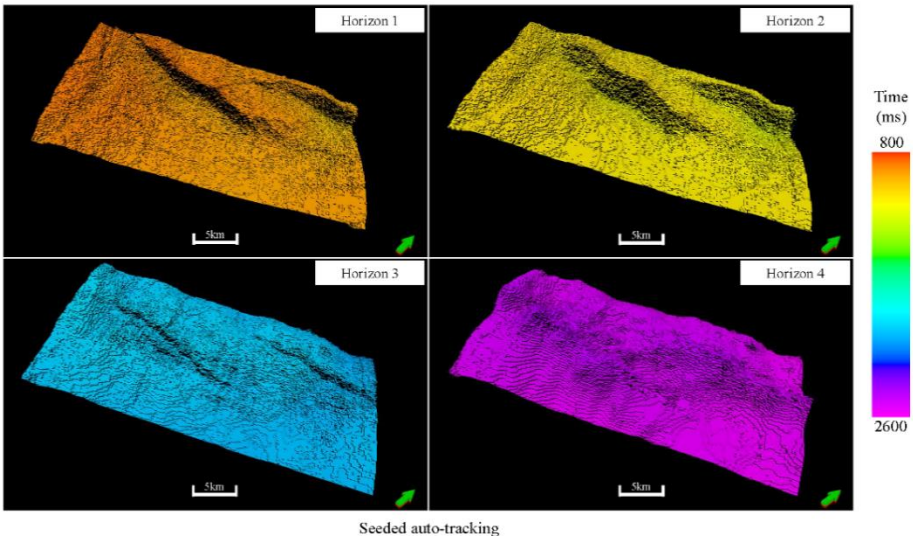
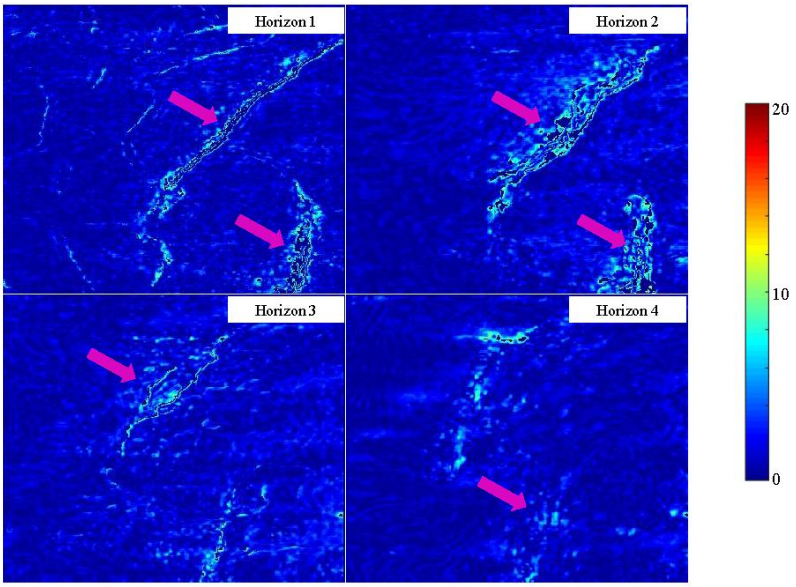
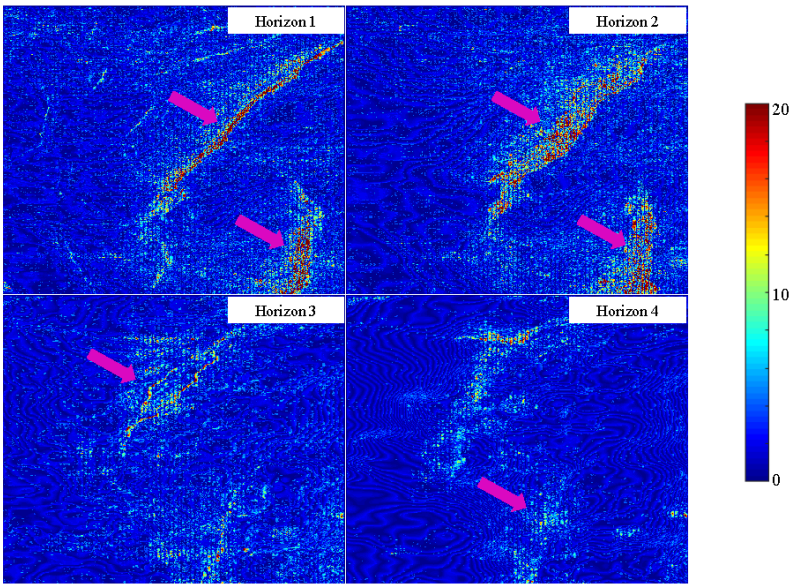


Figure 4.24. The interpreted seismic horizons using seeded auto-tracking of the second field data.



CNN

Figure 4.25. The absolute error maps between the ground truth and the results using our proposed method for the second field data.



Seeded auto-tracking

Figure 4.26. The absolute error maps between the ground truth and the results using seeded auto-tracking for the second field data.

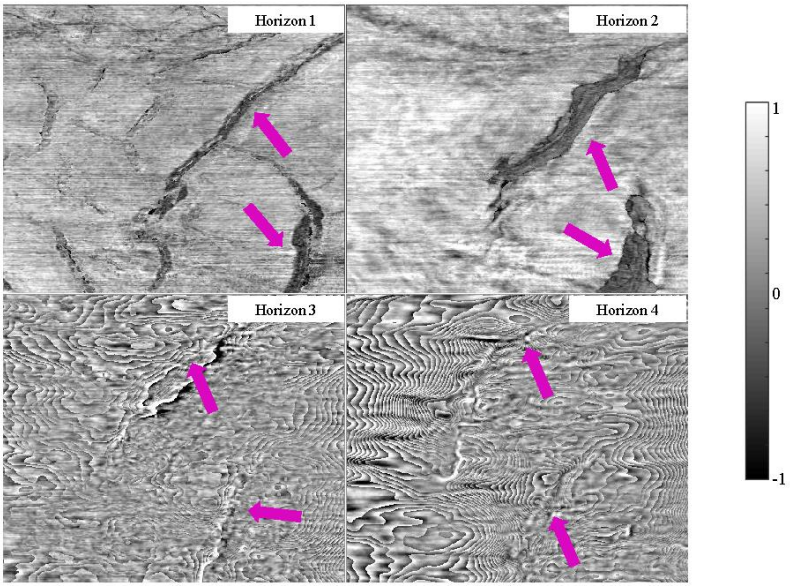


Figure 4.27. The seismic amplitude maps on the horizon surfaces of ground truth of the second field data.

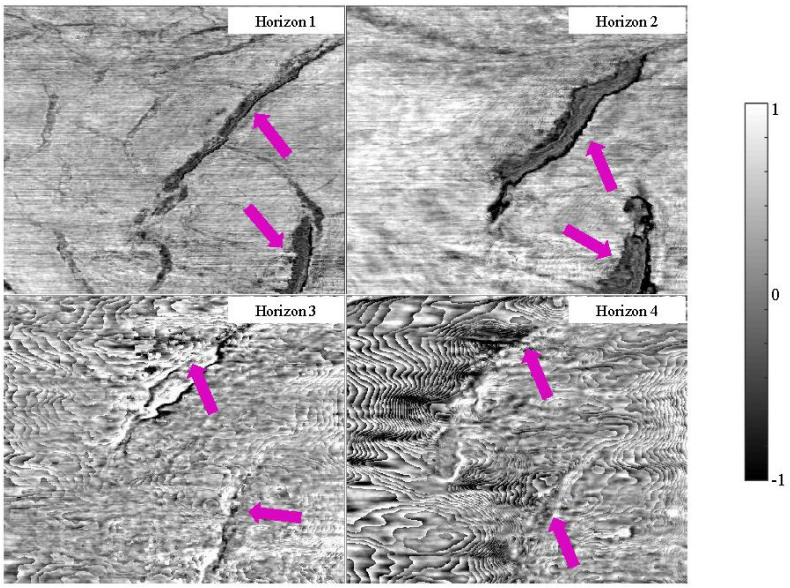


Figure 4.28. The seismic amplitude maps on the predicted horizon surfaces using our proposed method for the second field data.

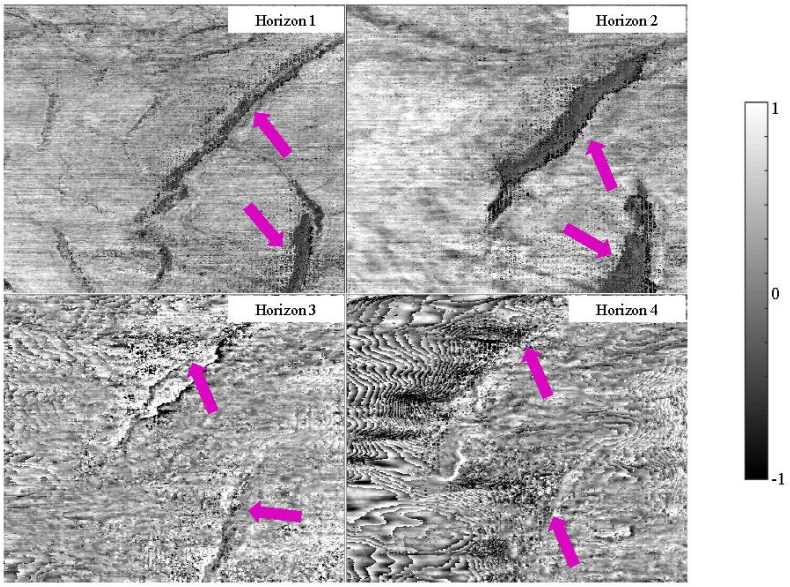


Figure 4.29. The seismic amplitude maps on the predicted horizon surfaces using our seeded auto-tracking for the second field data.

Figures 4.30 and 4.31 show the examples of seismic sections from the direction of inline and crossline, respectively. The interpretation results using our proposed method, seeded auto-tracking and manual picking are represented by yellow, green and red curves, respectively. The training data and training labels (seeds) are represented by white dots. Note that the seismic horizons obtained by using our proposed method have a very good match with the ground truth, while the seeded auto-tracking failed to follow the seismic horizons on the area of faults that indicated by blue arrows. In addition, the second real data test demonstrate that my deep learning model can be trained and further applied to 1D images with an arbitrary size. One key factor is the number of layers for building the deep convolutional neural network architecture. We obtained the best performance when the ratio between the size of the last layer and label number is four. We suggest that the size of the last layer should be between two to three times larger than the label number.

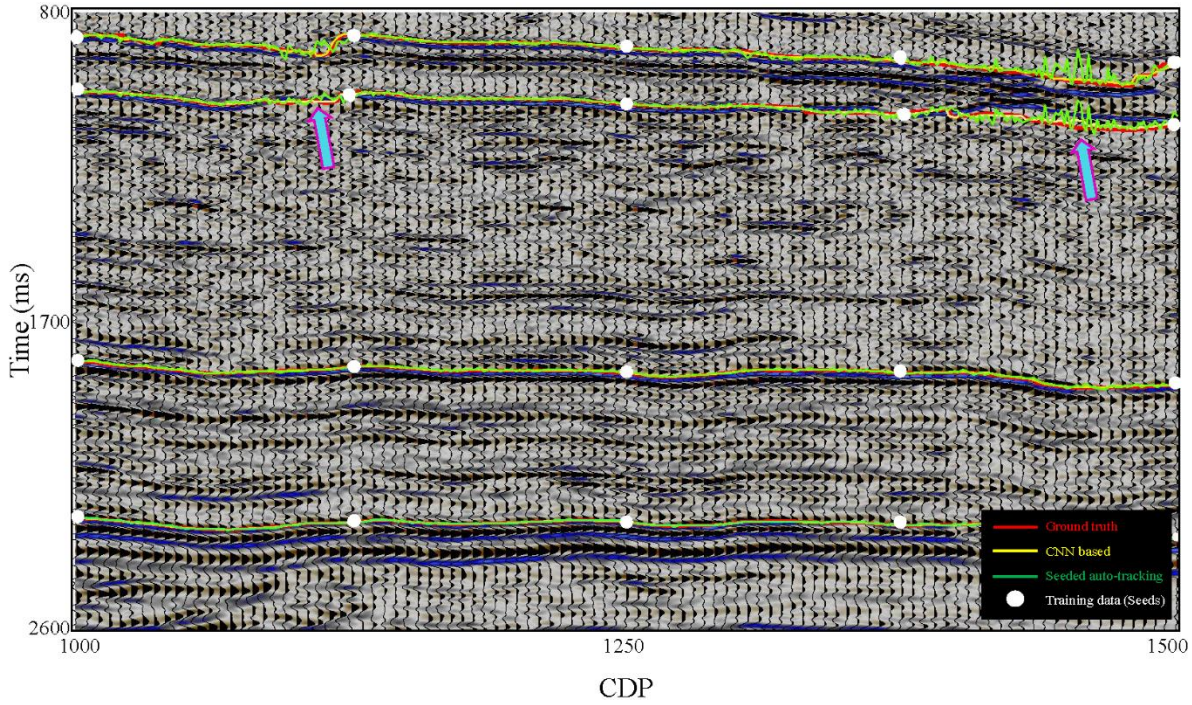


Figure 4.30. The comparison on an inline section of the second field data.

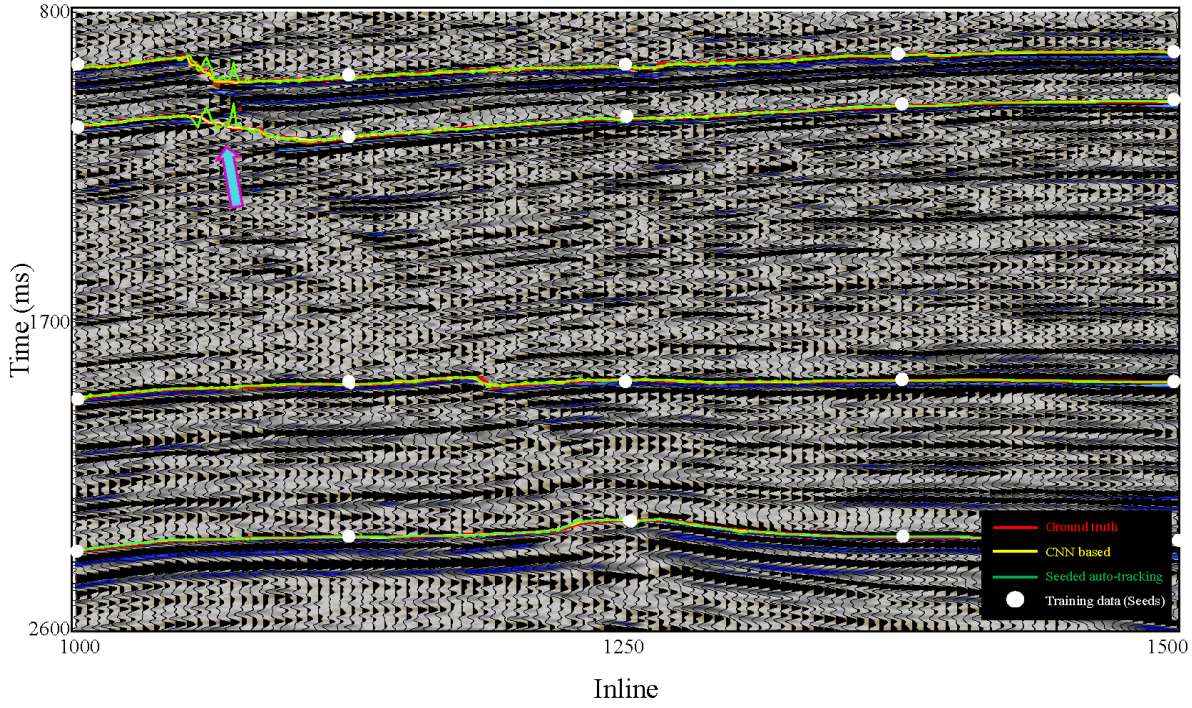


Figure 4.31. The comparison on a crossline section of the second field data.

CONCLUSIONS

We treat each seismic trace as a 1D image where the seismic horizon interpretation can be viewed as the task of image segmentation. We use the encoder-decoder convolutional neural network to successfully interpret the seismic horizons. The variable size of convolution kernel is more preferable than the commonly used fixed size in dividing the seismic traces into different segments where the boundaries of the segments are regarded as the horizons. We also designed a two-step filters to remove the spikes and keep the faults with rapid change. Two field data examples illustrate that our proposed method can produce accurate seismic horizons over the whole seismic survey even if we have faults and unconformities within the seismic surveys. The applications also demonstrate that our proposed method has comparable computation efficiency over the manual horizon interpretation.

REFERENCES

Badrinarayanan, V., A. Kendall, and R. Cipolla, 2015, SegNet: A deep convolutional encoder-decoder architecture for image segmentation, arXiv preprint.

Di, H., M. shafiq, and G. AlRegib, 2018, Patch-level mlp classification for improved fault detection: 88th Annual International Meeting, SEG, Expanded Abstracts, pp. 2211-2215.

Dorn, G. A., 1998, Modern 3-D seismic interpretation, The Leading Edge, vol. 17, pp, 1262-1262.

Fomel, S., 2010, Predictive painting of 3D seismic volumes, Geophysics, vol. 75, no. 4, pp. A25-A30.

Girshick, R., J. Donahue, T. Darrell, and J. Malik, 2014, Rich feature hierarchies for accurate object detection and semantic segmentation, IEEE Conference on Computer Vision Pattern Recognition, pp. 580-587.

He, K., X. Zhang, S. Ren, and J. Sun, 2015, Deep residual learning for image recognition, IEEE Conference on Computer Vision Pattern Recognition, pp. 770-778.

He, K., G. Gkioxari, P. Dollar, and R. Girshick, 2017, Mask R-CNN, arXiv preprint.

Huang, L., X. Dong, and T. E. Clee, 2017, A scalable deep learning platform for identifying geologic features from seismic attributes, The Leading Edge, vol. 36, pp. 249-256.

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

Kingma, D. P., and J. L. Ba, 2015, Adam: A method for stochastic optimization, International Conference on Learning Representations.

Krizhevsky, A., I. Sutskever, and G. Hinton, 2012, Imagenet classification with deep convolutional neural network, Advances in Neural Information Processing Systems, vol. 27, pp. 1090-1098.

LeCun, Y., L. Bottou, Y. Bengio, and P. Haffner, 1998, Gradient-based learning applied to document recognition, Proceeding of IEEE, vol. 86, no. 11, pp. 2278-2324.

Lomask, J., A. Guitton, S. Fomel, J. Claerbout, and A. A. Valenciano, 2006, Flattening without picking, Geophysics, vol. 71, no. 4, pp. P13-P20.

Long, J., E. Shelhamer, and T. Darrell, 2015, Fully convolutional networks for semantic segmentation, IEEE Conference on Computer Vision Pattern Recognition, pp. 3431-3440.

Parks, D., 2010, Seismic image flattening as a linear inverse problem, M.S. thesis, Colorado School of Mines.

Paszke, A., A. Chaurasia, S. Kim, and E. Culurciello, 2016, ENet: A deep neural network architecture for real-time semantic segmentation, arXiv preprint.

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, Springer, pp. 234-241.

Stark, T. J., 2003, Unwrapping instantaneous phase to generate a relative geologic time volume, 73rd SEG Annual International Meeting Expanded Abstracts, pp. 1707-1710.

Simonyan, K., and A. Zisserman, 2015, Very deep convolutional networks for large-scale image recognition, International Conference on Learning Representations.

Szegedy, C., W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Auguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, 2015, Going deeper with convolutions, arXiv preprint.

Wu, X., and D. Hale, 2015, Horizon volumes with interpreted constraints, Geophysics, vol. 80, no. 2, pp. IM21-IM33.

Wu, X., and G. Zhong, 2012, Generating a relative geologic time volume by 3D graph-cut phase unwrapping method with horizon and unconformity constraints, Geophysics, vol. 77, no. 4, pp. O21-O34.

Wu, X., and S. Fomel, 2018, Least-squares horizons with local slopes and multi-grid correlations, Geophysics, vol. 84, no. 4, pp. IM29-IM40.

Wu, X., and Y. Shi, S. Fomel, and L. Luming, 2018, Convolutional neural networks for fault interpretation in seismic images, 88th Annual International Meeting, SEG, Expanded Abstracts, pp. 1946-1950.

Zeng, H., M. M. Backus, K. T. Barrow, and N. Tyler, 1998, Stratal slicing; Part 1, Realistic 3D seismic model, Geophysics, vol. 63, no. 2, pp. 502-513.

Zhao, T., and P. Mukhopadhyay, 2018, A fault-detection workflow using deep learning and image processing, 88th Annual International Meeting, SEG, Expanded Abstracts, pp. 1966-1970.

CHAPTER 5
CONCLUSIONS

My dissertation illustrates that it is very promising to employ deep learning in assisting seismic data processing and interpretation. Properly transfer the geoscience problem to objects identification problem is the key to successfully apply deep learning in geoscience world. Meanwhile, the applications demonstrates that post processing using geoscience knowledge is necessary to further improve the accuracy. The main conclusions of this dissertation are summarized as follows.

In Chapter 2, I transfer the first arrival picking problem into an image segmentation problem and apply the pixel-wise convolutional method to address this problem. Most current convolutional neural networks are using constant size of convolution filters. The application demonstrate that the accuracy of predicted first arrival using constant size of convolution filters are very low. I propose to use variable size of convolution filters to build the neural network architecture. The comparison illustrates that using variable size of convolution filters has successfully improve the accuracy of predicted first arrival. Recently, many automatic first arrival picking methods have been proposed. However, those method are very hard to obtain the accurate result in the really noisy environment. Both synthetic and real data applications illustrate that our proposed method is superior to the traditional method of STA/LTA and can save massive labor work than manual interpretation.

In Chapter 3, I propose a novel seismic noise attenuation method (VMD-MDnCNN) by integrating our MDnCNN with VMD. Current CNN based denoising methods either require the

label of clean seismic data or the label of noise contained in the seismic data. My method does not require clean seismic label nor purely noise label. The applications demonstrate that the white noise contained in the seismic can be simulated by enough user-generated white noise realizations. In addition, the applications demonstrate that the MDnCNN can obtain a more accurate estimation of the noise feature from the decomposed bandlimited seismic data. Both synthetic and real seismic data applications illustrate that our method is superior to the traditional denosing method of *f-x* deconvolution. The applications also demonstrate that our method not only effectively reject the white noise but also the migration artifacts contained in the seismic data.

In Chapter 4, I solved several challenges of using deep learning techniques for seismic horizon interpretation. First, treat the seismic trace as a 1D image eliminates the effect of size of seismic survey on the hierarchy building. Second, cropping the seismic traces into a constant size reduce the effect of recording length on the hierarchy building. Third, using a downward-upward trend size of convolution filters help to improve the prediction accuracy. Fourth, the built hierarchy is universal and applicable to other horizon interpretation. Finally, post process with the aid of geoscience knowledge is necessary to further improve the accuracy of horizon interpretation.