GitHub repository: https://github.com/acse-2019/irp-acse-pt919

Imperial College London

Department of Earth Science and Engineering

MSc in Applied Computational Science and Engineering

Independent Research Project

Project Plan

Weakly-supervised Learning of Pixel-level Labelling for Salt Dome

by

Ping-Chen Tsai

Ping-chen.tsai19@imperial.ac.uk

GitHub login: acse-pt919

June 2020

First Supervisor: Dr. Erik Barrow

Second Supervisor: Navjot Kukreja

Ovation Data Ltd

Unit10, Crayfields Industrial Park

Main Road, St Paul’s Cray

Kent, BR53HP

GitHub repository: https://github.com/acse-2019/irp-acse-pt919

Background and Introduction

The Oil and Gas Industry has long shown an interest in the interpretation of seismic structure to find new hydrocarbon reserves. Potential oil and gas reserve exist when hydrocarbon accumulates in a geological structure called trap. There are two major categories of traps: structural traps and stratigraphic traps. Common seismic traps are salt domes, channels, folds(anticline) and faults. These traps are generally combinations of rock structures that have upper and lower boundaries to keep oil and gas from escaping. The detection of traps from seismic surveys is an indispensable aspect of seismic interpretation since useful information about oil and gas reserves could be extracted for the interpretation. Manual seismic interpretation is labour-intensive and time-consuming due to the large volume of data and complexity of subsurface structures. Thus emerges the need for automatic traps detection to improve the exploration of oil and gas reserves.

Seismic attributes extracted from seismic slices capture subsurface responses and identify geologic features. When seismic reflection amplitudes are either noisy or affected by parameters such as porosity and fluid content, texture attribute analysis is highly recommended for automatic seismic trap detection, since texture attribute are not sensitive to amplitude *(Chopra and Alexeev, 2006)*. As the quantity of seismic attributes increases, there is an explosion in the complexities of relationships regarding attributes interdependencies. This urges machine-learning-based solutions to automate the process of seismic traps detection. Today, the proposed methods for supervised seismic trap detection and segmentation have achieved impressive performance.

Kumar & Sain (2018) discriminated faults from an entire seismic data cube using attribute amalgamation through a fully connected multilayer perceptron (MLP) based on artificial neural network (ANN). Similar work has been conducted by *Shafiq et al. (2015)* to illuminate salt bodies based on 3D gradient of texture (GoT) in salt boundaries.

Another popular approach to illuminate traps is using image segmentation techniques based on deep learning. A simple approach is to use a bounding box around the identified trap, but bounding box is not enough for fine-grained tasks when there is need to outline the precise trap boundaries. Applying semantic segmentation to interprete stratigraphic traps at pixel level has been attempted widely based on simple CNN model (*Chevitarese et al., 2018)*, U-Net combined with ResNet and DenseNet model *(Milosavljević, 2020)* and fully convolutional network *(Long, J., Shelhamer, E. & Darrell, T., 2015)*.

However, full supervision for training seismic trap segmentation relies on enormous labelled data. It would be a very costly task to prepare large volume of labelled data for training a neural network. Plus, the limited number of geological domain experts constrains the efficiency of manual production of labelled seismic data and leaves large quantity of datasets behind. To overcome such obstacle, unsupervised approaches are proposed by many others such as self-organizing map *(Miller M.S. and Powell K.S., 2001)* and deep convolutional autoencoder *(Qian et al., 2018)*. The challenge is that many of the proposed unsupervised works involved training deep learning models that have complex architecture and millions of free parameters. In addition to unsupervised detection using bounding box, unsupervised segmentation would be more computationally expensive and unforeseeable regarding testing accuracy.

Apart from unsupervised approach, weakly-supervised approaches use coarsely labelled and less-informative (weakly-labelled) seismic slices to perform even more promising results in trap classification, detection, and segmentation on pixel-level *(Alaudah et al., 2018). Papandreou et al. (2015)* suggested that labelling by bounding box is 15 times faster/cheaper than labelling at pixel-level for a seismic interpreter, and it is even cheaper for image-level labelling. Therefore, a weakly-supervised approach is more economical to interpret seismic data when the labelled dataset is limited to either quantity or image-level annotation. In the past few years, weakly-supervised approaches have achieved a great success in many computer vision applications due to its low-demanding training input *(Papandreou et al. 2015).* The pixel-level annotations produced by weakly-supervised approach can then support training fully supervised trap detection/segmentation algorithms for more accurate models.

Project Description

To relax the training dependency on huge volumes of pixel-level labels, the goal of this project is to automatically provide precise positional information of a seismic trap in an image label by segmenting the seismic trap, specifically salt dome regions, at pixel levels. Due to low amount of people power on manual labelling at Ovation, using weak label, i.e. bounding box, for training would be more efficient than using pixel-based labelling. The project results ought to help the company reduce the effort and time put on manual trap segmentation. In addition to work reduction, the results could further be used to train more powerful and informative trap segmentation models by fully supervised methods.

Aim

The project aims to reproduce *Alaudah& AlRegib (2017)*’s work to obtain pixel-level segmentation from coarse image-level classes. The project will take bounding box annotated seismic images as weak label. In the end, the project solution will develop a weakly-supervised learning approach using Orthogonal Non-Negative Matrix Factorization (ONMF) with sparseness constraints. The algorithm can ultimately help with the detection of hydrocarbon reserve when facing a lack of finely labelled seismic dataset.

Method

On basis of Non-Negative Matrix Factorization (NMF, a dimensionality reduction method in machine learning), the proposed ONMF method can produce sparse representation of data *(Hoyer, P. O., 2004)*. A confident number of training dataset would be 1000 seismic slices per class. Since the input data matrices of seismic slices are nonnegative, ONMF outbids classical dimensionality reduction tools such as singular value decomposition (SVD) and principle component analysis (PCA) to maintain nonnegative matrices entries at a lower rank from original input matrices. K-means clustering centroids are taken to form the basis vectors at initialisation stage of the factorization to represent features of each input seismic slice. A feature matrix therefore could be constructed by combining basis vectors in columns. Then use multiplicative update algorithm to optimize the feature matrix. The sparseness of the resulting matrix signifies the salt dome features in the weakly-labelled images to facilitate seismic interpretation *(Alaudah& AlRegib, 2017)*. With weak labels of trap, the automatically generated pixel-level segmentations will be able to support fully supervised models training for various seismic interpretation tasks. As for result evaluation, internal validation will be implemented to validate the quality of algorithm.

Workflow and Objectives

The workflow of the project mainly is image pre-processing, dimensionality reduction and noise elimination. The raw dataset from Ovation is told to be segy and JSON files. Before Ovation providing the original labelled dataset, a backup dataset 'LANDMASS' (CeGP, 2015) is always ready for experiments.

Basic objectives:

1. Pre-process raw seismic data to obtain image-level label for salt domes in .png format, i.e. images.
2. Review concepts ONMF-related concepts such as clustering and sparse coding.
3. Implement ONMF algorithm to obtain pixel-level labels for every seismic images.
4. Visualize results by overlaying the smoothed (noise reduction) labels to the original images

Advanced objective:

1. Perform label smoothing process on stacked 2D slices.

GitHub repository: https://github.com/acse-2019/irp-acse-pt919

Work Progress to Date and Schedule

The Gantt chart in figure 1 lists work progress and future plan for this project starting from 01/06/2020.

			w1-w4	week 5	week 6	week 7	week 8	week 9	week 10	week 11	week 12	week 13
ID	Task	Duration(Day)	1-26/6	29/6- 3/7	6/7 - 10/7	13/7 - 17/7	20/7 - 24/7	27/7 - 31/7	3/8 - 7/8	10/8 - 14/8	17/8 - 21/8	24/8 - 28/8
1	Project background reviews, find opensource datasets and Python package tutorials	Planned: 26										
		Actual: 26										
2	Learn to work on Linux VM; Labelling + saving dataset on Ovation VM	Planned: 7										
		Actual:										
3	Data preprocessing to obtain image-level labels: Turn .segy to .png. Use .JSON to generate single-label images	Planned: 13										
		Actual:										
4	Concept reviews: clustering & ONMF	Planned: 20										
		Actual:										
5	Clustering experiment on backup dataset in case Ovation's data isn't ready	Planned: 7										
		Actual:										
6	Simulate ONMF based on clustering result	Planned: 10										
		Actual:										
7	Implement ONMF to obtain pixel-level salt dome labels	Planned: 15										
		Actual:										
8	Use image smoothing method, i.e. median filter, to reduce noise in pixel-level labels.	Planned: 7										
		Actual:										
9	Apply Internal validation in algorithm	Planned: 7										
		Actual:										
	Document code, clean up code	Planned: 35										
		Actual:										
10	Write final report	Planned: 45										
		Actual:										

Figure 1: Gantt chart for the project. Skin colour label for planned work duration, red label for actual works duration. The duration of the whole project lasted for 13 weeks. Future schedule starts from week 5, 29/6/2020, to week 13, 28/8. The final submission is due on 28/8 17:00 UK time.

GitHub repository: https://github.com/acse-2019/irp-acse-pt919

Bibliography

Alaudah, Y., Gao, S. & AlRegib, G. (2018) Learning to label seismic structures with deconvolution networks and weak labels. In: SEG Technical Program Expanded Abstracts 2018. *27 August 2018 Society of Exploration Geophysicists*. p.

Alaudah, Y. & AlRegib, G. (2017) A weakly supervised approach to seismic structure labeling. *In: SEG Technical Program Expanded Abstracts 2017*. 17 August 2017 Society of Exploration Geophysicists. p.

CeGP (2015) LANDMASS: large north-sea dataset of migrated aggregated seismic structures. Available at: http://cegp.ece.gatech.edu/codedata/LANDMASS/default.htm.

Chevitarese D., Szwarcman D., Silva R. M. D, & Vital Brazil E. (2018) Seismic facies segmentation using deep learning. *In American Association of Petroleum Geologists* (AAPG).

Chopra, S. & Alexeev, V. (2006) Applications of texture attribute analysis to 3D seismic data. *The Leading Edge*, 25(8), pp.934-940.

Hoyer, P. O. (2004) Non-negative Matrix Factorization with Sparseness Constraints. *The Journal of Machine Learning Research*. 5, pp1457-1469.

Kumar, P.C. & Sain, K. (2018) Attribute amalgamation-aiding interpretation of faults from seismic data: An example from Waitara 3D prospect in Taranaki basin off New Zealand. *Journal of Applied Geophysics.* 159, 52–68.

Long, J., Shelhamer, E. & Darrell, T. (2015) Fully convolutional networks for semantic segmentation. *In: 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. June 2015 IEEE. p.

Miller M.S.& Powell K.S. (2001) Seismic interpretation and processing applications. *Handbook of Geophysical Exploration: Seismic Exploration*, Vol. 30, Elsevier, pp. 101-118

Milosavljević, A. (2020) Identification of Salt Deposits on Seismic Images Using Deep Learning Method for Semantic Segmentation. *ISPRS International Journal of Geo-Information*. 9 (1), 24.

Pitas, I. & Kotropoulos, C. (1992) A texture-based approach to the segmentation of seismic images. *Pattern Recognition*. 25 (9), 929–945.

Qian, F., Yin, M., Liu, X.-Y., Wang, Y.-J., et al. (2018) Unsupervised seismic facies analysis via deep convolutional autoencoders. *GEOPHYSICS*. 83 (3), A39–A43.

GitHub repository: https://github.com/acse-2019/irp-acse-pt919

Shafiq, M.A., Wang*, Z., Amin, A., Hegazy, T., et al. (2015) Detection of Salt-dome Boundary Surfaces in Migrated Seismic Volumes Using Gradient of Textures. *In: SEG Technical Program Expanded Abstracts 2015*. 19 August 2015 Society of Exploration Geophysicists. p.

Papandreou G., Chen L., Murphy K., Yuille AL. (2015) Weakly and semi-supervised learning of a DCNN for semantic image segmentation. CoRR abs/1502.02734, 1502.02734