Learning to Label Seismic Structures with Deconvolution Networks and Weak Labels

Yazeed Alaudah, Shan Gao, and Ghassan AlRegib
{alaudah,gaoshan427,alregib}@gatech.edu
Center for Energy and Geo Processing (CeGP) at Georgia Tech and KFUPM

SUMMARY

Recently, there has been increasing interest in using deep learning techniques for various seismic interpretation tasks. However, unlike shallow machine learning models, deep learning models are often far more complex and can have hundreds of millions of free parameters. This not only means that large amounts of computational resources are needed to train these models, but more critically, they require vast amounts of labeled training data as well. In this work, we show how automatically-generated weak labels can be effectively used to overcome this problem and train powerful deep learning models for labeling seismic structures in large seismic volumes. To achieve this, we automatically generate thousands of weak labels and use them to train a deconvolutional network for labeling fault, salt dome, and chaotic regions within the Netherlands F3 block. Furthermore, we show how modifying the loss function to take into account the weak training labels helps reduce false positives in the labeling results. The benefit of this work is that it enables the effective training and deployment of deep learning models to various seismic interpretation tasks without requiring any manual labeling effort. We show excellent results on the Netherlands F3 block, and show how our model outperforms other baseline models.

INTRODUCTION

Many machine learning based methods have been proposed for various seismic interpretation tasks. Many of these techniques are used to extract or interpret localized seismic structures such as salt domes, faults, horizons, and so on. AlRegib et al. (2018) provides a good overview of such methods—and new emerging techniques—from a signal processing perspective. Given all this interest in automating various seismic interpretation tasks, not much work has been done on labeling *entire* seismic volumes based on their dominant structures. Alaudah and AlRegib (2016) presented an early attempt at this using similarity-based image retrieval and a support vector machine (SVM) classifier. Additionally, Rutherford Ildstad and Bormann (2017) recently published code for facies classification using a basic 5-layer convolutional neural network (CNN).

In recent years, deep learning has witnessed great success in wide-ranging applications and has revolutionized the fields of machine learning and computer vision. This success was not only due to the growing use of powerful GPUs, or the advent of deep learning models that can learn their own hierarchical data representations; but also, and more importantly, the arrival of very large labeled datasets, such as ImageNet (Russakovsky et al., 2015). Deep learning models are often far more complex than traditional machine learning models and can have hundreds of millions of free parameters. This not only means that they need large amounts of computational resources to train

these models, but more critically, they require vast amounts of labeled training data. Labeled data can be extremely costly and time-consuming to obtain. In practice, the high cost of acquiring labeled data is a critical bottleneck to the successful application of deep learning to many application domains. This bottleneck is especially true for the field of seismic interpretation, where very few *labeled* datasets are freely available.

There have been quite a few techniques published recently that apply deep learning algorithms, such as CNNs, to seismic interpretation problems (e.g., Waldeland and Solberg (2017); Araya-Polo et al. (2017); Huang et al. (2017); Di et al. (2018). However, all *these methods require "strong" labels*[1] that are obtained by manual labeling from an interpreter. Manually-labeling data for training deep learning models can be as laborious and time-consuming as manual interpretation workflows. Furthermore, over-training a network on a relatively small amount of manually labeled data can easily lead to over-fitting, and therefore poor generalization performance.

One solution is to use weakly-supervised methods that do not require manual labeling efforts and can produce orders of magnitude more labeled data than manual labeling by domain experts. However, these labels are usually of lower quality than manually obtained ones. In this paper, we build on our previous work on generating weak labels for seismic interpretation (Alfarraj et al., 2016; Alaudah and AlRegib, 2017; Alaudah et al., 2018). Here, we show that despite their lower quality, weak labels are extremely useful in enabling the training of complex deep learning models—such as deconvolution networks—with no manual labeling effort required. The only requirement is for the interpreter to "define" the various structures of interest to her by selecting at least one image for every seismic structure (as is shown in the *first row* of Figure 1).

To summarize, our main contributions in this paper are:

- We use automatically-generated weak labels to train deep learning models with no manual labeling effort required.

- We propose using a deconvolution network, trained on weak labels, to accurately localize and classify seismic structures.

- We modify the network's loss function to avoid putting too much trust in our weak labels, and therefore, reduce false positives.

In the next section, we will explain these contributions in detail; and in the results section, we will show how these contributions help achieve excellent results on the Netherlands North Sea F3 block.

---

[1] Here, "strong" labels mean high-quality labels generated by a domain expert. This is opposed to automatically-obtained "weak" labels that convey far less information than strong ones, and are usually much noisier and less accurate, but are much easier to obtain. Please see (Alaudah and AlRegib, 2017) for a more detailed explanation.
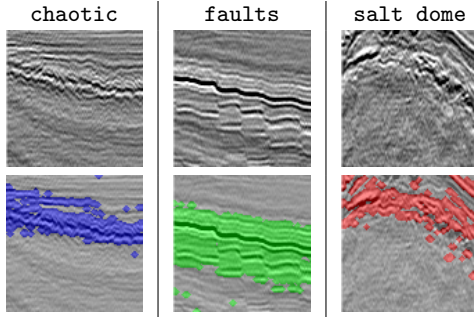
Figure 1: Sample results from our weakly-supervised label mapping technique in (Alaudah and AlRegib, 2017). The first row shows the original seismic images, with their image-level labels above. The second row shows the resulting pixel-level labels for these images.

METHOD

Weak labels

In our previous work (Alfarraj et al., 2016) we showed how seismic image similarity measures could be used to retrieve thousands of images based on their visual similarity. Using this technique, we can obtain thousands of image-level labeled seismic images that contain structures such as horizons, faults, chaotic layers, and salt domes. We then showed how we these image-level labels can be mapped into much more accurate pixel-level labels by solving a simple non-negative matrix factorization problem (Alaudah and AlRegib, 2017)[2].

Figure 1 shows examples of the results obtained from our previous work. The figure shows seismic images containing chaotic structures, faults, and salt domes with their corresponding pixel-level labels in blue, green, and red, respectively. While these labels are not as accurate as ones obtained from an expert interpreter, obtaining vast amounts of weak labels, such as these, is rather easy as we've shown in our previous work. Additionally, our method (Alaudah and AlRegib, 2017) allows us to easily quantify the confidence in the various pixel-level labels generated for every image. We refer to this confidence in the weak pixel-level labels as $q(x)$, where x is the pixel index.

In addition to our thousands of weakly-labeled images, we further increase the size of our training dataset by using several data augmentation techniques. These techniques include random horizontal flipping, and random rotations of up to $\pm 15°$, of all the images in our dataset and their corresponding labels. This data augmentation step helps to prevent our model from over-fitting to the training data, and helps it generalize to data it hasn't seen before.

Deconvolution networks

A major hurdle for the successful end-to-end application of CNNs for labeling visual data was what seemed like a trade-off between classification and localization accuracy. Deeper

networks that have many convolution and pooling layers have proven to be the most successful models in image classification tasks. However, their large receptive fields and increased spatial invariance (due to pooling and convolutional layers) make it difficult to infer the locations of various objects within the image. In other words, the deeper we go into a network, the more it seemed we lose the location information of various objects within the image. Some researchers have attempted to overcome this hurdle by using various pre- or post-processing techniques. However, the recent introduction of fully convolutional network architectures, such as FCN (Long et al., 2015) and DeconvNet (Noh et al., 2015) have shown that it is possible to achieve good labeling results using a convolutional network only, with no pre- or post-processing steps required. FCN accomplish this by replacing the fully-connected layers of the CNN with 1D convolutional layers that produce coarse feature maps. These coarse feature maps are then upsampled, and concatenated with the scores from intermediate feature maps in the network to generate the output. These upsampling steps, however, result in a blurred output that loses some of the resolution of the original image.

Deconvolution networks overcome this problem by using a symmetric encoder-decoder style architecture composed of stacks of convolution and pooling layers in the encoder, and stacks of deconvolution and unpooling layers in the decoder that mirror the encoders architecture. The role of the encoder can be seen as doing object detection and classification, while the decoder is used for accurate localization of these objects. This architecture can achieve finer and more accurate results than those of the FCN, and therefore is adopted in our work.

Figure 2 outlines the architecture of the deconvolution network used in our work. Every convolution or deconvolution layer (in white) is followed by a rectified linear unit (ReLU) non-linearity. The layers in red perform 2×2 max pooling to select the maximum filter response within small windows. The indices of the maximum responses for every pooling layer are then shared with their respective unpooling layers (in green) to undo this pooling operation and get a higher resolution image.

Adapting the loss function for learning with weak labels

Since our weak labels are generated automatically, they are not of the same quality as labels obtained from an expert interpreter. However, since obtaining such labels does not require any manual labor nor expensive computational resources, we can use these labels to train our model and modify our network loss function to not trust these weak labels too much. To achieve this, we use a recently introduced loss function called the focal loss (Lin et al., 2017) that was recently proposed for dense detection of objects in computer vision tasks. If we write the widely-used cross-entropy loss as

$$CE(p(x), q(x)) = -\sum_x q(x) \log(p(x)), \qquad (1)$$

where $p(x)$ is the confidence of the network output, $q(x)$ is the confidence of the weak labels, and x are the pixels in the image. Then, he focal loss can be written as

$$FL(p(x), q(x)) = -\sum_x (1 - p(x))^\gamma CE(p(x), q(x)), \qquad (2)$$

---

[2]Both works are better explained in our journal paper currently under review (Alaudah et al., 2018).
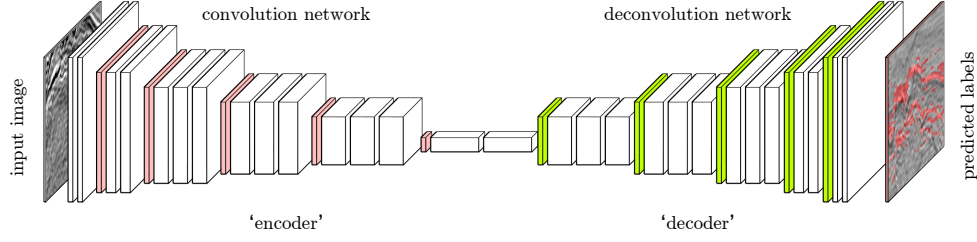
Figure 2: The architecture of the deconvolution network used in this work. White layers are convolution or deconvolution layers. Red layers are max-pooling layers, while green layers are unpooling layers.
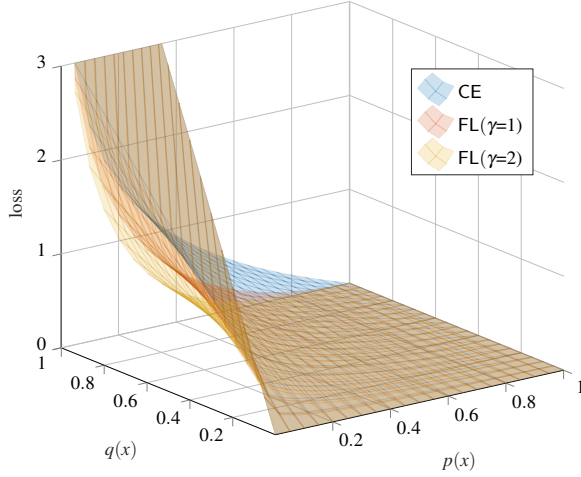

Figure 3: An illustration of the difference between the cross-entropy loss (CE) and the focal loss (FL) for different values of γ. $q(x)$ is the confidence of the weak training labels, while $p(x)$ is the confidence of the network output.

where γ is a parameter that controls how much weight is given to regions with low predicted confidence. We use this loss function, as opposed to the more cross-entropy (CE) loss, to put more weight on misclassified regions in the images and not trust our weak labels as much. Later in the results section, we compare the two loss functions and show how the focal loss can greatly enhance the results when training with weak labels. Figure 3 shows a comparison of CE with FL for two different values of γ. As the value of γ increases, less emphasis is put on regions where the network has learned the seismic structures relatively well, but not to the degree where they match the weak labels exactly. Instead, more emphasis is put on regions where the network has not learned to classify the underlying structure effectively.

RESULTS

We train our deconvolution network, shown in Figure 2, on thousands of automatically generated weak labels similar to those shown in Figure 1. Throughout our training, we set aside 25% of the training data for model selection and validation

purposes. Once our model's parameters are selected, we re-train our network on the entirety of the training data.

Once our deconvolution network is trained, we apply it to the Netherlands F3 block (dGB Earth Sciences, 1987) in a sliding window fashion to label the various subsurface structures in the data. This is done both in the inline and the crossline direction; then the final results are obtained by taking the element-wise product of the two. This step helps reduce any false-positive classifications. Figure 4 shows the final results obtained by our model for highlighting `chaotic`, `faults`, and `salt dome` structures in the F3 block. Due to the limited space, only results for crossline #150 are shown. Additionally, Figure 5 shows a 3D cross-section of the F3 block with the boundaries of several salt domes highlighted. We note that our model highlights only the salt dome boundaries, and that there are hardly any false positives present in the entire volume.

Additionally, Figure 6 shows the results of labeling the `faults` class in inline #1 of the F3 block using the FCN-8s network or the deconvolution network with CE or FL losses. FCN-8s is the best performing variant of the FCN architecture proposed by Long et al. (2015). We notice that due to the upsampling operation in FCN, several faults in the inline where not labeled. In addition, by comparing (b) and (c) we note how FL helps reduce false positives by not trusting the weak training data too much.

CONCLUSIONS

In conclusion, we showed how automatically-generated weak labels can be used to train a deep deconvolution network for labeling various subsurface structures within large seismic volumes. We also showed that modifying the loss function of the network to account for the weak labels can help the network avoid false positives, and increase the overall accuracy and robustness of the model.
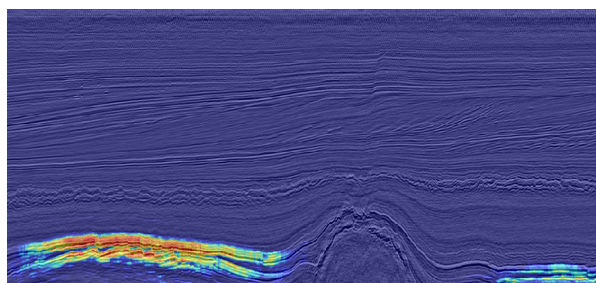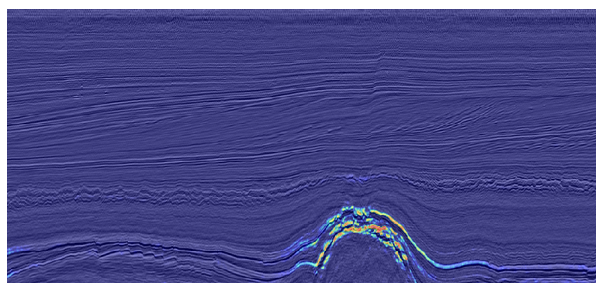
ACKNOWLEDGEMENTS
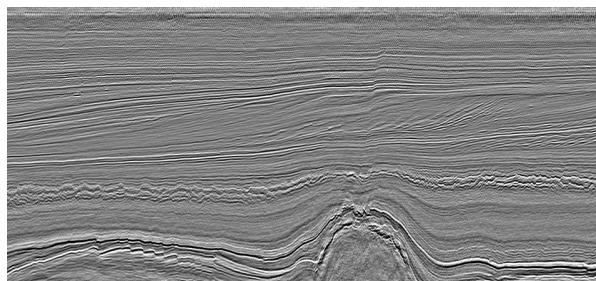
(a) original seismic
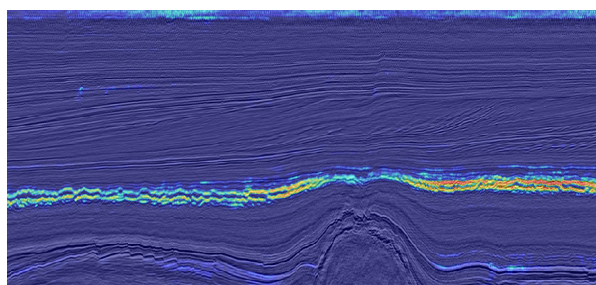
(b) chaotic class highlighted

(c) faults class highlighted

(d) salt dome class highlighted

Figure 4: Results using our model to highlight various subsurface structures in crossline #250 of the Netherlands F3 block.
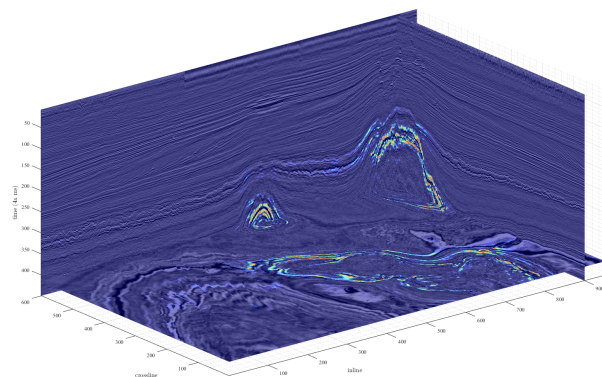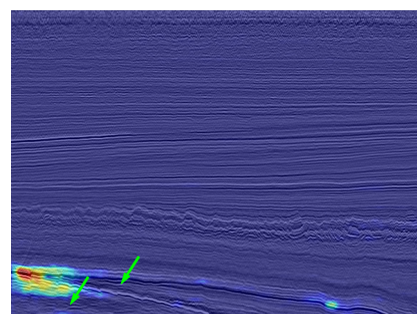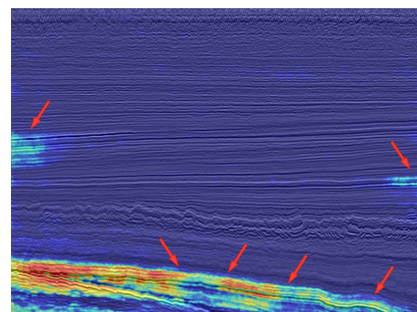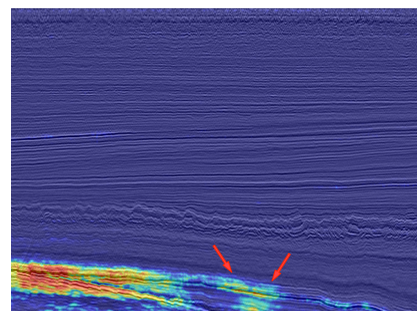

Figure 5: A 3D view of the Netherlands F3 block, with our model highlighting different salt dome structures.

(a) FCN-8s using CE

(b) Deconvolution network using CE

(c) Deconvolution network using FL (ours)

Figure 6: Fault structures in inline #1 highlighted using either deconvolution network or FCN-8s, and using either the cross entropy loss (CE) or the focal loss (FL). Green arrows indicate false negatives, while red arrows indicate false positives.

REFERENCES

Alaudah, Y., M. Alfarraj, and G. AlRegib, 2018, Structure label prediction using similarity-based retrieval and weakly-supervised label mapping: Geophysics [under review].

Alaudah, Y., and G. AlRegib, 2016, Weakly-supervised labeling of seismic volumes using reference exemplars: 2016 IEEE International Conference on Image Processing (ICIP), 4373–4377.

———, 2017, A weakly supervised approach to seismic structure labeling, *in* SEG Technical Program Expanded Abstracts 2017: Society of Exploration Geophysicists, 2158–2163.

Alfarraj, M., Y. Alaudah, and G. AlRegib, 2016, Content-adaptive non-parametric texture similarity measure: 2016 IEEE 18th International Workshop on Multimedia Signal Processing (MMSP), 1–6.

AlRegib, G., M. Deriche, Z. Long, H. Di, Z. Wang, Y. Alaudah, M. A. Shafiq, and M. Alfarraj, 2018, Subsurface structure analysis using computational interpretation and learning: A visual signal processing perspective: IEEE Signal Processing Magazine, **35**, 82–98.

Araya-Polo, M., T. Dahlke, C. Frogner, C. Zhang, T. Poggio, and D. Hohl, 2017, Automated fault detection without seismic processing: The Leading Edge, **36**, 208–214.

dGB Earth Sciences, B., 1987, The Netherlands Offshore, The North Sea, F3 Block - Complete: https://opendtect.org/osr/pmwiki.php/Main/Netherlands/OffshoreF3BlockComplete4GB.

Di, H., Z. Wang, and G. AlRegib, 2018, Seismic fault detection from post-stack amplitude by convolutional neural networks: Presented at the 80th EAGE Conference and Exhibition 2018.

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

Lin, T.-Y., P. Goyal, R. Girshick, K. He, and P. Dollár, 2017, Focal loss for dense object detection: arXiv preprint arXiv:1708.02002.

Long, J., E. Shelhamer, and T. Darrell, 2015, Fully convolutional networks for semantic segmentation: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 3431–3440.

Noh, H., S. Hong, and B. Han, 2015, Learning deconvolution network for semantic segmentation: Proceedings of the IEEE International Conference on Computer Vision, 1520–1528.

Russakovsky, O., J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, et al., 2015, Imagenet large scale visual recognition challenge: International Journal of Computer Vision, **115**, 211–252.

Rutherford Ildstad, C., and P. Bormann, 2017, MalenoV: Tool for training classifying SEGY seismic facies using deep neural networks: https://github.com/bolgebrygg/MalenoV.

Waldeland, A., and A. Solberg, 2017, Salt classification using deep learning: Presented at the 79th EAGE Conference and Exhibition 2017.