Insights on Domain Adaptation for fault identification

Mariana Lume, Marcelo Guarido, David Emery and Kris Innanen

ABSTRACT

Conventional deep learning has proved to be successful for fault identification tasks if the dataset selected for training and testing a model belong to the same domain. When it does not occur, the classifier usually degrades its performance. This study focuses on explaining the successes and failures of applying a model, previously trained with synthetic seismic images, on two different datasets of synthetic and real seismic images. When considering synthetic datasets, the classifier was very accurate; however, in the second scenario, a significant number of misinterpreted faults appeared. These outcomes are a direct consequence of the similitudes or discrepancies between both datasets used, hence Domain Adaptation techniques are usually applied to overcome the encountered challenges.

INTRODUCTION

Traditional deep learning approaches are those where a dataset is divided into a training and a test set, then a neural network learns from the training dataset and the pre-trained model is applied to the test set. These methods have been successful, but they are only valid under the assumption that both sets are drawn from the same probability distribution. The term domain comprises the feature space and the marginal probability distribution of the dataset. Hence, for usual learning approaches, the source (training data) and target (test data) domains and tasks are the same (Redko et al., 2019; Xu et al., 2020). Nevertheless, in some cases it is desirable to train a model with one or more source domains and then apply it to a target dataset with different but related distributions (Ben-David et al., 2010). In that sense, the performance of the classifier would degrade for the target task (Rozantsev et al., 2019), being necessary to find a learning paradigm that remains robust to a changing environment (Redko et al., 2019).

On the other hand, Deep Neural Networks, such as CNNs, learn representations from a vast amount of labeled data until the model achieve an acceptable performance. Particularly, DNN have brought improvements for fault detection in seismic volumes, allowing the recognition of complex structures in high-dimensional data. However, manually collecting and annotating faults is a laborious task and these labels are not always available (Xu et al., 2020; Zhou et al., 2021). Researchers, such as Sankaranarayanan et al. (2018), have solved the issue of missing labels through Domain Adaptation for a city objects semantic classification task. Other authors, such as Zhou et al. (2021), have additionally used similar Domain Adaptation techniques to address this topic in terms of seismic fault identification. Both studies focus on doing a pixel-wise prediction of the desired classes over a real dataset while training with features and labels of a synthetic dataset, which has related features but exhibits a different probability distribution. Subsequently, after using Domain Adversarial Neural Networks (DANN) architectures, a mapping function is learnt and domain-invariant features are found, allowing a good classification in the real dataset.

This paper focuses on introducing theoretical aspects of Transfer Learning and Domain Adaptation in order to explain the successes and failures when performing a fault identification task with a traditional deep learning approach in two scenarios: (1) considering two different synthetic datasets, as source and target domains and (2) using a synthetic dataset as the source domain and a real seismic dataset from a different geologic area as the target domain. Finally, a solution to address the encountered challenges in the second scenario is proposed according to the reviewed literature, aiming to apply it in future work.

TRANSFER LEARNING AND DOMAIN ADAPTATION

A method used to address the lack of labels of a dataset is Transfer Learning. This is a research topic of Machine Learning by which a new task is solved with the knowledge gained from an old task, improving generalization after finding transferable representations between both datasets that have different probability distributions. This technique is based on the real way humans learn, i.e., solving new tasks with the knowledge obtained from past experiences (Xu et al., 2020). The domain with available and enough annotated data is referred as source domain and the one with limited or none labels is referred as target domain. A type of an unsupervised scenario occurs when the source and target domains consist of synthetic labeled images and no annotated real images, respectively (Rozantsev et al., 2019).

Deep Domain Adaptation is a branch of Transfer Learning, commonly used in Computer Vision and in particular for image segmentation. Here, source and target domains are different but related and both tasks are the same. This technique aims to learn a mapping function that reduces the distribution discrepancy between both domains, thus the knowledge learned from the source domain is applied to the target domain. Representations are learnt with backpropagation from the source and target data for domain adaptation, finding domain-invariant schemes or sharing features from both datasets (Xu et al., 2020).

Differences between both domains can be caused by a covariate shift when the marginal distributions of source and target data change, but the predictive dependency is the same. Particularly, in image classification, a covariate shift occurs when objects of the same class, drawn from different domains (different datasets), are subject to high variability (image quality and representation), as can be observed in Figure 1; meaning that a classifier trained with one of these domains will fail on the other domain. Moreover, in image segmentation, differences can be related to a domain shift when the mapping function changes between two domains, leading to discrepancies in the marginal distributions of the observed samples due to variations in the measurement system or method of description (Redko et al., 2019). DANNs are typically introduced to address these problems; these are mainly based in variants of Generative Adversarial Networks (GANs), normally comprising: (1) a generator, which reconstructs fake images from an input space and (2) a discriminator, which distinguishes real from fake images. Both networks improve their abilities by alternatively training them, until the data distribution learned by the generator is close to the true data distribution (Xu et al., 2020).

CONVENTIONAL LEARNING WITH SYNTHETIC DATA

The Residual Neural Network (ResNet) model trained by Guarido et al. (2021) to perform a supervised fault identification task on 2D synthetic inline and crossline images of

FIG. 1. Each keyboard image comes from a different dataset and they differ drastically from one domain to another (Redko et al., 2019).

size 128x128 pixels (source domain), was applied for the same task on the inline Equinor synthetic model (target domain) provided in the FORCE 2020 Machine Learning Contest. This experiment was treated as an unsupervised scenario, although labels of the FORCE dataset were available. To achieve this classification, each section of the target dataset was cropped into small seismic data slices of size of 128x128 pixels to avoid altering the resolution of the original image. However, slices next to the boundaries were re-sized to the desired scale. Predictions were done over each of these slices after filtering them with a sharpening kernel with form:

$$\mathbf{K} = \begin{bmatrix} -1 & -1 & -1 \\ -1 & 9 & -1 \\ -1 & -1 & -1 \end{bmatrix} \tag{1}$$

Figure 2 shows that the source and target images are very similar in terms of quality, resolution, and frequency of the seismic events. Later, the full predicted 2D section was reconstructed, organizing and re-sizing back each prediction. Figure 3 illustrates some seismic sections with their corresponding predictions and labels. It is observed that this model classified with high accuracy the listric faults at the bottom of the images, as well as some extension of the normal faults in the upper part. Most of the missing faults, as well as the observed artifacts can be removed after applying a 2.5D prediction (Guarido et al., 2021). These results indicate that a traditional learning approach worked properly, meaning that after filtering the target images, the probability distribution of both datasets was very similar, allowing the appropriate application of the pre-trained model.

CONVENTIONAL LEARNING WITH SYNTHETIC AND REAL DATA

The same pre-trained ResNet model was applied for an image segmentation task on the unlabeled PSDM Poseidon seismic volume of the Browse Basin, Australia (target domain). Once more, each 2D section of the target dataset was cropped into small windows, as described previously, and the best results were obtained after computing their instantaneous amplitude attribute followed by the application of the sharpening filter **K**. The prediction was done over each modified window and full images were reconstructed. Figure 4 shows that the source and target images have significant differences in terms of quality, frequency of the events, and dipping and definition of the fault planes. Conversely to the previous scenario, this classification was not accurate since many artifacts were introduced and a large misidentification of faults occurred, especially in shallower depths, as can be observed in Figure 5.

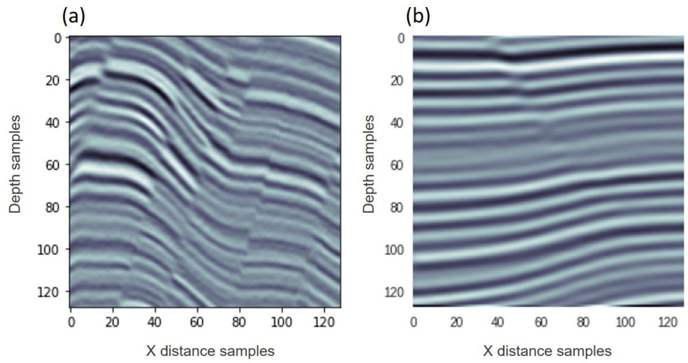

FIG. 2. Comparison between images from the (a) source dataset and (b) target dataset (FORCE dataset).

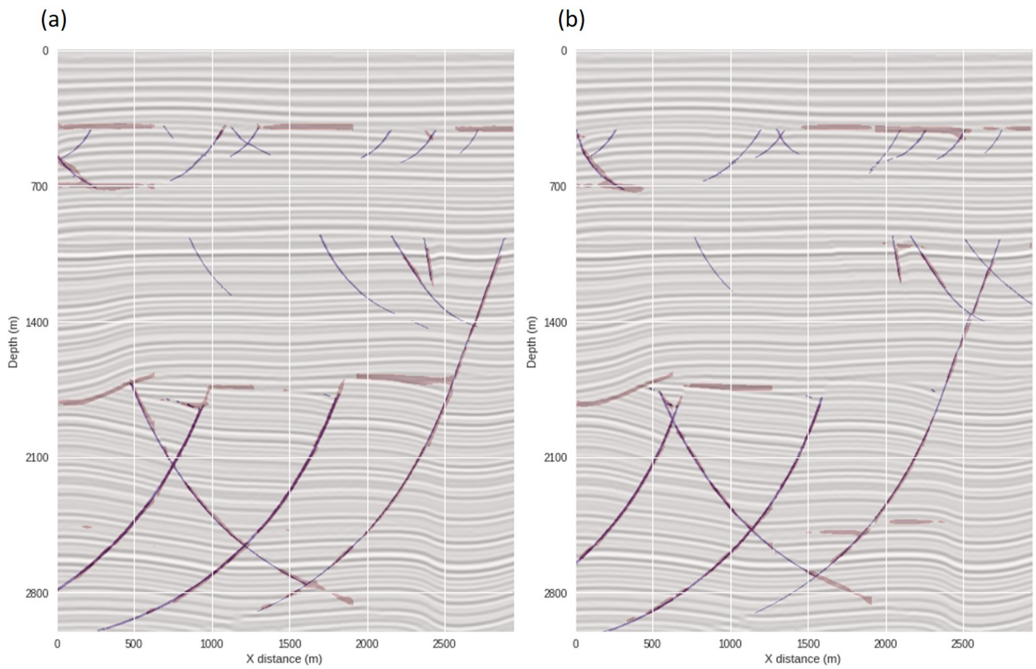

FIG. 3. Two different inline sections: Red lines are the predicted faults and blue lines are the provided labels.

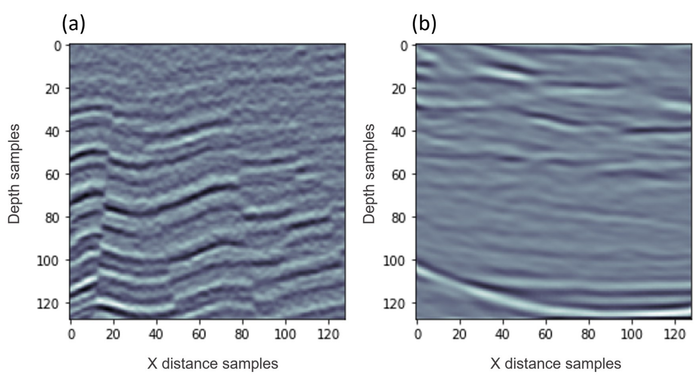

FIG. 4. Comparison between images from the (a) source dataset and (b) target dataset (Poseidon seismic volume).

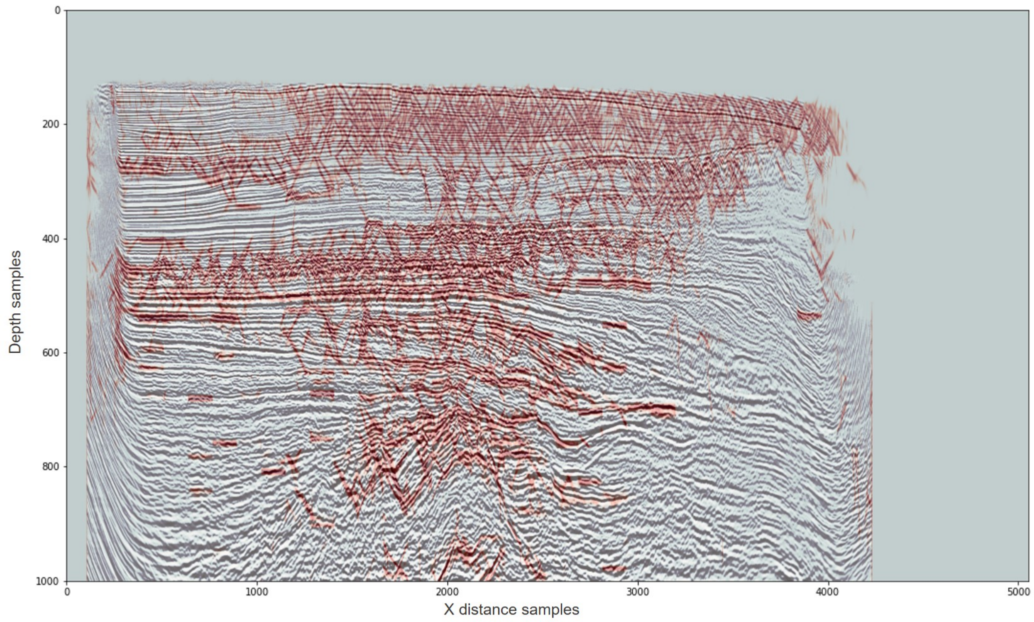

FIG. 5. Crossline section: Red lines are the predicted faults.

This time, no modification applied to the seismic slices made that the source and target domains were similar enough to perform a usual deep learning approach. As indicated by Zhou et al. (2021), although the source and target datasets share the same general geological structure law, a DNN trained with synthetic data can fail to produce a good fault identification in field data due to important differences on seismic signal frequency, seismic energy, seismic geometries (dip, azimuth), frequency of fault distribution, and noise level between both domains.

PROPOSED SOLUTION

The same ResNet architecture proposed by Guarido et al. (2021) could be adapted to a DANN to produce a more accurate fault classification in the Poseidon volume, while using the previous source dataset in the training process. Initially, the DANN network proposed by Zhou et al. (2021) will be reproduced to have a deeper understanding of how it works and later modify it according to this study. This architecture is mainly composed by: (1) a feature extractor, which generates features used as inputs for a fault and domain classifier; (2) a fault classifier, which predicts if the pixel is a fault or not; and (3) a domain classifier, which recognizes if the features come from the real or the synthetic dataset. With training, the domain classifier improves its abilities and the feature extractor finds the common-features confusing the domain classifier. Hence, it will not be able to recognize if the feature comes from real or synthetic data and the model will be sufficiently general to do a good performance on the real dataset.

Nevertheless, before training the model, feature engineering needs to be done to make the source and target datasets consistent on their aspect ratios, hence the dipping of the faults would be more similar. In addition, amplitude spectra of real and synthetic images should be studied to determine how large are these differences and apply additional modifications to the synthetic images if needed, in order to reduce these discrepancies and ease the extraction of the common-features. As suggested by Zhou et al. (2021), if the synthetic data have a higher frequency content than the real dataset, pre-processing the source images to reduce its frequency could be helpful.

CONCLUSIONS

The application of a conventional deep learning approach, based on a ResNet model trained with synthetic seismic images, proved to be successful for a fault identification task on a different but related synthetic dataset, as long as both sets of images are similar enough. However, testing this pre-trained model on a real dataset of a different geologic area yielded to a significant number of artifacts and misinterpreted faults due to large discrepancies in terms of frequencies, definition of fault planes, and quality between both datasets. Nevertheless, the future adjustment of this ResNet architecture to a different learning paradigm, such as Domain Adaptation, could help to overcome the issue of lack of labels in the real dataset and reduce differences of probability distributions between the synthetic and the real images by finding their common-features and allowing generalization.

ACKNOWLEDGMENTS

The authors would like to thank the sponsors of the CREWES project as well NSERC (Natural Science and Engineering Research Council of Canada) under the grant CRDPJ 543578-19 for making this work possible through their financial support.

REFERENCES

Ben-David, S., Blitzer, J., Crammer, K., Kulesza, A., Pereira, F., and Vaughan, J. W., 2010, A theory of learning from different domains: Mach Learn, **79**, 151–175.

Guarido, M., Wozniakowska, P., Emery, D. J., Lume, M., Trad, D. O., and Innanen, K. A., 2021, Fault detection in seismic volumes using a 2.5d residual neural networks approach: First International Meeting for Applied Geoscience Energy Expanded Abstracts, 1626–1629.

Redko, I., Morvant, E., Habrard, A., Sebban, M., and Bennani, Y., 2019, Advances in Domain Adaptation Theory: ISTE Press - Elsevier.

Rozantsev, A., Salzmann, M., and Fua, P., 2019, Beyond sharing weights for deep domain adaptation: IEEE Transactions on Pattern Analysis and Machine Intelligence, **41**, 801–814.

Sankaranarayanan, S., Balaji, Y., Jain, A., Lim, S. N., and Chellappa, R., 2018, Learning from synthetic data: Addressing domain shift for semantic segmentation: Presented at the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 3752–3761.

Xu, W., He, J., and Shu, Y., 2020, Transfer learning and deep domain adaptation, *in* Fernandez, M. A. A., Ed., Advances and Applications in Deep Learning: IntechOpen.

Zhou, ., Yao, X., Hu, G., and Yu, F., 2021, Learning from unlabelled real seismic data: Fault detection based on transfer learning: Geophysical Prospecting, **69**, 1218–1234.