Introduction

Automatic interpretation has long been an active topic of research and innovation for seismic exploration and reservoir characterization workflows. Currently, most of the practical implementations, especially in commercial software, often rely on choosing a set of attributes in advance and then training the classification algorithm to discriminate between different classes. However, recently, more elegant architectures of supervised deep learning (e.g. Convolutional Neural Networks or CNNs, Fully Connected Networks or FCNs) have been developed. Such architectures have the advantage of combining attribute extraction and classification in one go and they can achieve state-of-the-art performance for automatic interpretation (e.g. Waldeland et al., 2017). We review the application of Conditional Random Field (CRF) models on the segmentation maps obtained from the deep learning techniques. This approach promises to further improve the quality of the prediction for the deep learning models mentioned here.

Conventionally, computer vision tasks such as semantic segmentation are handled by probabilistic models such as the conditional random fields. The wide usage of CRF models in many state-of-the-art semantic segmentation pipelines is because of their ability in modelling the structural information. Despite CRFs successful application in natural and medical data, the application to seismic data, however, is limited. In this paper, we explain CRFs and how they can be incorporated into deep learning pipelines to make better predictions by acknowledging that we are predicting a structured output and explicitly modelling the problem to include our prior knowledge about the spatial image architecture.

Method

For a semantic segmentation task where the aim is to predict for each pixel the class it belongs to, the most traditional methods typically rely on a fully convolutional network (FCN) as displayed in Figure 1. A typical FCN architecture involves a down-sampling path achieved through pooling operations, which increases the receptive field size in order to capture more context. This down-sampling path is often accompanied by an increasing number of output channels to extract more attributes from the input images. However, despite gaining context, the resolution of the output maps is reduced and hence localization of the context is lost. For this reason, the down-sampling path is often followed by a symmetric up-sampling path achieved by transposed convolution operations which learn the localization of the context in low-resolution kernel maps. Sideway connections that preserve spatial

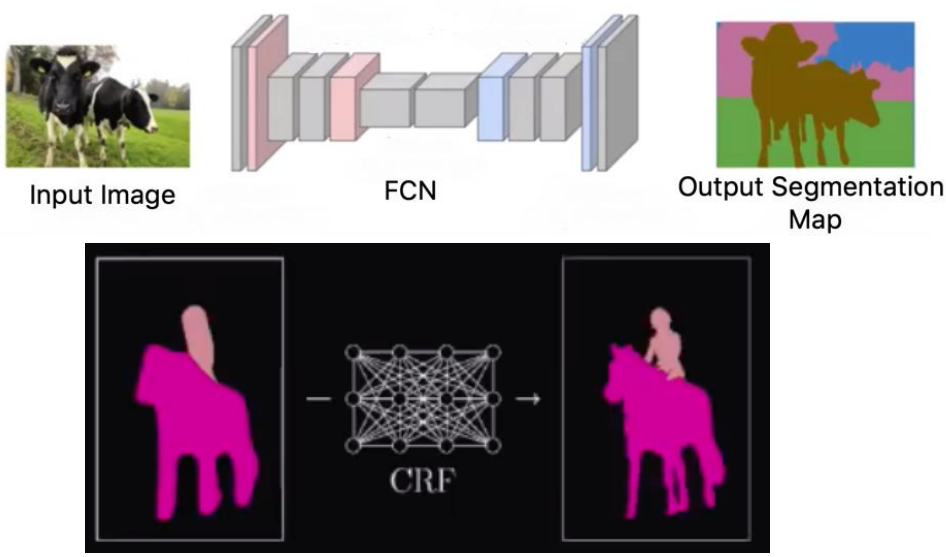

Figure 1 *Schematic display of FCN architecture that maps the input image directly to a segmented image of the same size. The bottom figure shows a CRF model applied to the segmented image to structure the output space (Luc et al., 2016).*

information are also present in the most state-of-the-art FCN architectures such as the U-Net architecture (Ronneberger et al., 2015). Given the success of U-Net architecture in recent semantic segmentation tasks, a similar model to that of Ronneberger et al. has been implemented in this paper.

Although deep learning models, including U-Net, have been proven to excel at many computer-vision tasks, one common drawback is that they predict label variables at each sample (or pixel) independently from each other. CRF models, on the other hand, belong to the family of probabilistic graphical models which are powerful tools for modelling the correlations between the samples being predicted. For this reason, combining the CRF models with deep learning models can result in better predictions by acknowledging that we are predicting a structured output and explicitly modelling the problem to include our prior knowledge about the spatial image architecture. In our pipeline, we consider fully connected CRFs with Gaussian edge potential (Krahenbuhl et al., 2012) as a conditioning tool for our FCN prediction as displayed in Figure 1. We make use of the highly efficient approximate inference algorithm proposed by Krahenbuhl et al. (2012) which is publicly available under MIT Licence on their GitHub repository (https://github.com/lucasb-eyer/pydensecrf, last accessed on Jan 2020). The CRF model we used, considers the Gibbs energy,

$$E(x) = \sum_i \psi_u(x_i) + \sum_{i<j} \psi_p(x_i, x_j), \qquad (1)$$

where ψ_u is the unary-potential term that describes the cost of assigning a certain label to a given pixel. This term is taken as the reciprocal of the class proportion. This is obtained from our FCN predicted segmentation maps which we feed into our CRF model. The term ψ_p is the pairwise gaussian potential term which forces nearby pixels to be in the same class to make the objects in our segmentation map more continuous. The pairwise potential has the form,

$$\psi_p(x_i, x_j) = \mu(x_i, x_j) k(\boldsymbol{f}_i, \boldsymbol{f}_j), \qquad (2)$$

where, $\mu(x_i, x_j)$ is the compatibility function given as,

$$\mu(x_i, x_j) = \begin{cases} 0 \ [x_i = x_j] \\ 1 \ [x_i \neq x_j] \end{cases}. \qquad (3)$$

f_i and f_j are position vectors for pixels i and j, and K is a gaussian kernel given by,

$$k(\boldsymbol{f}_i, \boldsymbol{f}_j) = w \exp\left(-\frac{|\boldsymbol{f}_i - \boldsymbol{f}_j|^2}{2\theta^2}\right), \qquad (4)$$

where w and θ are parameters that are set by the user. Then, the job of the inference algorithm is to find the configuration of the segmentation map that minimizes the Gibbs energy in equation 1. This way we penalize for the nearby pixels having different labels and thus make the object boundaries in our predictions smoother and more continuous as shown in Figure 2.

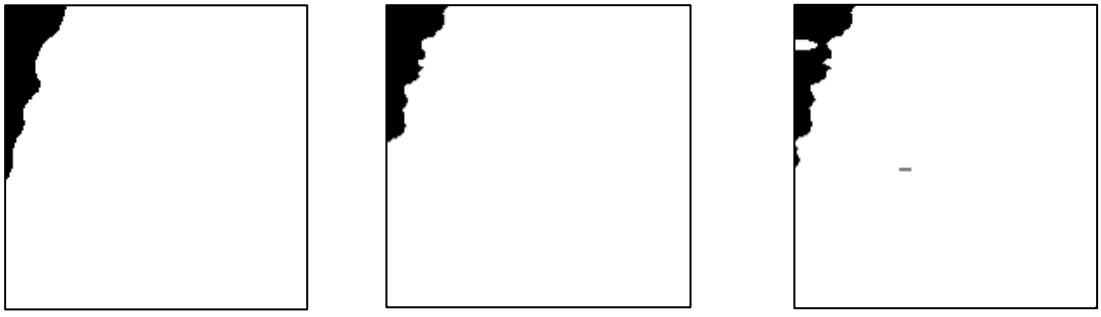

Figure 2 *Depth slice of ground truth segments obtained via manual interpretation of the seismic (left), CRF output (middle) and raw FCN output (right). It can be observed that CRF output matches the actual label better by removing the undesired holes and also improving the continuity.*

Real Data Example

We performed our experiments on a seismic cube located on the Norwegian Barents Sea, on the western part of the Bjarmeland Platform. The area covered by the seismic consists of the northward dipping normal fault striking in an east-west direction and nicely exhibits various geological sequences.

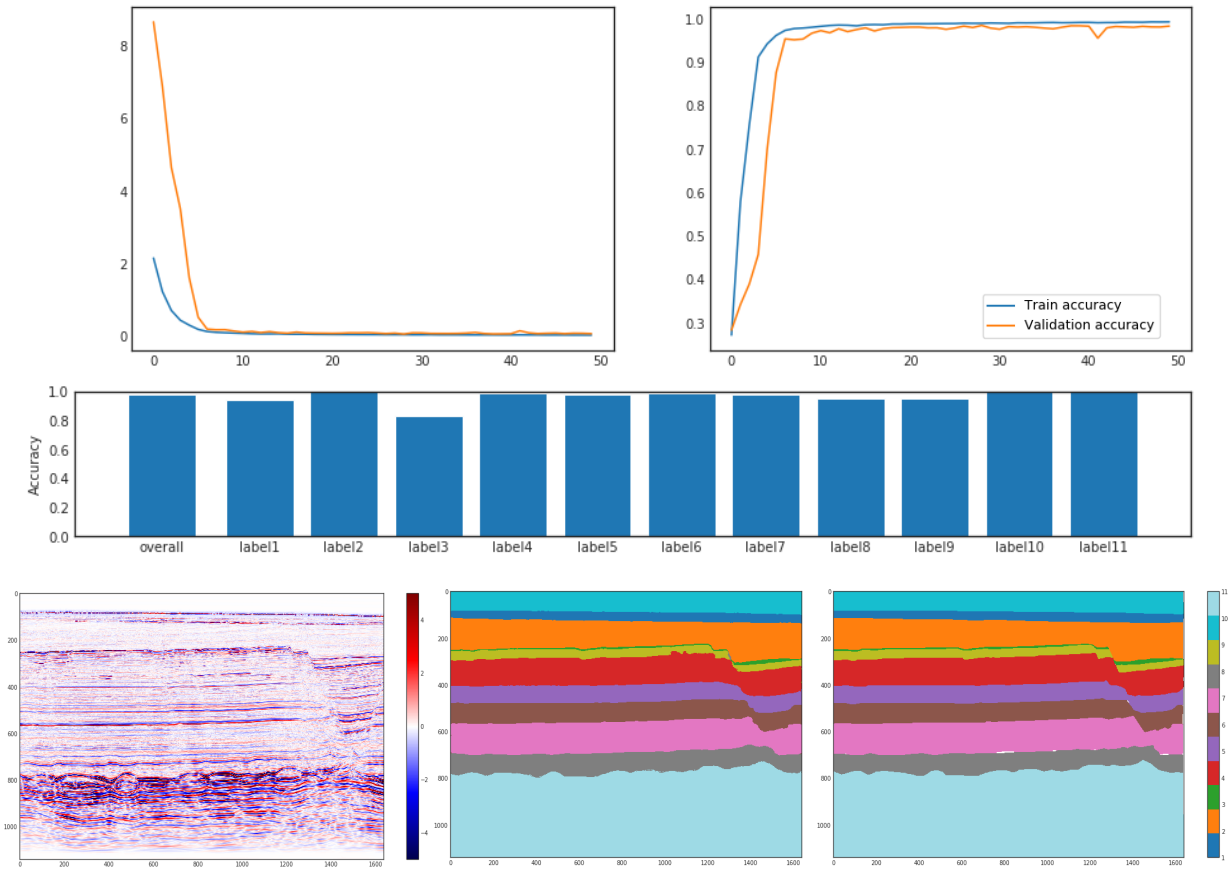

Figure 3 *The top figure shows the training and validation accuracy and loss for the U-Net model for 50 epochs of training. The second figure shows the accuracy of each class predicted for the test set where the overall pixel-wise accuracy is 0.969 (validation accuracy was 0.982). The third figure shows the seismic amplitudes for a test inline, the U-Net predicted labels, and the ground truth labels respectively.*

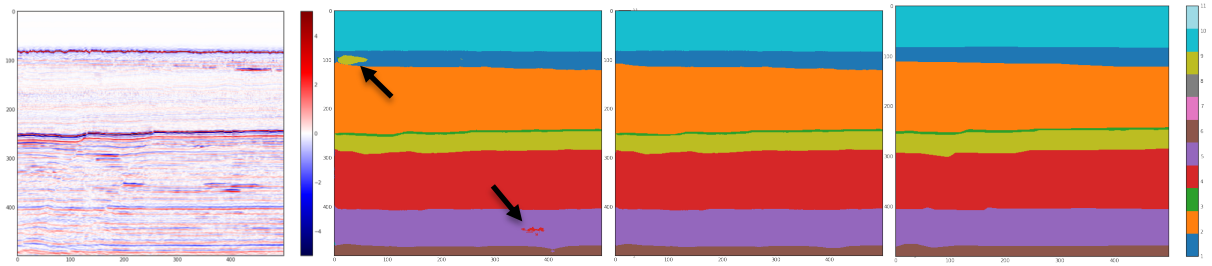

Figure 4 *The left figure is the seismic amplitudes from a blind inline slice. The second figure from the left is the segmentation map obtained from the FCN model (overall pixel-wise accuracy is 97.6%). The third figure from the left is the segmentation map after applying the CRF model (overall pixel-wise accuracy is 98.10%). The figure on the right is the ground truth labelling. Note CRF is applied in 3D across inline, crossline and finally depth directions. One can observe the improvements in removing the undesired patches within sequences as shown by the arrows.*

We trained our U-Net model using only six inline slices. We also use two inline sections as our validation set. Data augmentation, consisting of horizontal and vertical flips was applied to the training set to increase its size by four-fold. As the loss function, we use categorical cross-entropy and we weight the loss for every class proportionally to their reciprocal proportion in the training dataset. This was done since it was observed that classes with lower abundance (e.g. class 3) were being predicted much less accurately than classes with very high abundance (e.g. classes 2 and 11). We use the Adams algorithm as the optimiser algorithm which was observed to get better accuracies than the standard SGD algorithm. Then, we train the model for 50 epochs using 18 batches, each containing four images and their corresponding segmentation maps. The training curves and the predictions from the U-Net are displayed in Figure 3 where one can observe that the model is trained appropriately with good overall validation and test accuracy in the order of 0.98%.

Once the segmentation maps were obtained for the seismic cube by the U-Net approach, we applied the CRF model to the segmentation map in 3D (inline, crossline, depth directions) as shown in Figure 4. For our Gaussian kernel (equation 4), we set the values of θ to 10 and w to 7. These values were obtained by visually inspecting the output of the CRF model as displayed in Figures 2 and 4. It is observed that the CRF model improved the quality of the output segments significantly by removing some of the undesired patches and jitters and better conformance to the seismic events.

Conclusions

A U-Net model was trained using 6 randomly selected inline sections and their corresponding 11 labelled lithological units within a seismic cube from the Norwegian Barents Sea. Using data augmentation techniques, an overall accuracy of about %97 was achieved. However, it was observed that some of the segmented zones have undesired patches of other classes. In other cases, jittery behaviour was observed in the boundary between two classes. Our proposed CRF conditioning led to significant improvements in those areas. In practice, the proposed CRF conditioning can be applied on the fly where the user can input the prior geological knowledge and observe the outcome seamlessly.

References

Waldeland, A.U., Solberg, A.H.S.S. [2017]. Salt Classification Using Deep Learning. Conference: 79th EAGE Conference. 10.3997/2214- 4609.201700918.

Ronneberger, O., Fischer P., Brox, T., [2015]. U-Net: Convolutional Networks for Biomedical Image Segmentation, arXiv preprint arXiv:1505.04597**.**

Krahenbuhl P and Vladlen Koltun, [2012], Efficient Inference in Fully Connected CRFs with Gaussian edge potentials. In *NIPs,* pages 109-117.

Luc, P., Couprie C., Chintala S., and Verbeek, J., [2016], Semantic segmentation using adversarial networks. In NIPS Workshop on Adversarial Training.