Complete identification of complex salt-geometries from inaccurate migrated images using Deep Learning

Ana Paula O. Muller¹ 2 6, Jessé C. Costa³ 4, Clecio R. Bom¹ 5, Elisangela L. Faria¹, Matheus Klat¹, Gabriel Teixeira¹, Marcelo P. de Albuquerque¹, Marcio P. de Albuquerque¹

ABSTRACT

Delimiting salt inclusions from migrated images is a time-consuming activity that relies on highly human-curated analysis and is subject to interpretation errors or limitations of the methods available. We propose to use migrated images produced from an inaccurate velocity model (with a reasonable approximation of sediment velocity, but without salt inclusions) to predict the correct salt inclusions shape using a Convolutional Neural Network (CNN). Our approach relies on subsurface Common Image Gathers to focus the sediments’ reflections around the zero offset and to spread the energy of salt reflections over large offsets. Using synthetic data, we trained a U-Net to use common-offset subsurface images as input channels for the CNN and the correct salt-masks as network output. The network learned to predict the salt inclusions masks with high accuracy; moreover, it also performed well when applied to synthetic benchmark data sets that were not previously introduced. Our training process tuned the U-Net to successfully learn the shape of complex salt bodies from partially focused subsurface offset images.

¹Centro Brasileiro de Pesquisas Físicas (CBPF), Rua Xavier Sigaud, 150, Urca, Rio de Janeiro, Brasil.
²Petróleo Brasileiro S.A. (PETROBRAS), Edifício Senado, Av. Henrique Valadares, 28, Centro, Rio de Janeiro, Brazil.
³Universidade Federal do Pará, Belém, PA, Brasil.
⁴National Institute of Petroleum Geophysics (INCT-GP).
⁵Centro Federal de Educação Tecnológica Celso Suckow da Fonseca (CEFET-RJ), Av. Maracanã, 229, Maracanã, Rio de Janeiro, Brasil.
⁶E-mail: anapmuller@petrobras.com.br

INTRODUCTION

Velocity model building (VMB) is essential to make accurate subsurface images, especially in regions with high contrast velocities and complex structures. The conventional model building process uses seismic ray-based tomography (Bishop et al., 1985; Lambare et al., 2014) and Full Waveform Inversion (FWI) (Tarantola, 1984; Pratt, 1999) to determine the velocity model. Each iteration of Tomography or FWI requires a high level of regularization and a good input model to avoid local minima or cycle skipping. The result is often bounded to the maximal resolution power of the method, which often does not resolve the geological complexity of the area. One good example of the limitation of conventional inverse methods is the inclusion of salt bodies in the velocity model, a complex non-automated method that poses a high cost and is subject to uncertainties.
Salt inclusion is critical during the VMB. Whence a mistake in salt geometry makes the image below the salt unfocused or distorted (Dellinger et al., 2017), generating a wrong structure of the subsurface. Therefore, it can lead to economic consequences, especially in petroleum provinces where the reservoirs are below complex salt structures. Salt presents a great diversity of possible geometries (Hudec and Jackson, 2007) and has a high velocity contrast with the enclosing sediments. In complex areas, the definition of the salt geometry is estimated in an iterative process called salt flood (Mosher et al., 2007). Salt flood iterations require massive interpretation, geological knowledge of the sedimentary basin, and scenarios testing (Herron, 2013; Jones and Davison, 2014). Moreover, each iteration requires a lot of computational effort for migrating the data and observing the respective response of the interpreted salt geometry in the quality of the image generated. Some works try to overcome such difficulties focusing on FWI to correct the badly interpreted salt (Michell et al., 2017; Wang et al., 2019). However, it still requires data acquired with long offset, wide-azimuth, and low frequency, or the imposition of geological constraints over the method (Esser et al., 2016).

The problem of identifying the presence of salt in seismic images was previously studied using Deep Learning (DL) (Shi et al., 2019; Zhang et al., 2019). Nonetheless, this salt-segmentation is made over a well focused migrated image. Therefore, the focalization and accuracy of a seismic image are directly dependent on an accurate velocity model used in migration, where the model must include a reasonable estimation of the salt structures above the interpretation zone. Therefore, the available algorithms for salt-segmentation do not eliminate the salt flood steps during the VMB. These segmentation techniques have a high value for interpretation over seismic images and for refinement of salt structures. However, these structures must be somehow present in the velocity model used to migrate the seismic data.

Another process that attracts attention in applying DL is velocity estimation from seismic data (Araya et al., 2018). Many recent works are using DL aiming to make the complete prediction of velocity models, being particularly successful in predicting the salt-bodies embedded in such models, even those with very complex geometries (Araya-Polo et al., 2019; Yang and Ma, 2019; Li et al., 2020). These works use the raw seismic shot data to train the network to fully predict the correct velocity model that generated those seismic shots. Their results are very encouraging; however, they can not be easily extended to current seismic field data. The first reason is the size of real seismic acquisitions. DL needs to hold the training data in the device memory, and real seismic shots are much large than the most modern device could support, and even if they could be supported, the number of convolution operations would be prohibitive. Another difficulty is the irregularity of seismic shots, and the geometric non-correspondence of this data with the seismic images or the velocity models, i.e., source and receivers, varies their positions during the acquisition process. Concurrent to our work, Geng et al. (2022), have independently proposed a new approach that bypasses some problems mentioned here. Their work uses the common image gathers (CIG) in the angle domain and presents good results in recovering with DL structurally simple models from synthetic data sets.

Similarly, in this work, we propose to reduce the input data to the imaging domain to deal with more realistic seismic acquisition geometries, making a hybrid approach of the conventional VMB workflow and DL velocity estimation from seismic data. We use DL only in the most challenging step of VMB flow: the inclusion of salt structures. Instead of using
raw shots as input to the network, we migrated these shots with a reasonable sediment velocity model, without any salt inclusions. Our migration process uses an extended image condition, which generates a set of images with different subsurface offsets, which partially focus incorrectly migrated salt reflections. These images are the input channels for a DL algorithm to solve a segmentation problem, which has as outputs the salt masks. The trained U-Net returned accurate predictions over the test set and also over benchmarks salt models, indicating that this approach has robust properties when compared with traditional methods. It requires only kinetic information in the migrated domain, being less sensitive to the velocity model, coherent and incoherent noise, and salt rugosity. Moreover, when the salt mask predictions are not very accurate, according to our metrics, the estimated shape of the salt bodies are not drastically off, and still can help to improve interpretation.

RTM WITH EXTENDED IMAGE CONDITION

Reverse Time Migration (RTM) is one of the best algorithms for seismic imaging data in areas of high complexity (Etgen et al., 2009). It can handle high structural dip and velocity contrast, conditions common in salt basins. The inputs for RTM are the seismic data shots and the parameters models of the subsurface. For simplicity and to validate the idea, we will study only isotropic 2-D models in this work.

RTM uses the full wave equation to propagate the fields of the receivers \( W_r(x, t) \) (the registered seismic traces) and the field of the source \( W_s(x, t) \) into the earth using the full-wave equation with time running backward, where \( x = \{x, z\} \) for our 2-D case. This propagation is made individually for each shot, applying an image condition at each time step of the wave equation solution. In this work, we will use a cross-correlation extended image conditions (Sava and Vasconcelos, 2011), whose generalized form can be written as:

\[
R(x, \lambda, \tau) = \sum_{\text{shots}} W_s(x - \lambda, t - \tau) W_r(x + \lambda, t + \tau).
\]

(1)

Thus, the source and receivers fields are dislocated in extended image conditions to find correlations in the reflection point and in its surroundings. Correlations outside the reflection point indicate that the velocity model is incorrect. Equation (1) opens the possibility to vary the time and the space lag. For simplicity, we will focus only on horizontal spatial extension, thus keeping \( \tau = 0 \), and applying \( \lambda \) only over the \( x \) direction. \( \lambda \) is known as the subsurface offset, and it has one interesting property that differs from the well-known surface offset, which was decisive in choosing this representation in our study.

A gather presenting the subsurface offset, migrated with the correct velocity model, concentrates the reflections’ energy in \( \lambda = 0 \). The focus of energy in zero subsurface offset occurs because the source and receiver fields correlate better at the correct position when using the right model. To illustrate this property, in Figure 1(a), we show a gather migrated with the right velocity model. When the velocity used in RTM is not the correct one, the unfocused energy due to the error in the model spreads through the far offsets. In Figure 1(b), we can see a gather migrated with high velocity, and in Figure 1(c), the case with a low velocity. Few studies make use directly of the subsurface offset gathers (Frigério and Biondi, 2021); usually, the gather is converted to angle domain, and these angles gather continue the flow of VMB (Sava and Fomel, 2003).
Figure 1: Space-lag common-image-gather migrated with (a) correct velocity, (b) low velocity, and (c) high velocity.

Figure 2: Example of one model used to simulate the shots and the correspondent smoothed version without salt inclusion used for migration. The second line shows the effect of the lack of salt in the migration model along the common subsurface offset panels ($\lambda$), where the salt reflections dominates the large offsets due to the velocity error.
In this work, we proposed using the property of the spreading energy along $\lambda$ to detect the high contrast between the velocities of salt and sediment. Consider that we built reasonably well the sediment velocity using tomography, and we migrate the seismic data with this velocity without any salt body included in the sediment. In this case, majorly the salt energy will spread through the non-zero offset.

In Figure 2 we illustrate an example of inaccurately migrated salt. On the first line, we show the model used to generate synthetic shots and the model used in RTM migration. The model used in RTM is a smoothed version of sediment without salt inclusion. On the second line, we show the common subsurface offset panels, with $\lambda = 0$, $\lambda = 60$ m and $\lambda = 120$ m. As can be seen in Figure 2 the energy of the unfocused salt will dominate the high offsets. Of course, the shape of salt bodies will be completely distorted and compressed due to the error in the migration velocity, and these distortions vary along the $\lambda$. Indeed to infer the salt geometry using only the inaccurate migrated images presented in Figure 2 is a task which requires mapping the non-linear relations between input and output data. We will show in the following sections that a CNN can map the images over the subsurface offsets to the right salt-body shape.

DATA SET DEFINITION

We generated a set of 1505 synthetics, marine 2-D models, to validate the idea of using subsurface offset as input of a CNN to determine salt geometry. Models have an area of 3 km of depth and 5 km of lateral extension. The following sections describe the complete process of data set definition, from choosing the appropriated salt geometries, velocity model definition, modeling of the shots, and finally migration with the extended image condition.

Salt geometries and the velocity model

In order to get good responses from the network when presented to geological salt geometries, we used real models built with tomography and FWI in the past decade in regions with complex salt structures and high exploration interest. We extracted salt masks from these models as slices of the original model since our data set is composed of 2-D models. To achieve great varieties of structures we take slices with a minimal increment of 0.5 km in the original models. Furthermore, we distorted the proportion between vertical and horizontal sizes to fit the salt body in the size of our experiments (3 km x 5 km). With this strategy, we generated a total of 215 salt masks.

Besides, we applied distortions in these geological salt masks to increase the size of the original data set and the variability of geometries presented to the network. The distortions applied consist of randomly combined operations, like vertical and horizontal shifts, zoom-in or zoom-out, and rotations. First, we cut our mask in the top and bottom in 400 m to make the distortions, working over a 2.2 km of depth mask. These cuts were made to avoid salt extrusions and ensure that there will be at least one sediment layer below the bottom of the deepest salt in each model. Then, we included top and bottom layers of 400 m without salt to return the masks generated to the interest size. In the first line of Figure 3 we can see examples of the geological salt masks, and in the lines below, we see the masks generated by the distortion process.
There were generated six extra masks from the original one, making a total of 1505 salt masks to define the training, validation, and test set. With this process, we hope to increase our network's ability to generalize to any salt geometry. We included each salt mask in a different sediment velocity model. The sediment model is randomly generated, but obeying some rules. The first layer has a velocity of 1500 m/s and the last one has a velocity of 3500 m/s. Each model has eleven velocity layers with random shapes and depths. The salt body is then inserted with a velocity of 4500 m/s.

Modeling seismic acquisition

For each model, we simulated a synthetic seismic acquisition, using a finite-difference wave propagator, isotropic, acoustic, with second-order in time and eight-order in space (with spatial sampling equals to 10 m), and CPML boundaries (Martin et al. 2008 Ko- matitsch and Martin 2007). To source simulation, we used a Ricker wavelet with a peak frequency of 20Hz. The shots are recorded by 601 receivers with a distance of 10 m between them, in a split spread acquisition with the maximum offset relative to the source of 3 km. There were recorded 125 shots, the first shot occurs at the beginning of the interest region and the last shot at the end, with 40 m of interval between shots. Receivers record at a rate of 4 ms for 3.5 s. We applied the mute of direct wave in shots for migration. The models were laterally extrapolated to allow the full shot covering inside the interest area.
Migrated images for training

As mentioned in the previous sections, we migrated our shots with an RTM cross-correlation extended image condition defined in equation 1. The subsurface offset $\lambda$ varies between zero, which is equivalent to the non-extended condition, and 120 m, with an increment of 20 m.

Our main objective in introducing the use of RTM was to preserve the ability that CNN has to predict the salt geometries, but with the capacity to reduce the amount of information to input in the network. However, when we proposed to migrate the shots with RTM, it is necessary to input also a velocity model in depth. Thus we gain in reducing the size of inputs, but we lose with the requirement of such model. Considering that the conventional salt flood process occurs after the initial definition of the sediment model, it is also reasonable to include our method in this step of VMB flow. The sediment model for a salt flood can be a low-frequency model, like the ones obtained in initial interactions of tomography and does not need to honor all the details of the layers. To emulate a sediment velocity model in the early stages of a tomography flow, we smoothed the slowness with a Gaussian filter with a window of 12 samples. Smoothing in the slowness preserves most reflections’ travel time but leaves some errors in the fine-scale.

DEEP LEARNING PIPELINE FOR SALT MASK PREDICTION

The presented Deep Learning pipeline for salt mask prediction is based on an U-Net architecture. The U-Net (Ronneberger et al., 2015) is a type of fully convolutional network (Shelhamer et al., 2016) model widely used in tasks related to semantic segmentation (e.g., Badrinarayanan et al., 2017). The model’s encoder-decoder structure is ideal for tasks where network inputs and outputs must have the same spatial dimensions. At least a couple of DL applications in seismic analysis problems have used the U-Net topology (e.g., Alfarhan et al., 2020; Wang and Nealon, 2019), despite not being strictly related to semantic segmentation tasks. This is possible since these problems can often be formulated with 2D, or 3D, inputs and outputs of equal spatial dimensions, which works well with networks of this type.

Here we pose the problem of mapping salt inclusions from migrated sections as one of binary mask prediction. The U-Net receives as input a [256, 512, $N_\lambda$] volume, where $N_\lambda$ is the number of subsurface offsets panels, and produces as output a [256,512,1] binary section of the estimated salt mask. With respect to the original design, we have made only two changes to the architecture: 1) our convolutional block, in both encoder and decoder sections, include a batch normalization layer between each pair of convolutional and activation layers, which helps to stabilize the gradient flow through the network (Ioffe and Szegedy, 2015); 2) our network starts with 16 feature maps in the first stage of the encoder, which significantly reduces the number of network parameters to be trained.

The four stages encoder section of the network gradually downsamples the data at each stage and learns to extract relevant information to the task at different resolution scales of the data. Downsampling occurs only in the first three encoder stages, where it is performed by max-pooling layers after processing through the convolutional block. The decoder section gradually upsamples the data back to the original spatial dimensions. It also has four stages. At each stage the data is first upsampled with a transpose convolutional layer, then concatenated with the feature map coming from the skip connection of the corresponding
Figure 4: Schematic of our flow. We divided the flow into four foremost steps 1) the generation of data set training, 2) migration of shots to reduce the information and make the data domains compatible for the network, 3) network training, and finally 4) the salt mask prediction over unseen data.
encoding stage, before going through the convolutional block. After the last upsampling stage in the decoder, a final convolutional layer is applied to the data to reduce the number of channels to one, so it matches the single channel of the salt mask labels. Then a sigmoid activation function is computed over this single panel ([256, 512, 1]), which represent the network’s predictions.

In Figure 4, we summarized our flow from the data generation process to the predictions over unseen masks. Step 1 is usually adopted in the literature of VMB with DL, which consists in the generation of synthetic labeled data. The main novelties in our flow are in steps 2), where we changed the input data representation to migrated subsurface offsets panels, and in steps 3) and 4), with training and prediction of the salt mask over this new input data set. It is essential to clarify that the prediction step does not require any previous salt geometry information. The RTM CIGs reduced the original seismic data size by sixty times, bringing the network input to a more affordable size, even for realistic field data sets. Moreover, the network input and output are in the same domain and have the same size, which facilitating the DL approach.

**Training methodology**

We trained our network with $N_\lambda = 7$, i.e. with seven subsurface offsets panels, which varies from $\lambda = 0$ m to $\lambda = 120$ m with 20 m of increment. These panels work like channels of the input image for the CNN. The original size of salt masks and migrated panels is 300 samples in depth and 500 samples laterally. In order to fit the input and output images to the proposed U-Net design, we cut the images in depth to 256 and pad 6 samples at each lateral to have images with 256x512 samples. In addition, since the models have a minimal water layer, we cut the top portion of the images. We also clipped the amplitude of each subsurface offset panel in percentile 96 and then normalized the amplitude of each panel between 0 and 1.

We used the 1505 synthetic models for our study previously described. This data is split into three distinct sets: training, validation, and test. The training set is the one that contributes to updating the network’s weights. The validation set is used to control the metrics and loss during the training process to avoid overfitting. Finally, the test set is used to make the network predictions after the training process. We split our original set in twenty percent for the test set, and for a robust evaluation of network performance, we used $k$-fold cross-validation in training [Moreno-Torres et al., 2012]. We shuffled our validation set in five-folds and then trained the network ten times for each fold using random initial weights. Our training and validation loss always shows the statistics over the fifty training processes. Such strategies account for fluctuations inherent to the random nature of the process.

One crucial aspect to be considered when training a CNN to solve a segmentation problem is to define which loss function will be used. The loss is the inaccuracy measure of predictions made by the network and guides the update of the weights through an optimization algorithm. We used the Jaccard loss, especially because of its capability to address class imbalances (Duque-Arias et al., 2021).

Figure 5 shows the learning curve corresponding to the network performance and calculated by using the Jaccard loss. The learning curve calculated from the training data
Deep learning salt detection

Figure 5: Jaccard loss measured along the training process for the training and validation sets.

set presents how well the network is learning, while the one calculated from the validation data set indicates the generalization capacity of the network. By comparing both curves, it is possible qualitatively infer that the general behavior of the training curves shows an accurate fit, with a slight fractional difference between validation loss and training loss. The validation loss decreases until a point of stability, situated around 50 epochs, showing a loss median below 0.02. This suggests that, despite of the complexity of the problem addressed, the trained network generalizes well.

RESULTS

Images predicted over the original test set

Initially, we evaluated the results over the test set. The test set is composed of models from the original set that were not seen by the network during training. The results are presented in Figures 6 and 7, which shows the real salt mask, the predicted salt mask, and the difference between predicted and real masks, where red indicates false-negative and blue false-positive. We separated our results into two figures in order to easily discuss the good and bad examples.

In Figure 6 we captured models with complex salt structures, such as salt teardrops, overhangs, weld, and tongues, showing that the network can make surprisingly good predictions even for the more challenging cases. One astonishing prevision is the correct shape of allochthonous salt, especially the base of such bodies, and the correct separation from the salt below such structures. The predictions could capture even thin structures, which suggests that this approach can also be helpful to track basalt dikes inside the sediment. The predictions errors made in those masks could be easily corrected during the FWI process or with interpretation over the stack session migrated with a velocity model including these predicted salt geometries.

Although, in Figure 7 we can see the cases where the network made more mistakes during prediction. Below some overhangs, particularly the thick ones, is visible challenging to track the thin salt structures. This probably is due to the lack of illumination necessary to devise such complex structures. We can also note border effects in the bottom and laterals,
Figure 6: Examples of masks from the test set where the network made good predictions. Blue indicates false positive and red false negative in error figures.

Figure 7: Examples of masks from the test set where the network made extensive mistakes. We can see that they are associated with regions with low illumination or the geology is too complex to our simple cross-correlation image condition. Blue indicates false positive and red false negative in error figures.
Deep learning salt detection

Figure 8: The velocity model for shot simulation including the salt geometry of a cut in the left portion of the BP/SEG model, and the smoothed sediment model used in migration to generate the input for network predictions.

Figure 9: Network prediction made over the BP/SEG inspired data. Blue indicates false positive and red false negative in error figures.

probably because the energy needed to visualize those structures spread outside the image domains. Another point of attention is the inaccuracy in structures containing salt-tongues or diapirs that are upside down due to the distortion process applied over the geological structures. Our algorithm usually lacks accuracy in these regions, predicting smaller salt than the correct one. Such behavior can be explained by the fact that the bottom of a salt body, when migrated with a smaller velocity, that is our case, collapses in a small region than the actual one. Where the base of salt is a thin structure, probably these collapsed regions are smaller than the network can detect as indicative of salt.

Performance evaluation on benchmarks

To check for the ability of generalization of our trained network, we used two well-known geophysics benchmarks to generate input data sets to test our trained network. The benchmarks data sets were not used during the training/validation process. The first one was a rectangular cutout BP/SEG velocity model [Billette and Brandsberg-Dahl, 2005]. We extracted only the salt-mask, re-scaled, and inserted them inside a sediment model as shown in Figure 8. The sediment model was defined by the same algorithm we used to create our original sediment models. Remember that shape and velocities of interfaces are randomly defined but are restricted to a limited range of possibilities. Therefore, using this data, we want to check the generalization of salt geometries, not yet the generality for any sediment velocity. The predicted salt mask for BP/SEG adapted model is shown in
Figure 10: Example of one slice of SEG 3D model used for shot generation and the respectively model used for migration. One remarkable aspect of these figures is that we do not change sediment’s velocity range, which is entirely different from the velocity of our training data set. Only salt velocity was modified to 4500 m/s.

Figure 11: Network predictions made over inputs from SEG model. Blue indicates false positive and red false negative in error figures.

Figure 9, where it is possible to see that the network predicted the salt mask with high accuracy, restricting the mistakes to the past discussed regions, were there probably a lack of illumination to precisely devise the reflections.

Another test was performed over 2D velocity models extracted from the 3D SEG salt model (Aminzadeh et al., 2017). In this case, as the SEG model has dimensions more compatible with the ones we used in our work, we only cut and reshaped the model to fit in our interest area, leaving the sediment layers untouched. The only velocity manipulation we performed was to set the salt velocity to 4500 m/s, which in the original SEG model is equal to 4482 m/s. This was made because we assume that the network uses the salt reflections displacements to infer the salt geometry, so it should be necessary to have the same salt velocity to observe the same displacements due to the lack of salt in the model. In Figure 10 we show one slice of SEG 3D model used in this work, with the model for shot simulation and the correspondent model used in RTM. The velocity model used in RTM is a smoothed trace of sediment velocity, which we extended to fulfill our interest area. For this study, we extracted 24 slices from the original 3D model.

Figure 11 presents two examples of predictions made over SEG data set, where it is possible to see that the network generally made accurate previsions of the salt body location. Here we want to emphasize two differences from the previous predictions we made. First,
sediment velocity for the shot generation is entirely different from the one used to generate our training data set. Moreover, the sediment velocity used in migration is slightly worse than the previous case since we used a plane-parallel model in migration due to the lack of an isolated sediment model. The previously mentioned differences in data generated for predictions presented in Figure 11 are a strong indication that our trained network is robust to generalize the predictions to other interest areas.

**Does the number of subsurface offsets matter?**

In the initial input data configuration proposed in this work, we use some geophysical premises, such as the sampling of subsurface offset. When we build a velocity model using tomography, it is always desirable to have the offset information as denser as possible. Such high density avoids alias effects and makes the assignment of measuring velocity errors easier to be made. In our problem, we do not have to deal with alias effects because tracking of move-out information is not done. Nevertheless, if we consider the information amount, with a large number of offsets, we expect to be giving more information about the salt velocity errors to the network. Thus there are important questions to be answered. For example, which is the minimal offset density necessary for the network to make the correct predictions? Can we ignore kinematics effects and train the network using only one offset?

To verify these points, we changed the network input and measured the performance of the loss during the training process. We defined five different configurations for the input data, as listed below:

- **Dense**: $N_\lambda = 7$, with $\lambda = 0,20,40,60,80,100,120$ m
- **Sparse**: $N_\lambda = 3$, with $\lambda = 0,60,120$ m
- **Near**: $N_\lambda = 1$, with $\lambda = 0$ m
- **Mid**: $N_\lambda = 1$, with $\lambda = 60$ m
- **Far**: $N_\lambda = 1$, with $\lambda = 120$ m

In Figure 12 we show the performance of validation loss using the different inputs. The smallest loss is achieved when using the Dense input; the Sparse input has the second-best performance. This result suggests that the network uses all the available information to make the best predictions. Figure 12 also shows that using only one input offset achieves the worst performance, with almost the same performance for Near, Mid or Far offset. Such behavior suggests that the network accounts for the kinematics effects to make the salt predictions, being necessary to track the salt reflections positions over the offsets to detect the salt body shape.

The values of loss validation curves are separated by more than one standard deviation, but they have close values. In order to confirm the best way to define the input data, we measured the Dice Coefficient (DC) for each predicted model, 301 in the test set and 24 models in SEG set. Figure 13 shows the median and one standard deviation as error bars. DC for the test set and SEG models confirms what is observed in the validation loss curve, presenting the best prediction accuracy when using the Dense input. Thus, the
Figure 12: Effect of the number of subsurface offset ($N_\lambda$) in the input data over the median validation loss during fifty training processes.

Figure 13: Median Dice Coefficient calculated over each model in Test Set and in SEG models for the different inputs configurations proposed.
results degrade when the number of offsets decreases, reaching the worst results for the Far offset in the test set. Therefore, Figure 13 indicates that our U-Net needs the kinematic information of salt reflections to make accurate predictions. The worst result with Far offset, when we use only one, can be explained by the fact that the top of salt is approximately corrected imaged in the Near offset, due to the correct sediment velocity, and can guide the interpretation of the top of the salt body.

The effect of the input subsurface offset is more evident when we analyzed DC for SEG models. Figure 13 shows that the only input configuration that generates reasonable values of DC is the Dense input. With the others inputs, the network is not able to generalize to the SEG model.

**DISCUSSION**

Migrating the shots has two evident improvements over the known strategies of using the shots directly: reduces/regularizes information and brings the network input to the same domain of network output. With these improvements, applying DL methods to predict models features from real seismic data becomes probably a possible task, requiring additional studies to deal with noise, multiples, anisotropy, and other problems of real seismic data. As we converted the input data to the domain of the output data, this problem can be easily extended in size or dimensions. Moreover, we believe that an extension to 3D models can better handle the illumination problems that we have in some predictions because the structures could be illuminated in different directions.

Besides, migration introduced a generalization ability in the network, as this eliminates the effect of the background velocity; only the salt generates the energy that matters for predictions. We show with SEG models that we can use a trained network to predict salt in unseen sediment velocities, as long as the salt velocity is the same and the salt geometry in the training set is sufficiently diverse. The results with SEG models suggest that a trained network could be used for different geologies. Such generalization is relevant because the most expensive step is generating synthetic data to train the network. Once we have the trained network, predictions are made almost instantly.

When applying an imaging method to reduce our input information and focus only on the desirable feature (salt geometry), we introduced some well-know imaging problems in predictions, such as illumination, band-limited resolution, and border effects. These problems can be addressed in future investigations of our approach. For example, using seismic data with wide frequency-band, image condition varying the $\lambda$ horizontally and vertically, time-shift extended image condition, and processing improvements to CIGs.

**CONCLUSIONS**

This work presented an alternative approach to the conventional salt flood methodology, which consists of training a DL network (U-Net) to locate salt inclusions over incorrectly migrated subsurface offset panels. In the migrated images, salt events are entirely distorted from the original geometry due to the lack of salt velocity in the migration model. Here we showed that a U-Net could successfully learn to identify, track, and surprisingly make the non-linear transforms necessary to determine the right salt location. Despite our
predicted salt bodies having minor inaccuracies at this development point, they present a high correlation with the actual structures. Thus, even if they can not be used as a final interpretation of salt, they can be an excellent initial input in VMB workflow. Either for improving the salt body interpretation without the extensive testing of salt scenarios or as a better approximation of salt inclusions for FWI iterations.

ACKNOWLEDGEMENTS

APOM thanks Petrobras for sponsoring her postdoctoral research and for the permission to publish this work. CRB acknowledges the financial support from CNPq (316072/2021-4). The authors acknowledge the LITCOMP/COTEC/CBPF multi-GPU development team for all the support in the Artificial Intelligence infrastructure and Sci-Mind’s High-Performance multi-GPU system. The authors also would like to thank the SENAI CIMATEC Supercomputing Center for Industrial Innovation, for the cooperation, supply and operation of computing facilities."

REFERENCES

Alfarhan, M., A. Maalej, and M. Deriche, 2020, Concurrent detection of salt domes and faults using resnet with u-net: 2020 6th Conference on Data Science and Machine Learning Applications (CDMA), 118–122.
Aminzadeh, F., N. Burkhard, J. Longa, T. Kunz, and P. Duclos, 2017, Three dimensional seg/eae models — an update: The Leading Edge, 15, 109–156.
Araya, M., J. Jennings, A. Adler, and T. Dahlke, 2018, Deep-learning tomography: The Leading Edge, 37, 58–66.
Araya-Polo, M., S. Farris, and M. Florez, 2019, Deep learning-driven velocity model building workflow: The Leading Edge, 38, 872a1–872a9.
Badrinarayanan, V., A. Kendall, and R. Cipolla, 2017, Segnet: A deep convolutional encoder-decoder architecture for image segmentation: IEEE Transactions on Neural Networks and Learning Systems, 39, 2481 – 2495.
Billette, F., and S. Brandsberg-Dahl, 2005, The 2004 bp velocity benchmark: Presented at the 67th EAGE Conference & Exhibition, European Association of Geoscientists & Engineers.
Bishop, T. N., K. P. Bube, R. T. Cutler, R. T. Langan, P. L. Love, J. R. Resnick, R. T. Shuey, D. A. Spindler, and H. W. Wyld, 1985, Tomographic determination of velocity and depth in laterally varying media: Geophysics, 50, 903–923.
Dellinger, J., A. J. Brenders, J. R. Sandschaper, C. Regone, J. Etgen, I. Ahmed, and K. J. Lee, 2017, The garden banks model experience: The Leading Edge, 36, 151–158.
Duque-Arias, D., S. Velasco-Forero, J.-E. Deschaud, F. Goulette, A. Serna, E. Decenci`ere, and B. Marcotegui, 2021, On power jaccard losses for semantic segmentation: Presented at the VISAPP 2021: 16th International Conference on Computer Vision Theory and Applications.
Esser, E., L. Guasch, F. Herrmann, and M. Warner, 2016, Constrained waveform inversion for automatic salt flooding: The Leading Edge, 35, 235–239.
Etgen, J., S. H. Gray, and Y. Zhang, 2009, An overview of depth imaging in exploration geophysics: Geophysics, 74, no. 6, WCA5–WCA17.
Frigério, O., J., and B. Biondi, 2021, Tomographic waveform inversion (twi): Presented at the IMAGE - International Meeting for Applied Geoscience & Energy, SEG —AAPG.

Geng, Z., Z. Zhao, Y. Shi, X. Wu, S. Fomel, and M. Sen, 2022, Deep learning for velocity model building with common-image gather volumes: Geophysical Journal International, 228, 1054–1070.

Herron, D. A., 2013, Thoughts and observations on interpreting depth-imaged data: Interpretation, 1, B1–B6.

Hudec, M. R., and M. P. Jackson, 2007, Terra infirma: Understanding salt tectonics: Earth-Science Reviews, 82, 1–28.

Ioffe, S., and C. Szegedy, 2015, Batch normalization: Accelerating deep network training by reducing internal covariate shift: ArXiv, abs/1502.03167.

Jones, I. F., and I. Davison, 2014, Seismic imaging in and around salt bodies: Problems and pitfalls: SEG Technical Program Expanded Abstracts, 3684–3688.

Komatitsch, D., and R. Martin, 2007, An unsplit convolutional perfectly matched layer improved at grazing incidence for the seismic wave equation: Geophysics, 72, no. 5, SM155–SM167.

Lambare, G., P. Guillaume, and J. Montel, 2014, Recent advances in ray-based tomography: European Association of Geoscientists & Engineers, 2014, 1–5.

Li, S., B. Liu, Y. Ren, Y. Chen, S. Yang, Y. Wang, and P. Jiang, 2020, Deep-learning inversion of seismic data: IEEE Transactions on Geoscience and Remote Sensing, 58, 2135–2149.

Martin, R., D. Komatitsch, and A. Ezziani, 2008, An unsplit convolutional perfectly matched layer improved at grazing incidence for seismic wave propagation in poroelastic media: Geophysics, 73, no. 4, T51–T61.

Michell, S., X. Shen, A. Brenders, J. Dellinger, I. Ahmed, and K. Fu, 2017, Automatic velocity model building with complex salt: Can computers finally do an interpreter’s job?: SEG Technical Program Expanded Abstracts, 5250–5254.

Moreno-Torres, J. G., J. A. Sáez, and F. Herrera, 2012, Study on the impact of partition-induced dataset shift on k-fold cross-validation: IEEE Transactions on Neural Networks and Learning Systems, 23, 1304–1312.

Mosher, C., E. Keskula, J. Malloy, R. Keys, H. Zhang, and S. Jin, 2007, Iterative imaging for subsalt interpretation and model building: The Leading Edge, 26, 1361–1488.

Pratt, R. G., 1999, Seismic waveform inversion in the frequency domain — part 1: Theory and verification in a physical scale model: Geophysics, 64, 888–901.

Ronneberger, O., P. Fischer, and T. Brox, 2015, U-net: Convolutional networks for biomedical image segmentation: Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015, 234–241.

Sava, P., and I. Vasconcelos, 2011, Extended imaging conditions for wave-equation migration: Geophysical Prospecting, 59, 35–55.

Sava, P. C., and S. Fomel, 2003, Angle-domain common-image gathers by wavefield continuation methods: Geophysics, 68, 1065–1074.

Shelhamer, E., J. Long, and T. Darrell, 2016, Fully convolutional networks for semantic segmentation: IEEE Transactions on Pattern Analysis and Machine Intelligence, 39, 1–1.

Shi, Y., X. Wu, and S. Fomel, 2019, Saltseg: Automatic 3d salt segmentation using a deep convolutional neural network: Interpretation, 7, 1A–T725.

Tarantola, A., 1984, Inversion of seismic reflection data in the acoustic approximation: Geophysics, 49, 1259–1266.
Wang, E., and J. Nealon, 2019, Applying machine learning to 3d seismic image denoising and enhancement: Interpretation, 7, 1–34.
Wang, P., Z. Zhang, J. Mei, F. Lin, and R. Huang, 2019, Full-waveform inversion for salt: A coming of age: The Leading Edge, 38, 204–213.
Yang, F., and J. Ma, 2019, Deep-learning inversion: A next-generation seismic velocity model building method: Geophysics, 84, no. 4, R583–R599.
Zhang, L., M. Zhang, Z. Zhong, T. Zhao, Y. Wu, J. Wei, and C. Zhan, 2019, Deep learning approach in characterizing salt body on seismic images: SEG Technical Program Expanded Abstracts, 2594–2598.