XLIII CILAMCE
21-25, November, 2022
FOZ DO IGUAÇU - PR - BRAZIL

Automatic segmentation of breakouts in acoustic borehole image logs using convolutional neural networks

Gabrielle B. dos Anjos[1], Marcelo Gattass[1], Augusto I. Cunha[1], Candida M. de Jesus[2], Luiz F. Santos[1], Mayara Gomes[1], Nelia Reis[1], Raquel Guilhon[1], Renata Nascimento[1]

[1]*Tecgraf Institute, Pontifical Catholic University of Rio de Janeiro - PUC-Rio*
gabriellebrand@tecgraf.puc-rio.br, mgattass@tecgraf.puc-rio.br, acunha@tecgraf.puc-rio.br, mayaragomes@tecgraf.puc-rio.br, neliareis@tecgraf.puc-rio.br, raquelg@tecgraf.puc-rio.br, rlins@tecgraf.puc-rio.br, lsantos@tecgraf.puc-rio.br
Prédio Pe. Laércio Dias de Moura - R. Marquês de São Vicente, 22451-900, Gávea, Rio de Janeiro - RJ, Brasil
[2]*Petrobras*
Av. República do Chile, 330, 15th Floor, 20031-170, Centro, Rio de Janeiro - RJ, Brasil
candida.jesus@petrobras.com.br

Abstract. Breakouts are collapsed zones on borehole walls caused by compressive failure. The identification of breakouts in wellbores is fundamental for estimating the stability of the well and to obtain the magnitude of the maximum horizontal stress present in the rock formation. Traditionally, professional interpreters identify and characterize breakouts manually, which can be considered a very time-consuming task due to the massive size of the borehole data. Due to the complexity of the structures of interest and several noisy artifacts in the image log, traditional image processing methods are not very effective in solving the problem. The U-Net proposed by Olaf Ronneberger et al. is a convolutional neural network model commonly used in medical image segmentation that has been applied on several areas. This architecture is composed of two parts: the encoder, which is used to capture the image context, and the decoder, which is used to allow a precise location using transposed convolutions. A series of changes in this architecture has been proposed to improve the network's capacity to extract features in multiple scales and improve the skip connections between encoder and decoder. The DC-UNet (Dual Channel U-Net) is one of the U-Net based models designed to overcome some limitations of the original network. In this work, we present the application of DC-UNet for breakout segmentation in wellbore's amplitude image logs. Furthermore, we also discuss the problems related to the inaccuracy of the data's annotated masks and the image pre-processing strategies applied to reduce their effects.

Keywords: Image Logs, Semantic Segmentation, Breakouts, Deep Learning

1 Introduction

The drilling of a borehole can modify the local *in-situ* stress state of the rock formation, leading to potential well bore damage or failures. A stress concentration occurs in vertical wells with a principal stress axis, possibly causing the collapse of borehole wall zones in the direction of minimum horizontal stress, called Breakouts. The use of borehole breakouts is a recognised method for determining in-situ stress direction within boreholes.

Acoustic image logs are usually used for identification and characterisation of geomechanical structures such as breakouts. In image logs they are recognized as twinned, vertical, and irregular cavities approximately 180º apart. Due to the enormous magnitude of the borehole data, the manual identification of breakouts and other borehole wall structures by expert interpreters is generally a very time-consuming operation.

Even though the search for methods to automatically detect and segment these structures seems very important, there aren't effective solutions proposed for this problem yet. In a recent work, Dias [2] proposed a method for breakout detection using the Fast Region-based Convolutional Neural Network (fast-RCNN). The neural model is trained using synthetic acoustic logs with simulated breakouts. Even though the proposed solution achieves high accuracy in the synthetic images, it doesn't perform so well in real data. Also, the method only provides detection and not segmentation of the object. The complex characteristics of borehole structures and several noisy artifacts in the real well image logs are some of the challenges to be overcome.

In this paper we use a convolutional neural network to find a method for automatically segmenting breakouts in borehole acoustic image logs. The next topics of this document give information about the acoustic image logs and the problems regarding the annotation dataset available, followed by a brief description of the model's architecture used. After that, the method description is presented, including the pre-processing techniques applied to attenuate the data problems and, finally, the results achieved by the proposed work.

2 Data Information

Acoustic image logs are acquired through the emission of acoustic waves pulse. The log tool registers the travel time (msec) and amplitude (dB or mV) of the reflected pulse. The amplitude information is related to the variations due to impedance difference between well's fluid and the rock formation, while travel time is related to the distance of wall surface to the tool. For this work, only the amplitude log data is used as the single channel input of the network. The available dataset for this study was provided by Petroleo Brasileiro S.A. under privacy conditions.

2.1 Ground truth mask

The ground truth data used for the supervised training was generated from the histogram segmentation of the amplitude image and manual adjustments done by interpreters. Through an image log segmentation tool, the amplitude image histogram is segmented in multiple ranges of values (see details in Menezes [1]). Breakouts and other types of geological structures are typically located in the lowest amplitude values registered. Interpreters manually separate the log samples of breakouts from the others.

Because of the limitations caused by this semi-automatic method of annotation, the ground truth mask obtained is not 100% accurate. Also, the generation of segmented regions by histogram thresholding results in problems related to the noisy distribution of a real data such as fragmented regions, inaccurate borders and small false positive regions. To mitigate these problems, some pre-processing strategies were applied to the labeled image log and described in subsection 4.1.

3 Architecture Overview

The DC-UNet [8] architecture is briefly described in this section, emphasizing how it differs from traditional U-Net. Along with the loss function applied to every experiment.

3.1 DC-UNet

The convolutional neural network U-Net by Olaf Ronneberger [9] is a popular architecture for the segmentation of biological images. Due to its distinctive U-shaped encoder-decoder architecture, it is made to learn from little training samples. The blocks are organized in the shape of a ”U” with 4 encoder blocks on the left side and 4 decoder blocks on the right side. The encoder first receives the input image, after which it applies many convolutional layers to extract usable characteristics from the image. The decoder then uses transpose convolution to upsample the features and skip connection to concatenate them. As a result, the output of this network is a segmentation mask.

DC-UNet is a slightly modified version of MultiResUNet, which in turn is an evolved model of U-Net. The MultiResUNet [5] can produce a significantly superior output than the U-Net since it can supply different scaling properties. The MultiRes block's objective is to provide a variety of scale information that can be utilized to help separate an object from the full image. The MultiRes block was modified by the DC-UNet to add more practical features. This idea served as the author's new foundation for advancement.

The potential of image segmentation has been proved by features at various scales. Therefore, DC-Unet replaces the remaining connections in the MultiRes block with a series of three $3x3$ convolutional layers to address the issue of insufficient spatial information. Dual Channel is the name of this block, as seen in Figure 1.

DC-UNet uses the same connection (Res-Path) as MultiResUNet between encoder and decoder. Then use the Res-Path and Dual-Channel modules to build a new U-Net architecture, as shown in Figure 2.

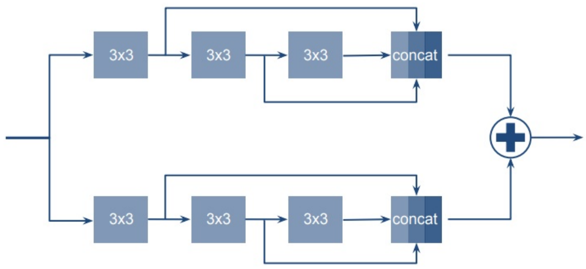


Figure 1. DC-UNet Dual Channel Block

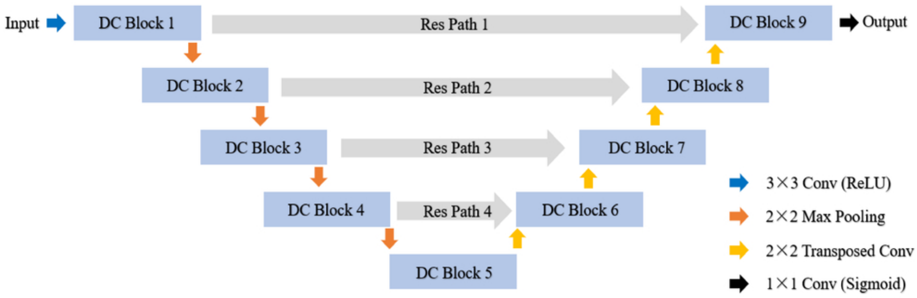


Figure 2. DC-UNet Architecture

3.2 Loss Function

Semantic segmentation aims to determine whether a pixel belongs to the object. As a result, this problem could be characterized as a binary classification problem at the pixel level. We choose to set the loss function to minimize the binary cross-entropy [10]. The model's forecast for the input picture X is y, but the actual value is $y\prime$. As a result, the loss function and binary cross-entropy for a batch of n images are defined as follows, respectively:

$$CrossEntropy(y, y\prime) = \sum -(y \log(y\prime) + (1-y)\log(1-y\prime)) \tag{1}$$

$$loss = \frac{1}{n}\sum_{i=1}^{n} CrossEntropy(y, y\prime) \tag{2}$$

We used the Adam optimizer by Kingma [6] to train the models, with the parameters $beta1$ = 0.9 and $beta2$ = 0.999. Early Stopping technique is used during training, therefore the amount of epochs was diverse.

4 Method Description

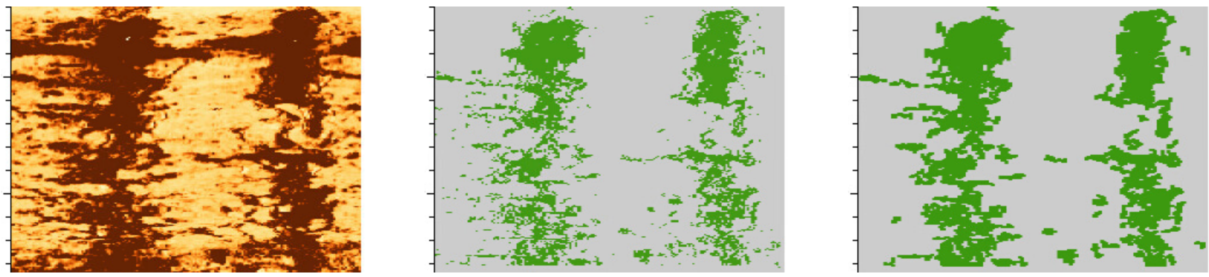
This section describes the proposed method in more detail. The adjustments required to process the input and ground truth data are listed in the subsection below.

4.1 Pre-processing

The quality of various well data collected over the years may vary due to logging parameters and technical limitations. Also, older wells tend to have more noise, especially at the deepest part of the profile. For these reasons, we emphasize the need for the preprocessing steps outlined below.

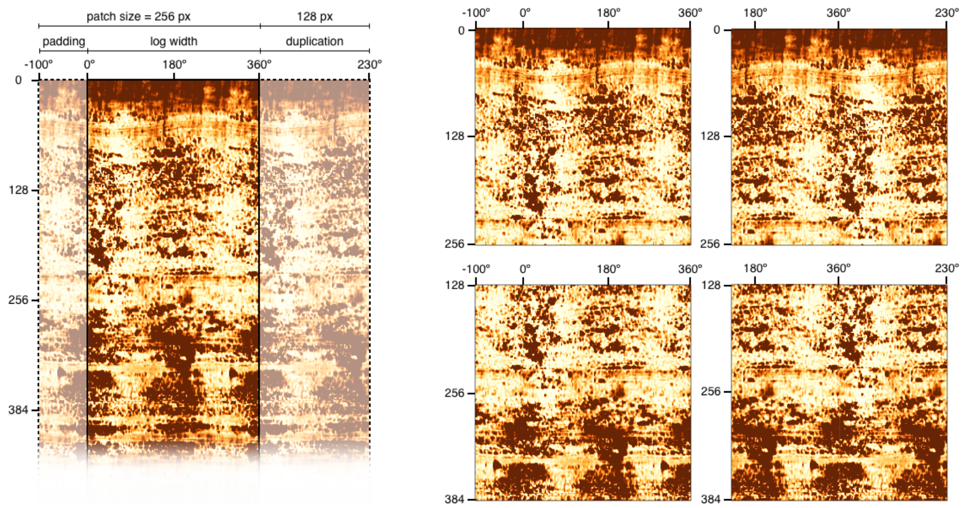
For the amplitude data, firstly we applied the Inter Quartile Range method for outlier removal, followed by a step of contrast adjustment to improve the variance inside the masks since there are masks in our dataset with a higher variance between the amplitude samples, implying lower likelihood for each sample. This preprocessing step is also commonly used in seismic interpretation software to improve the well data visualization. Final pre-processing step is normalization of the amplitude data to [0,1] range. This step is often used in machine learning applications to help decision models better estimate weights since all observed samples are in the same range of values.

For the segmentation mask data, a small regions filtering is used to reduce the noisy artifacts from the labeled data. The filter removes all regions with area smaller than 10 pixels. After that, a dilation morphological operation is applied to each connected region in segmentation mask. Due to the problems reported in subsection 2.1, this operation expand the regions boundary and better delimit the breakout areas, in our case applying a 2x2 kernel. Dilation also helps to join small fragmented regions that should be a single connected region.



(a) Amplitude data. (b) Original segmentation mask. (c) Segmentation mask after filtering and dilation operation.

Figure 3. Example of a mask section before and after preprocessing operations.



(a) Well log section with left extension to complete patch size and right extension of half patch size

(b) Extracted patches from well log section on the left with sliding window method and a stride of half patch size

Figure 4. Example of patch extraction from image log with 200 pixels width and patch size of 256 pixels

4.2 Patches extraction strategy

In order to send the data as input to the model, it is necessary to sample the well log in small patches. The sliding window method is used for extracting patches from the image data. The log width varies for each well between 180 to 256 azimuthal samples. A patch of size 256x256 pixels was defined so that every patch has at least the whole borehole wall horizontally. The left side of the image log was filled with the last columns in order to achieve the patch size. By concatening the first 128 columns to the right side of the log and defining a step of the same size, it was possible to generate a second patch with a different rotation, varying from 180 to 256 degrees. The vertical step of the sliding window was also half patch size, generating overlaping. Figure 4 illustrates the proposed technique. Considering the characteristics of the data's domain, this method should be considered a type of data augmentation technique.

4.3 Evaluation Metrics

Recall, Precision, F1-Score, and Intersection over Union (IoU) are four conventional metrics for the image segmentation task that we used to evaluate the proposed approach. The precision metric is a measure of quality, while the recall metric is a measure of quantity. High recall indicates that an algorithm provides the majority of the relevant results, while higher precision indicates that an algorithm delivers more relevant results than irrelevant ones. By taking their harmonic mean, the F1-Score integrates a classifier's precision and recall into a single metric. An evaluation metric called intersection over union is used to assess an object detector's precision on a certain dataset. The above metrics are indicated by the following expressions:

$$\text{Recall} = \frac{TP}{TP + FN} \tag{3}$$

$$\text{Precision} = \frac{TP}{TP + FP} \tag{4}$$

$$\text{F1-Score} = \frac{2 \mid Recall * Precision \mid}{\mid Recall \mid + \mid Precision \mid} \tag{5}$$

$$\text{IoU} = \frac{\mid Recall * Precision \mid}{\mid Recall \mid + \mid Precision \mid - \mid Recall * Precision \mid} \tag{6}$$

Where TP is the number of true positives (breakouts samples are classified as breakouts), TN is the number of true negatives (non-breakouts are classified as non-breakouts), and FP is the number of false positives (non-breakouts classified as breakouts) FN is the number of false negative samples (samples for which a breakout is classified as a non-breakout).

5 Experiments and Discussion

Cross validation is a frequent technique for assessing a model's performance. For the k-Fold cross-validation test, the dataset X is arbitrarily split into k roughly equal-sized, mutually exclusive subsets $(X1, X2, ..., Xk)$. The model is run k times, with each run using one of the k subsets as the training set and the rest as the validation set [3]. We calculated the model's performance by looking at the combined results of k training runs. For this work, the data from 18 wells of different sizes were used in 5-fold combination. The Table 1 below shows the train, validation and test set sizes (and its percentage ratio in relation to total samples) for each fold.

Table 1. K - Fold Description

Partition	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5
Train	7 wells (73%)	7 wells (73%)	6 wells (56%)	10 wells (72%)	11 wells (72%)
Validation	3 wells (16%)	4 wells (12%)	2 wells (26%)	5 wells (12%)	4 wells (12%)
Test	8 wells (11%)	7 wells (15%)	10 wells (18%)	3 wells (16%)	3 wells (16%)

Table 2. Metrics

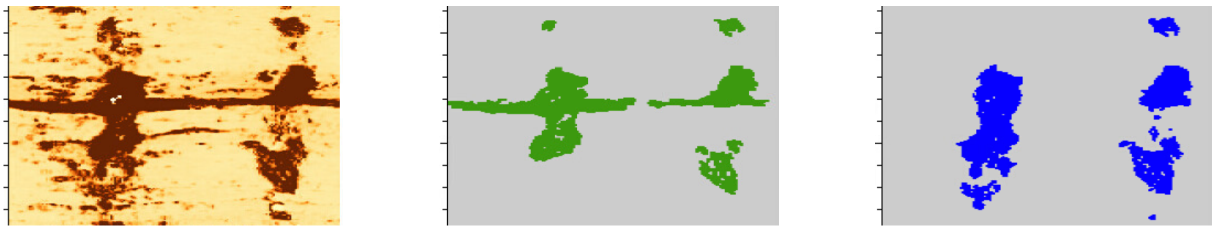
Partition	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Average
Precision	66.79%	80.86%	75.89%	95.16%	84.37%	80.61%
Recall	39.89%	75.74%	73.45%	72.63%	78.16%	67.97%
F1-Score	49.95%	78.22%	74.65%	82.38%	81.15%	73.27%
IoU	33.29%	64.23%	59.55%	70.03%	68.28%	59.08%

The first epochs of the training process often have the highest error rates. This is due to the model's weights being randomly assigned, while there are several approaches around this issue, such as utilizing learned weights for a similar task (fine-tuning), or using an initializer that still has the issue but to a lower degree. The warm-up technique [4] involves raising the learning rate by a factor over a predetermined number of repetitions from zero to the specified value. The training process will then start after a certain number of epochs to learn at the maximum pace, and then the decay using the cosine function will start.

By analyzing the result of the metrics presented in Table 2, the overall performance of the model denote the effectiveness of the suggested approach, achieving high scores in 4 of 5 fold subsets and predicting breakout zones with a low rate of false positives. The example of a well log section with breakout structures in figure 5 confirms the model's capability to perform a good segmentation of the breakout regions. The false positive examples are located at the border of the structures and justify the low values of Recall and IoU, indicating the need of a refinement during training to focus the learning on the hard examples, ie samples located at the boundaries limit.

(a) Amplitude log section. (b) Ground truth segmentation mask. (c) DC-Unet's prediction mask.

Figure 5. Example of prediction result from DC-UNet. F1-Score: 75.81% IoU: 61.04%



(a) Amplitude log section. (b) Ground truth segmentation mask. (c) DC-Unet's prediction mask.

Figure 6. A DC-UNet prediction result more correct than the ground truth. F1-Score: 73.20% IoU: 57.72%

(a) Amplitude log section. (b) Ground truth segmentation mask. (c) DC-Unet's prediction mask.

Figure 7. Example of prediction result from DC-UNet. F1-Score: 90.14% IoU: 82.04%

As explained in section 2.1, the ground truth was done semi-automatically using image processing techniques and manual corrections. This phenomenon resulted in incorrect marking in some cases, as can be seen in Figure 6.

Thus, the segmentation metrics were affected at the time of prediction, in scenarios where the model's prediction is better than the ground truth itself.

For breakout regions with less noisy artifacts in the background such as the example in figure 7, the model is able to predict an almost perfect segmentation result.

6 Conclusions

In this paper, a CNN-based method for finding breakouts in borehole acoustic image logs is described. To verify and evaluate the method, we used a set of real log well data that Petrobras owned privately and that included drilled wells with the breakout anomaly. The quantitative metrics and model predictions highlight the quality of the results.

Even with sparse training data, the suggested technique produced encouraging results, with good segmentation metrics and visual outcomes that corroborated the numerical ones. From this, it can be concluded that the offered method can help experts identify possible breakouts in well log image data. The interpretation procedure should then be optimized to reduce costs and the length of the interpretation phase.

We suggest using transit time details together with the input amplitude data in further study and also test loss functions that focus the learning on hard examples in order to improve detection at the borders such as the Focal Loss presented by Lin [7]. The data augmentation technique presented in this work should be further explored to raise the evaluation metrics. The addition of sophisticated self-supervised training methods, including contrastive learning [11] to avoid the fact that training labels are produced using basic image processing methods, in an effort to improve the model's generalization, would be another stage of progression.

Acknowledgements. Acknowledgements. The present work was carried out with the support of the Tecgraf Institute Petroleo Brasileiro S.A (Petrobras) and the CNPq (National Council for Scientific and Technological Development).

Authorship statement. The authors hereby confirm that they are the sole liable persons responsible for the authorship of this work, and that all material that has been herein included as part of the present paper is either the property (and authorship) of the authors, or has the permission of the owners to be included here.

References

[1] Candida de Jesus, Andre Compan, and Rodrigo Surmas. Permeability estimation using ultrasonic borehole image logs in dual-porosity carbonate reservoirs. *PETROPHYSICS*, 57:620–637, 12 2016.

[2] Luciana Olivia Dias, Clécio R. Bom, Elisangela L. Faria, Manuel Blanco Valentín, Maury Duarte Correia, Márcio P. de Albuquerque, Marcelo P. de Albuquerque, and Juliana M. Coelho. Automatic detection of fractures and breakouts patterns in acoustic borehole image logs using fast-region convolutional neural networks. *Journal of Petroleum Science and Engineering*, 191(June 2019):107099, 2020.

[3] T. Hastie, R. Tibshirani, and J.H. Friedman. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer series in statistics. Springer, 2009.

[4] Tong He, Zhi Zhang, Hang Zhang, Zhongyue Zhang, Junyuan Xie, and Mu Li. Bag of tricks for image classification with convolutional neural networks, 2018.

[5] Nabil Ibtehaz and M. Sohel Rahman. MultiResUNet : Rethinking the u-net architecture for multimodal biomedical image segmentation. *Neural Networks*, 121:74–87, jan 2020.

[6] Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization, 2017.

[7] Tsung-Yi Lin, Priya Goyal, Ross Girshick, Kaiming He, and Piotr Dollár. Focal loss for dense object detection, 2017.

[8] Ange Lou, Shuyue Guan, and Murray Loew. Dc-unet: Rethinking the u-net architecture with dual channel efficient cnn for medical images segmentation, 2020.

[9] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In Nassir Navab, Joachim Hornegger, William M. Wells, and Alejandro F. Frangi, editors, *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, pages 234–241, Cham, 2015. Springer International Publishing.

[10] Zhilu Zhang and Mert R. Sabuncu. Generalized cross entropy loss for training deep neural networks with noisy labels. *CoRR*, abs/1805.07836, 2018.

[11] Xiangyun Zhao and Philip Andrew Mansfield. Contrastive learning for label efficient semantic segmentation. 2012.