Lithology identification using semantic segmentation for well log data

http://dx.doi.org/10.1590/0370-44672023770096

Áttila Leães Rodrigues[1,2]
https://orcid.org/0000-0003-4524-4087
Fernanda Gontijo Fernandes Niquini[1,3]
https://orcid.org/0000-0003-1872-1466
Sandro Pinzon[1,4]
https://orcid.org/0000-0001-5951-1743
João Felipe Coimbra Leite Costa[1,5]
https://orcid.org/0000-0003-4375-370X

[1]Universidade Federal do Rio Grande do Sul – UFRGS, Departamento de Engenharia de Minas, Porto Alegre – Rio Grande do Sul – Brasil.

E-mails : [2]attila.leaes@ufrgs.br, [3]fernanda.gontijo.fn@gmail.com, [4]sandro.pinzon@ufrgs.br, [5]jfelipe@ufrgs.br

Abstract

In the past decade, machine learning techniques were responsible for a revolution in classification and regression tasks, making it possible to automate some laborious activities, saving time and reducing errors. It is known that the geological logging process is one of the most time-consuming activities accomplished by mining companies. Additionally, it is a subjective activity, and changes in the staff directly affect the geological databases due to different human log interpretation. By developing an automatic log classifier, a company can avoid problems related to the turnover of the staff by standardizing the criteria used to label an interval and can save time by avoiding manual log description. The proposed solution is: given the well log data containing the coordinates, resistivity and natural gamma, the model will be able to predict the presence or absence of coal, and its lithology. The innovation of the methodology proposed, considers not only the geophysical logging values, but additionally inserts the neighbourhood of a given depth as valuable input information, using Fully Convolutional Network. It performs a semantic segmentation using the well log data, which means that model´s input is the complete well log data curve and the trained model will return an output curve giving the probability of the presence of coal, by interval. The results showed good prediction for the binary problem (F1-score 0.79). The multi-class modelling suffers from the lack of data for each class, resulting in a F1-score from 0.38 for the worst result to 0.76 for the best.

Keywords: well log data, lithology identification, machine learning, semantic segmentation, Fully Convolutional Networks.

1. Introduction

Geophysical logging is widely used in mining and petroleum, especially for its ability to discriminate strata and the potential to replace laboratory analyses made for certain physical and chemical parameters of interest. In the case of coal deposits, geophysical logging could be used to delineate the coal/rock interfaces (Hoffman *et al.* 1982 Asfahani and Borsaru 2007), assisting in establishing stratigraphic correlation, and eventually generating estimates of quality parameters and geomechanical requested data. In short-term mine planning, due to the impossibility of using diamond drilling at operating benches and faces, geophysical logging can be used quickly at blasting holes as an alternative to obtaining lithological contacts and the other estimates mentioned above. In this study, geophysical logging was used to determine the coal and waste rock, by analyzing the contrasts between the measures given by the equipment along the drill hole. The negative aspect regarding the use of geophysical logging is related to the need to use radioactive sources, such as Cesium 137 and Americium–beryllium in these probes. The radioactive sources coupled with these probes give rise to environmental concerns because there is always the possibility of their getting trapped in the drill holes. Therefore, handling the radioactive source requires accredited personnel and special care. In spite of the negative points above-mentioned, this research area is on the uprise.

Another field of study that is under expansion in the mining context is machine learning, which according to Lary *et al.* (2016), is part of the artificial intelligence field: its algorithms can learn from the data using multivariate, nonlinear, nonparametric modeling for classification and regression. The quality of information, combined with a good choice of predictive models, will significantly increase the possibility of getting better forecasts to be used in the mining routine. The present study uses artificial intelligence methods capable of indicating the correspondences between the type of ore and its physical, chemical, and mineralogical characteristics. The hypotheses investigated are: i) can the lithotype (layer) be determined through the usage of trained algorithms, and ii) fed with readings of physical-chemical properties? Would classification/agglomeration/regression methods be useful in identifying lithologies?

It is expected that an almost automatic separation or individualization of the lithological units or their origins, with a given degree of certainty, can be obtained. This could help the geomodeler to quickly get a determination of a lithotype, using existing and available data. Additionally, this can be used as a checking tool for old campaigns where the input data is available or easily accessible, by highlighting intervals where the geological descriptions made do not match with the resistivity and natural gamma found. It can also be used in a descriptive way, by helping to understand which variables can best characterize or contribute to identify the geological material.

1.1 Analysing geophysical trace using machine learning algorithms

Lithological classification using machine learning methods have been explored (Imamverdiyev and Sukhostat 2019; Antariksa *et al*. 2022; Salehi and Honarvar 2014; Dev and Eden 2019) usually in a “patch based” modelling. We understand there is space for improvement in prediction when considering the curve for each measured variable. To accomplish this improvement, we propose modelling the well log data by using the full log curve as input to a Fully Convolutional Network based on 1d convolutional layers to perform a semantic segmentation.

Semantic segmentation is well known in image and pattern recognition: after taking a photo or image as input, the semantic segmentation algorithm is capable of recognizing the elements present in the image at the pixel-level. When considering well log data, the inputs are the 1d vectors (one vector for each log variable, one entry for a given depth value) and a semantic segmentation algorithm will determine where the lithologies occur.

A Fully Convolutional Network (FCN) (Fukushima 1988; Shelhamer *et al*. 2017) is a segmentation solution based on Deep Learning (Schmidhuber 2015). The main characteristic of an FCN is its exclusive use of convolutional layers in the network, with no downsampling, such as with max pooling layers. These convolutional layers are responsible for learning the spatial correlations. Considering well log 1-dimensional data, these spatial correlations are the possible patterns of the measured quantities across the 1d vector (across the depth). A FCN also permits inputs with different depth sizes, which is the case of this study. Fully Convolutional Networks were used in Zhu *et al*. (2020) to predict the occurrence of gas in a gas reservoir using well log data. Our study comes to enrich this field with some progress and differences.

The next sections will highlight some key aspects of the study: an explanation about the coal deposit under study will be given in section 2.1; details about the database used (section 2.2) and the geophysical logging values (section 2.3) will also be presented; a quick description about Fully Convolutional Network will be given in section 2.4; the results obtained through the application of the techniques herein presented will be shown in section 3 and the discussion about the results obtained and the conclusions will be given in section 4.

2. Materials and methods

2.1 Geological details about Candiota Sedimentary Basin

The Parana Basin is the largest intracratonic basin on the South American platform, distributed within Brazilian, Argentinian, Uruguayan and Paraguayan territories. It has a NE–SW direction with dimensions of approximately 1750 km long by 900 km wide. It represents a sedimentary-magmatic succession with ages between the NeoOrdovician (465 Ma) and the Neo-Cretaceous (65 Ma). It was marked by cyclical events of subsidence and uplift, with thicknesses reaching up to 7,000 m in its deepest portion. The Gondwana I supersequence, which bears the coal seams, is represented by sedimentary successions that define transgressive-regressive cycles linked to fluctuations in the relative sea level during the Paleozoic Era. The Seival and Candiota areas are briefly composed of the rocks of the Rio Bonito formation and Palermo formations. The Rio Bonito formation can be subdivided into three main horizons, with coal seams separated by sandstones or siltstones.

In the Candiota deposit, in addition to the two layers of the same name Candiota Superior (CS) and Candiota Inferior (CI), there are up to nine upper layers (from S1 to S9) and up to nine lower (from I1 to I9) (Projeto Candiota - CPRM, 2016). Particularly in the database used, there are S layers that may not be present due to the uplift of the Candiota Block, which caused the erosion of these layers. In the lower layers "I", separated by gray siltstones, we have identified layers I1 to I4 in the cores. Above the Candiota layers, we also have the Banco Louco (BL) layer, characterized by high ash content. Figure 1 shows the spatial locations of the drill hole collars.

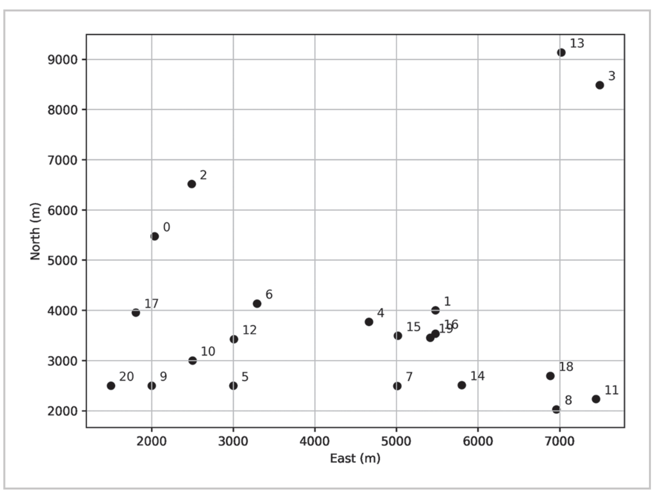

Figure 1 - Location of the drill hole collars. The dataset is comprised of 21 diamond drill holes.

2.2 Database information

A coal mining company, located in the South of Brazil, provided the database used, consisting of 21 drill holes with resistivity (Ω·m) and natural gamma (API Cs.) measurements. In the case of the database used, the number of layers observed is not constant in all holes, whereas for most holes, there are incomplete sequences of layers S6 to I4, with the greatest absences being the layers S6, BL and I4. The layers “S” and “I” have an average thickness of around one meter, and the layers CS and CI combined, present 3.85 meters in average. As shown in Figure 1, the holes are arranged on an irregular grid, with an average depth of 52 m, ranging from 22 to 74 m. The shallow depth of the deposit results in an open pit extraction in the region.

2.3 Obtaining geophysical logging values

The geophysical profiles were obtained using a GLOG® (Focussed Electric Guard Log Sonde). The natural gamma and resistivity values were processed with the Winlogger software in a suitable format to obtain the values along the coal seams. An experienced professional previously identified the coal seams to obtain a correct correlation between the seam and the values received. The instrument reads continuously along the hole. However, when processing, we used a one-centimeter window interval.

Table 1 shows the lithologies for each well, where the numbers represent the length of the correspondent lithology in centimetres. Not all lithologies are present in all wells. The S6 lithology samples, for example, exist only in eight different wells. The most abundant lithology is CI, which is present in all wells. The well log data from 0 to 10 meters depth were discarded due to noise and the high number of outliers.

Table 1 - Length by lithology (in centimeters) for each well. The ‘Coal’ column represents the sum of the length of all lithologies. The ‘Other’ column represents all depth points that are not coal.

Well	BL	CI	CS	I1	I2	I3	I4	S2	S3	S4	S5	S6	Other	Coal
0	0	130	241	0	0	155	140	58	0	0	88	0	6393	812
1	77	153	257	111	99	92	90	89	117	84	101	153	3007	1423
2	0	143	0	130	92	120	55	64	107	0	0	0	6672	711
3	94	152	279	115	78	40	0	75	0	0	91	49	6196	973
4	100	108	253	86	78	49	0	68	102	67	83	0	6389	994
5	0	140	236	105	82	70	120	0	0	0	0	95	4438	848
6	97	125	240	110	75	79	88	0	0	0	0	89	6480	903
7	79	165	0	105	120	93	120	60	109	60	0	0	6472	911
8	0	163	275	98	113	82	63	76	109	73	85	0	6246	1137
9	117	109	225	120	83	96	0	96	110	111	109	241	5966	1417
10	131	93	205	101	89	111	215	92	100	108	101	238	5799	1584
11	105	153	265	137	0	0	0	77	130	101	90	0	6325	1058
12	89	172	263	114	88	133	52	75	94	0	0	0	6303	1080
13	0	142	243	69	95	100	0	68	97	76	83	85	6325	1058
14	0	146	224	112	80	0	0	0	0	0	0	0	6821	562
15	0	158	282	108	78	111	127	63	110	83	93	0	6170	1213
16	0	143	259	125	98	0	0	74	114	105	110	230	4028	1258
17	0	145	0	0	82	122	82	0	0	0	0	0	6952	431
18	0	144	230	103	76	115	0	70	0	0	69	0	6421	807
19	0	144	211	116	66	118	135	88	94	90	65	0	6256	1127
20	83	146	194	97	0	78	77	69	0	0	0	0	6639	744

The methodology uses the core logged data to train a neural network to predict where the coal appears. To predict a given core lithology, the input are the two curves obtained from geophysical logging (resistivity and natural gamma, Figure 2) and the output is the probability of each interval to contain coal. Next, the same neural network was trained to predict each specific coal seam, this being a multi-class scenario. The network adjusted was a Fully Convolutional Neural Network, which will be better explained in section 2.4. The training procedure and performance assessment of the trained model are detailed in section 2.4.3.

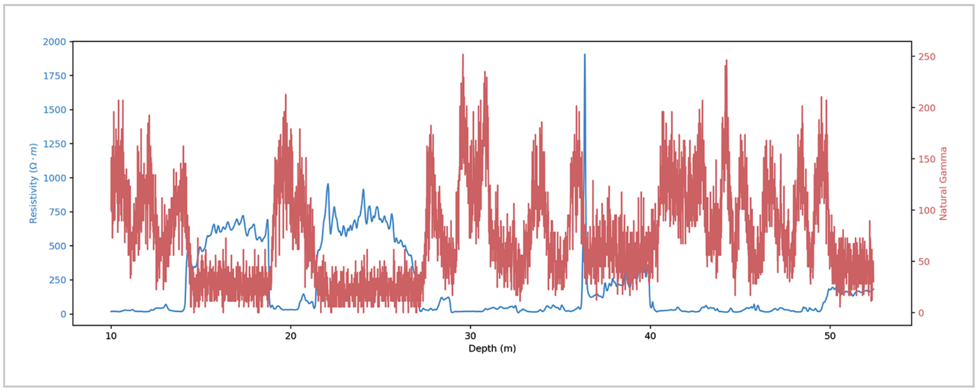

Figure 2 - Example of captured well log data (well 16). The resistivity and natural gamma curves are provided as input to the fully convolutional neural network for a given well.

2.4 FCN - Fully Convolutional Networks

A Fully Convolutional Network (FCN) permits inputs with different sizes, which is necessary in this study since each drill hole contains different thicknesses. The chosen architecture is detailed in Figure 3. It comprises 10 identical convolutional layers with the same size (length) as the input.

After 10 convolutional layers, the tensor is processed by the last convolutional layer, which also has the same size of the input, but with kernel size equal to 1 and the number of filters equal to the number of classes. The data is processed by a Softmax activation function, which normalizes the probability for a given interval, so all class probabilities add 1.

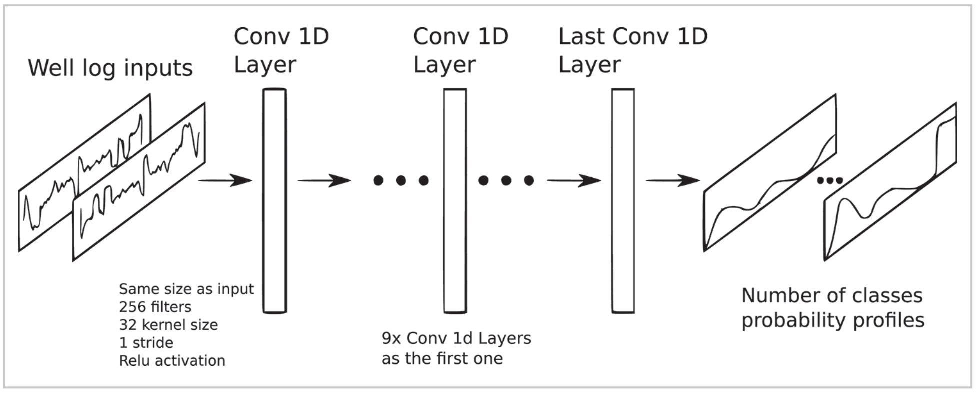

Figure 3 - Architecture of the fully convolutional network. The inputs are two variables from the well logs, followed by ten convolutional layers with the same size as the inputs. The last part consists of a 1D convolutional layer with kernel size equal to 1 and number of filters equal to the number of classes presented in the dataset. The result is a probability profile for each class.

2.4.1 Input tensors for FCN

The input tensor for a given well has the two 1d vectors, resistivity, and natural gamma, and its dimensions are $n_i \times 2$, where n_i denotes the number of data points for well *i*. Thus, the input tensors will have different lengths depending on the specific well.

For the output tensor, there are two different scenarios of segmentation of well log data: the first one is a binary segmentation, which means the neural network will classify the input data as coal or not coal. In this binary case, the output is $n_i \times 1$, which means one probability curve of the same size as the input (values between 0 and 1).

The second scenario has the input configuration but was retrained to predict the different coal classes. In this case, the output tensor will have the dimensionality $n_i \times n_{classes}$, where nclasses is the number of different coal classes plus 1 (not coal, or "other"). Again, each 1D vector for each class in the output tensor is a probability result (values between 0 and 1), of which the sum over all classes, at a given depth, equals 1. In the database used to train the network, the output tensors are filled with zeros or ones, because there is knowledge about the presence or absence of coal (and its lithology).

2.4.2 Softmax activation function and Loss function

Inside the FCN, two functions were adjusted: the Softmax activation function which normalizes the sum of the probabilities over the classes and the loss function, calculated across the classes and averaged over the entire well log size during the training. A brief explanation about them will be given below.

The Softmax function normalizes a vector to a probability distribution, i.e., each component lies in the interval (0,1] and the sum adds 1. Given a vector *y* with size *N*, the Softmax normalization becomes

$$\sigma\,(y)_j = \frac{expz_j}{\sum_{i=1}^{N} expz_i}, \qquad (1)$$

where N denotes the number of classes and the index j indicates the specific depth.

The purpose of the loss function is to calculate the distance between a probability tensor generated by the Softmax function and the true values. The loss function chosen was the categorical cross entropy, defined by

$$Loss(y, \hat{y}) = -\sum_{i=1}^{N} y_i . \log \hat{y}_i , \quad (2)$$

where N is the number of possible classes, y is the target vector, and $\hat{y}$ is the probability vector. The loss function is calculated considering y in equation (2) as the tensor with dimension number of classes × size of the input vector (well depth).

2.4.3 Model training and evaluation

The training steps and the model evaluation phase for the two scenarios (binary and multi-class) are described as follows:

1. Initiate the Leave One Out Cross-Validation (LOOCV) by well. Select one well for the test set and the remaining wells for the training set.
2. Standardize each variable in the training set. The learned transformation is applied to the test set.
3. Calibrate the FCN model parameters using the training set.
4. Evaluate the segmentation model using the test set.
5. Repeat the steps using another well log as the test set with the training set being composed of all the remaining wells.
6. Evaluate the FCN model performance for all test sets.

To clarify the step by step presented above, illustrated now is the methodology naming the wells: the LOOCV by well means, for example, that the first well (id 0) is used as the test set, and the remaining 20 wells compose the training set. The training set is standardized, and the mean and standard deviation of the training variable are used to standardize the test set (well 0). No data imputation was needed in the preprocessing step. The performance of the trained model is assessed using the test set, namely, well 0. The process is then repeated for the next well (id 1), and the standardization and training are repeated for the new training and test sets. The performance is measured using well 1 as the test set. The process is repeated until all wells belong once to the test set. This procedure is a possibility for spatial cross-validation regarding wells individually. This spatial cross-validation field is getting increased attention due to the machine learning models which are being used to predict estimates and are prone to overfit (Wang *et al*. (2023) Brenning (2022) Ploton *et al*. (2020) Valavi *et al*. (2018) Deng *et al*. (2017)).

2.4.4 Performance Metrics

The FCN model performance is scored by the Precision, Recall, and F1-score metrics, described below.

The Precision score is defined as T_p / (T_p+F_p). It represents the number of correct predictions of "coal" (true positives, T_p) divided by the total number of predictions of "coal".

The Recall score (also called true positive rate) is defined as the number of correct predictions of coal layers divided by the total number of coal layers existent in the database. The expression for the recall is T_p/ (T_p+F_n), where T_p is the number of true positives and F_n is the number of false negatives.

The Precision and Recall scores measure different aspects of the model predictions. A way to summarize these scores is to calculate the F1-score, defined by the harmonic mean of the Precision and Recall (equation 3):

$$F1 \text{ - score} = \frac{2 \; . \; Precision \; . \; Recall}{Precision + Recall} \; . \quad (3)$$

2.4.5 Code

The workflow was programmed in Python using Keras and Tensorflow libraries. The script and the dataset are available in the fully-conv-network-1d-coal Github repository (Rodrigues, 2022) where graphical results for all wells can be seen.

3. Results

Several combinations of the FCN parameters were tested, mainly the number of convolutional layers, the kernel size, and the number of filters. A specific set of parameters was chosen when the performance metrics increased.

Table 2 shows the main parameters for the Fully Convolutional Network used in this study. The same set of parameters were used to train the model for the binary case and for the multi-class case.

Table 2 - Chosen parameters for the Fully Convolutional Network.

epochs	4000
number of 1d convolutional layers	10
kernel size	32
number of filters	256
optimizer	Adam
learning rate	10^{-5}

3.1 Binary results

The binary model labels the lithologies as "Coal" or "Not Coal". The fully convolutional neural network was trained following the step-by-step presented in section 2.4.3.

Table 3 shows the results (F1-score, Precision and Recall) for all wells in the test set. The mean score is in the last row, being 0.79 for the F1-score, 0.79 for the Precision, and 0.80 for the Recall. The results were rounded to the second place after the decimal point.

Figure 4 shows the true data and the model prediction for well 19 when it was used as the test set, a case of good prediction. Although all the coal layers are predicted by the trained model, the performance score is not perfect. Even in this good case, the F1-score is 0.89, the precision is 0.90, and the recall is 0.88. For a perfect score, the prediction of the coal layers must match exactly the beginning and the end for each layer. Even a small translation of a layer prediction will worsen the score.

Table 3 - Results for the binary modelling ("Coal" or "Not Coal").
The well id column means that the given well is the test set and all others make up the training set.

well id	F1-score	Precision	Recall
0	0.78	0.77	0.78
1	0.69	0.80	0.61
2	0.74	0.62	0.90
3	0.80	0.81	0.79
4	0.81	0.86	0.76
5	0.55	0.63	0.49
6	0.86	0.92	0.80
7	0.69	0.66	0.72
8	0.81	0.85	0.77
9	0.78	0.74	0.82
10	0.77	0.90	0.68
11	0.83	0.75	0.92
12	0.86	0.95	0.79
13	0.76	0.74	0.77
14	0.93	0.88	0.99
15	0.87	0.83	0.92
16	0.73	0.68	0.80
17	0.74	0.69	0.79
18	0.83	0.81	0.84
19	0.89	0.90	0.88
20	0.87	0.81	0.94
all (mean)	0.79	0.79	0.80

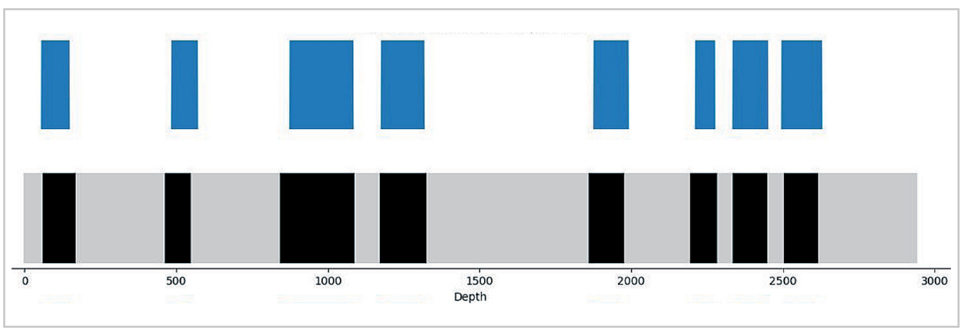

Figure 4 - True data for well 19 (above) and the model prediction (below). Selected from visual inspection of the outcomes as an example of good prediction. The scores when using well 19 as test set are: F1-score 0.89, precision 0.90 and recall 0.88.

Figure 5 shows the worst case of prediction in terms of the F1-score (0.55): the case of well 5. In this case, there are 2 predicted coal layers which do not exist in the true data, making the precision score lower (0.63). The width of these predicted layers directly impacts the precision score. There is also one coal layer in the true data, which is not correctly predicted by the model, the widest one. Although the model predicts a thinner version of this layer, the size of the unpredicted length of the true layer heavily penalizes the recall score (0.49). It is worth emphasizing that the accuracy metric may be misleading due to the imbalance between classes. For example, for the worst well, 5, the accuracy score is 0.85, which is because there is a high count of successful predictions of “not coal”.

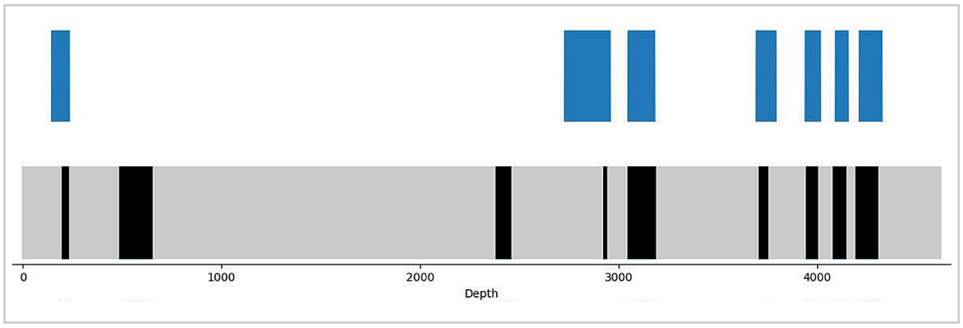

Figure 5 - True data for well 5 (above) and the model prediction (below). Selected from visual inspection of the outcomes as an example of a not-so-good prediction. The scores when using well 5 as test set are: F1-score 0.55, precision 0.63 and recall 0.49.

As explained in section 2.4.3, the cross-validation is performed well in a LOOCV fashion. Figure 6 shows how model’s performance varies, since the test well is far from the remaining wells, used as the training set (Brenning (2022)). This distance is defined as the cartesian distance, in XY plane, between the well in the test set to the nearest well in the training set. The results show that the F1-score maintains its average value across all distances. Some might imagine that the model’s performance could degrade for a well distant to the training data, which did not happen in this case study. This behaviour can be explained by the coal’s spatial continuity, which is well-known as high compared to other commodities. Another factor that can be accounted for is the nature of the prediction realized by this fully convolutional neural network, which is based on the coal’s resistivity and gamma natural signature in the well log. Although the signature is related to the spatial continuity, the well log curves may be less sensitive to variations in the anisotropy’s angle.

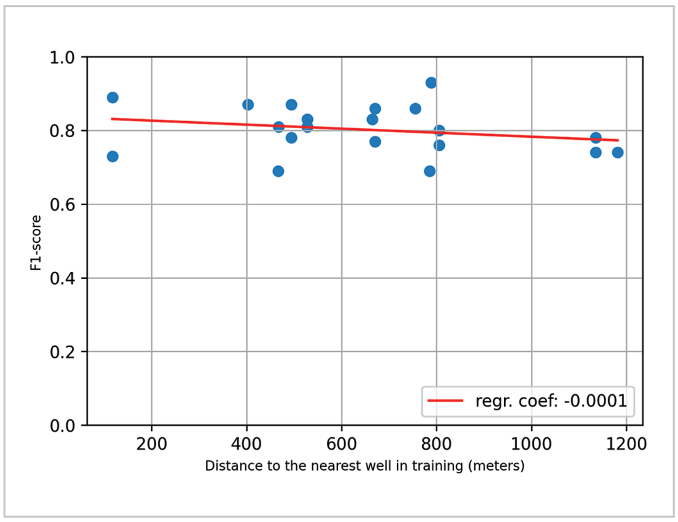

Figure 6 - Scatterplot of the F1-score and the distance of the well in test set to the nearest well in training set. In this study the trained model shows approximately the same performance regardless of the distance to the well in test set.

3.2 Multi-class results

For the multi-class results, the coal layers are labelled according to the corresponding class (lithology). There are 12 different classes of coal in the dataset, and one last class, which is the “not coal” class. This results in a total of 13 classes.

Not all classes appear in every well. In these cases, the performance scores are undefined. A performance score for a given class is not defined in a well where there is no data corresponding to this class in the test set.

Table 4 shows the F1-score for each well when it is present in the test set. Overall, the classes which are better predicted are (from easier to harder): CS, CI, S5, I2, I1, S3, S2, S4, I3, BL, I4, S6.

Table 4 - F1-scores using the individual well data as test set for each class. When the test set does not contain a given class to be tested, the score is represented by a dash. The last line is the mean F1-score calculated across the well data used as test set.

well id	BL	CI	CS	I1	I2	I3	I4	Other	S2	S3	S4	S5	S6
0	-	0.96	0.98	-	-	0.00	0.61	0.93	0.77	0.92	0.78	0.91	-
1	0.48	0.52	0.63	0.00	0.50	0.44	0.05	0.86	0.65	0.00	0.14	0.62	0.00
2	-	0.97	-	0.70	0.62	0.36	0.00	0.94	0.00	0.33	-	-	-
3	0.00	0.91	0.92	0.78	0.87	0.63	-	0.96	0.62	0.91	0.92	0.94	0.00
4	0.48	0.00	0.88	0.83	0.89	0.81	-	0.95	0.88	0.92	0.90	0.91	-
5	-	0.00	0.35	0.00	0.00	0.20	0.00	0.91	-	-	-	-	0.00
6	0.00	0.93	0.92	0.79	0.63	0.00	0.00	0.94	-	-	-	-	0.00
7	0.61	0.87	-	0.89	0.69	0.52	0.42	0.87	0.00	0.11	0.00	-	-
8	-	0.82	0.86	0.20	0.77	0.95	0.85	0.94	0.70	0.87	0.85	0.90	-
9	0.63	0.93	0.41	0.69	0.41	0.00	-	0.91	0.57	0.55	0.77	0.31	0.32
10	0.03	0.00	0.44	0.70	0.59	0.00	0.00	0.89	0.65	0.70	0.73	0.00	0.33
11	0.85	0.92	0.93	0.95	-	-	-	0.94	0.95	0.00	-	-	-
12	0.00	0.92	0.85	0.85	0.75	0.64	0.00	0.94	0.00	0.58	-	-	-
13	-	0.31	0.88	0.00	0.90	0.68	-	0.92	0.96	0.00	0.00	0.43	0.00
14	-	0.93	-	0.94	0.74	-	-	0.96	-	-	-	-	-
15	-	0.96	0.96	0.73	0.69	0.65	0.91	0.96	0.96	0.96	0.97	0.79	-
16	-	0.94	0.77	0.37	0.57	-	-	0.92	0.00	0.88	0.22	0.86	0.85
17	-	0.62	-	-	0.79	0.67	0.20	0.94	-	-	-	-	-
18	-	0.93	0.46	0.39	0.14	0.00	-	0.94	0.00	0.65	0.00	-	-
19	-	0.84	0.94	0.40	0.81	0.72	0.47	0.92	0.63	0.69	-	-	-
20	0.68	0.65	0.73	0.76	-	0.00	0.84	0.96	0.88	-	-	-	-
mean	0.38	0.71	0.76	0.58	0.63	0.40	0.33	0.93	0.54	0.57	0.52	0.67	0.19

There are classes (like CI class) that are almost "all or nothing" regarding the F1-score. The F1-score is above 0.90 or 0.00 for most of the wells, and this is due to the lesser amount of samples of this class presented in the wells.

Figure 7 shows the best test well (id 15) for the multi-class case. The trained model predicts all classes in the true well data. Also, it predicts two classes which are not in the test well data. It is worth mentioning that the I3 F1-score for this test well is only 0.65, but it is clear from Figure 7 that the I3 layer prediction is a very good one (dark green). The F1-score in this case is heavily penalized due to the fact that the I3 layer in the test well data starts a bit earlier than the prediction.

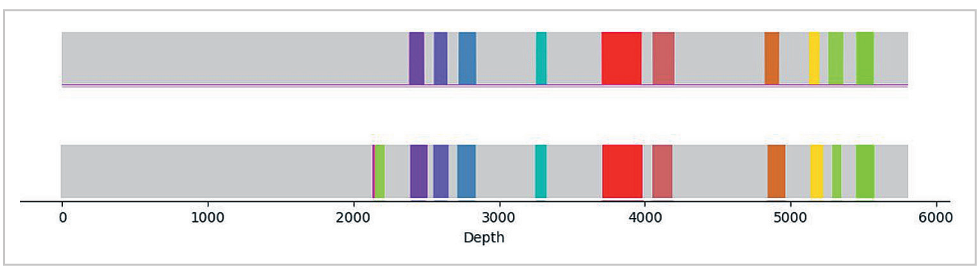

Figure 7 - True data for well 15 (above) and the model predictions (below). Selected from visual inspection of the outcomes as an example of good prediction.

Figure 8 shows the test well with the worst prediction performance (well 5). Although the model correctly predicts when the coal layers appear in the test well data, the classes are mismatched. Additionally, the model incorrectly predicts three layers which are not present in the test well data.

A poor prediction for a given well may have one or more causes: the behaviour of the variables (resistivity and natural gamma) for well 5 may be unique and the training dataset may not have a similar pattern. Another possibility is that well 5 has a more subtle pattern or noiser data. Thus, the chosen FCN parameters may not detect it in this case and more data may be needed to capture this information. Finally, it is expected that a multi-class version performs worse than the binary version on the same dataset because the amount of data available to train each class is diminished. Discussing again the accuracy score, for the worst well, 5, the accuracy is 0.78. This value does not reflect the fact that almost no coal classes were correctly predicted by the model. The F1-score better represents the model's ability to predict.

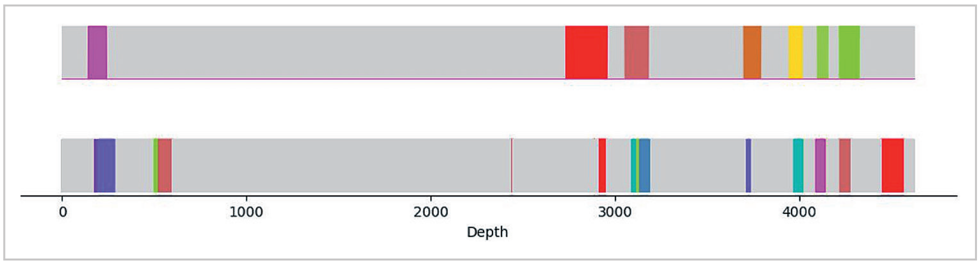

Figure 8 - True data for well 5 (above) and the model predictions (below). This example is considered the worst. Although the model correctly predicts some coal layers, the labels are not correct.

4. Discussion and conclusion

In this study, a Fully Convolutional Network was built to perform semantic segmentation (binary prediction coal or not coal, or multi-class prediction of coal lithologies) using two variables from well logs: the resistivity and the natural gamma.

The results show good performance for the binary ("coal" and "not coal") prediction, with a mean F1-score of 0.79. For the multi-class modelling, the results vary from an F1-score of 0.76 for the CS coal class to an F1-score of 0.19 for the S6 coal class.

For practical use, both trained models can be used: the binary trained model may establish where the coal layers are and the multi-class trained model may provide expectations for the classes of the layers, especially for the classes with highest score, such as CI and CS.

For further development, there is the idea to explore other neural network designs, adjust the loss function to add more weight to the correct and incorrect predictions of the class when training. Another area to explore is to find a way to encode spatial variables in the input. In our study, the depth is available, but it is not used in the input because the neural network may memorize the outputs to specific depths. A clever way to encode depth may be relative distances, and a well encoded depth may play an important role for angled drill holes. Another idea of spatial encoding in the training is to use as input for a given well the data from the wells in the neighborhood. The use of other measured variables is also a good object further investigation.

Other information that may be encoded in the modelling is the order of the layers with respect to the depth. Seismic data is another interesting field to apply semantic segmentation with Fully Convolutional Networks for classification tasks, although FCN may also be used for regression.

References

ANTARIKSA, Gian; MUAMMAR, Radhi; LEE, Jihwan. Performance evaluation of machine learning-based classification with rock-physics analysis of geological lithofacies in Tarakan Basin, Indonesia. ***Journal of Petroleum Science and Engineering,*** v. 208, 109250, 2022. DOI: 10.1016/j.petrol.2021.109250

ASFAHANI, J.; BORSARU, M. Low-activity spectrometric gamma-ray logging technique for delineation of coal/rock interfaces in dry blast holes. ***Applied Radiation and Isotopes,*** v. 65, n. 6, p. 748-755, 2007. DOI: 10.1016/j.apradiso.2007.01.019

BRENNING, Alexander. Spatial machine-learning model diagnostics: a model-agnostic distance-based approach. ***International Journal of Geographical Information Science,*** v. 37, n. 3, p. 584-606, out. 2022. DOI: 10.1080/13658816.2022.2131789

DENG, Xiaolong; LI, Lihui; TAN, Yufang. Validation of spatial prediction models for landslide susceptibility mapping by considering structural similarity. ***ISPRS International Journal of Geo-Information,*** v. 6, n. 4, mar. 2017. DOI: 10.3390/ijgi6040103.

DEV, Vikrant A.; EDEN, Mario R. Formation lithology classification using scalable gradient boosted decision trees. ***Computers & Chemical Engineering,*** v. 128, p. 392-404, 2019. DOI: 10.1016/j.compchemeng.2019.06.001

FUKUSHIMA, Kunihiko. Neocognitron: a hierarchical neural network capable of visual pattern recognition. ***Neural Networks,*** v. 1, p. 119-130, 1988. DOI: 10.1016/0893-6080(88)90014-7

HOFFMAN, G. L.; JORDAN, G. R.; WALLIS, G. R. ***Geophysical borehole logging handbook for coal exploration.*** Canada: Coal Mining Research Centre, 1982. ISBN: 978-0969104803.

IMAMVERDIYEV, Yadigar; SUKHOSTAT, Lyudmila. Lithological facies classification using deep convolutional neural network. ***Journal of Petroleum Science and Engineering,*** v. 174, p. 216-228, mar. 2019. DOI: 10.1016/j.petrol.2018.11.023

LARY, David J.; ALAVI, Amir H.; GANDOMI, Amir H.; WALKER, Annette L. Machine learning in geosciences and remote sensing. ***Geoscience Frontiers,*** v. 7, n. 1, p. 3-10, 2016. DOI: 10.1016/j.gsf.2015.07.003

PLOTON, Pierre *et al.* Spatial validation reveals poor predictive performance of large-scale ecological mapping models. ***Nature Communications,*** v. 11, n. 1, set. 2020. DOI: 10.1038/s41467-020-18321-y.

RODRIGUES, Áttila. Fully convolutional network example application to coal layers. GitHub, 2022. Available in: <https://github.com/attilalr/fully-conv-network-1d-coal>. Access at: 16 maio 2022.

SALEHI, Seyyed Mohsen; HONARVAR, Bizhan. Automatic identification of formation lithology from well log data: a machine learning approach. ***Journal of Petroleum Science Research,*** v. 3, n. 2, p. 73-82, 2014. DOI: 10.14355/JPSR.2014.0302.04

SCHMIDHUBER, Jürgen. Deep learning in neural networks: an overview. ***Neural Networks,*** v. 61, p. 85-117, 2015. DOI: 10.1016/j.neunet.2014.09.003

SHELHAMER, Evan; LONG, Jonathan; DARRELL, Trevor. Fully convolutional networks for semantic segmentation. ***IEEE Transactions on Pattern Analysis and Machine Intelligence,*** v. 39, n. 4, p. 640-651, 2017. DOI: 10.1109/TPAMI.2016.2572683

VALAVI, Roozbeh *et al*. BlockCV: an R package for generating spatially or environmentally separated folds for k-fold cross-validation of species distribution models. ***Methods in Ecology and Evolution,*** v. 10, n. 2, p. 225-232, 2018. DOI: 10.1111/2041-210X.13107.

WANG, Yanwen; KHODADADZADEH, Mahdi; ZURITA-MILLA, Raúl. Spatial+: a new cross-validation method to evaluate geospatial machine learning models. ***International Journal of Applied Earth Observation and Geoinformation,*** v. 121, jul., 2023. DOI: 10.1016/j.jag.2023.103364

ZHU, Kai *et al*. DeepLog: identify tight gas reservoir using multi-log signals by a fully convolutional network. ***IEEE Geoscience and Remote Sensing Letters,*** v. 17, n. 4, p. 568-571, 2020. DOI: 10.1109/LGRS.2019.2930587

Received: 22 August 2023 - Accepted: 24 February 2024.