ANALYSIS OF CLASTIC SEDIMENTARY ROCKS USING IMAGE SEMANTIC SEGMENTATION: EXPLORING A COMPUTATIONAL PROCESSING METHOD IN PETROGRAPHIC STUDIES

Julián Bardelli[1], Pablo Joaquín Alonso-Muruaga[2], Marisol Suárez Cruz[1]

1: Inlab S.A. jbardelli@inlab.com.ar

2: IGEBA (CONICET-UBA). pablojoaquin3@gmail.com

Palabras clave: photomicrographs, semantic segmentatation, petrography

ABSTRACT

Las imágenes de sección delgada representan una valiosa fuente de información para los estudios petrográficos aplicados. En este marco, se analiza un método de procesamiento computacional basado en redes neuronales artificiales para la obtención de información cualitativa y cuantitativa a partir de microfotografías de cortes de delgados en muestras areniscas. El método utilizado incluye dos arquitecturas principales de segmentación semántica, correspondientes a los modelos U-net y DeeplabV3+. El primero esta basado en redes neuronales totalmente convolucionales y una estructura de codificador-decodificador, mientras que el segundo incluye la convolución de Atrous, la agrupación de pirámide espacial y un módulo decodificador simple. Los modelos fueron entrenados con imágenes que incluyen por muestra un par de fotos con luz polarizada paralela y cruzada junto a una "máscara" que diferencia los píxeles correspondientes a cada clase. Las clases corresponden a los componentes principales de arenisca, incluyendo cuarzo, feldespatos, líticas y "de fondo" (incluyendo poro y cemento), y fueron diferenciadas a partir del proceso de segmentación semántica. El modelo entrenado resultante permitió el análisis de métricas evaluadas en muestras de prueba y comparar los resultados con la técnica estándar de conteo de puntos. El análisis indica que el método de procesamiento puede ser aplicado, brindando preliminarmente datos cualitativos y cuantitativos generales de la sección delgada en estudio. Además, el modelo puede mejorarse sustancialmente en función de su entrenamiento con datos nuevos y complementarios, y por lo tanto, refinar el método analítico.

INTRODUCTION

Thin section sedimentary rock images represent a valuable source of information for applied petrographic studies in oil exploration and reservoir characterization. Computational methods generated in order to extract component data from photomicrographs were tested by previous authors (eg. Ładniak and Młynarczuk, 2015; Budennyy *et al.*, 2017; Izadi *et al.*, 2017, Rubo *et al.*, 2017) arriving to the conclusion that automatic quantitative and qualitative analysis can be suitable for image processing. Therefore a complementary analysis tool that favors thin section

description and interpretation of sedimentary rock main components, including grain composition, might provide valuable assistance to geoscientists in the process of petrographic analysis. In this framework, the present contribution explores and test a computational processing method based on convolutional neural networks (CNN) in order to analyze the feasibility of its application in obtaining qualitative and quantitative petrographic data from thin section photographs.

Consequently, a model was trained with a pair of images (Parallel Polarized Light PPL and Crossed Polarized Light XPL) and its mask indicating the component (e.g. mineral) corresponding to each pixel. Once trained, the model can be directly fed with a pair of images (PPL and XPL). The resulting image analysis is characterized by a 2D array of values, each one indicating the component recognized per pixel. This process of classifying an image in a pixel by pixel fashion is usually known as Semantic Segmentation. Finally, the model was mainly trained to work with sandstone images, and the components selected to be identified were Feldspar, Lithic Fragments and Quartz, in order to relate the results with the arenite ternary diagram of Folk´s (1970) classification.

MATERIALS

Materials included sets of images of twenty-seven randomly selected thin sections of sandstones (arenites). The selection comprises arenites with different composition, diagenetic evolution and grain size for better model generalization. Each set of images consists of a pair (PPL and XPL) of 1600 by 1200 pixel photomicrographs together with its respective mask image obtained from a clast rich area in a given thin section (Figure 1) per sample. Regarding the generation of the masks, a one-hot-encoded format representing a 2D array of values per pixel was used. In this way each pixel of the image was assigned to a value ranging from 0 to 3, which in turn represent the class or component type. Then, the image is annotated by a geologist by mapping the different classes using an annotation tool specific for Semantic Segmentation (eg. Apeer software). The component types consist of: Background (0) including background typified by matrix, cement and pore space; Feldspars (F); Lithic Fragments (L); and Quartz (Q). Therefore, the resulting output of an image analysis based on the convolutional neural network (CNN) herein generated correspond to a 2D array (1600 x 1200 in size) where each value represents the main selected classes of the pixel in the input image (see results section). The hardware used was a Desktop PC with a Rayzen 7 CPU, 32 Gb of RAM and a GPU with 12Gb memory. The program was made in Python programming language and the deep learning framework used was Tensorflow. Each pixel in the XPL image corresponded to the same pixel position in the PPL image, meaning there was no rotation between them, only a polarizer was added in between.

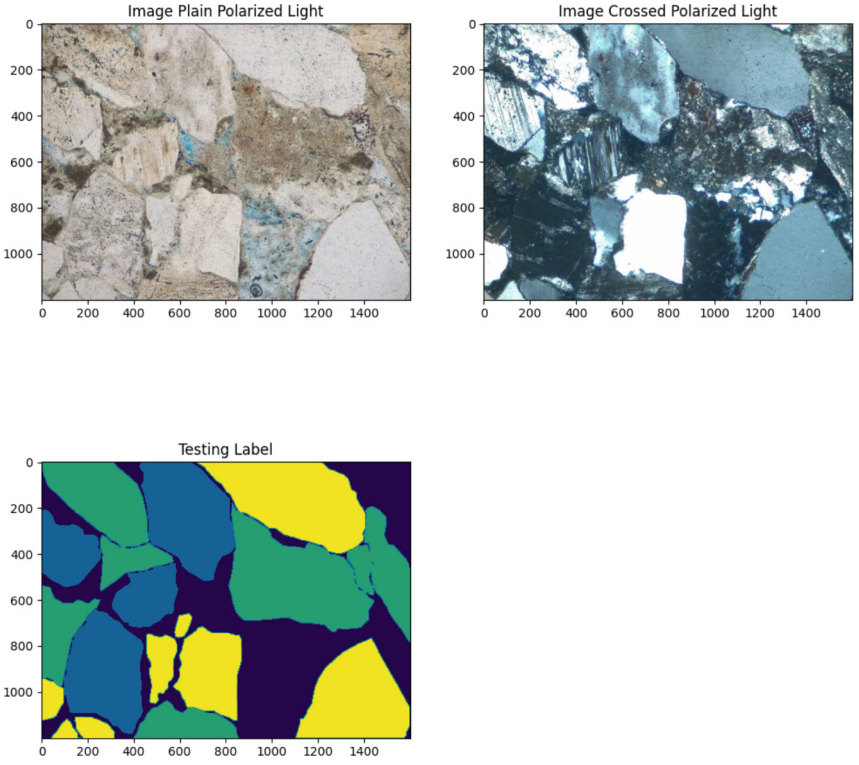

Figure 1. Example of a pair of Plain Polarized Light image and a Crossed Polarized Light image, together with the annotated mask by the geologist, where yellow is Quartz, blue is Feldspars, green is Lithic Fragments and dark violet is background.

CONSTRAIN IN MODEL GENERATION AN ANALYSIS

It is necessary to consider some limitations while acquiring data to feed and train the model, and the possible results. In this way it is important to mention that image annotation by a geologist is a very time and resources consuming task. A main limiting factor in our evaluation was the amount of annotated images available, which depending on the objectives and the resolution may reach hundreds in order to generate an good image segmentation model. In addition, identification of the component types (eg. minerals) from a photomicrograph is a very difficult task compared to, for example, segmenting cars or people from a street image Therefore, considering that the scope of this work is limited to assess the feasibility of a semantic segmentation model to classify the main minerals from a photomicrograph, it was difficult to expect good segmentation results. However once the models are trained with the available data and tested, it is possible to select the exemplar that can be improved with more training images. In this manner it is possible to generate a model that could consistently classify the selected component types and assist the petrographic analysis process.

CONVOLUTIONAL NEURAL NETWORK ARCHITECTURES APPLIED IN IMAGE ANALYSIS

Two semantic segmentation architectures were used, called Unet (Ronneberger *et al.*, 2015) and the DeeplabV3+ (Chen *et al.*, 2018) models, shown in Figure 2 and Figure 3 respectively. Unet architecture is based on fully convolutional neural networks and encoder-decoder structure, while Deeplab relies on Atrous convolution, Spatial Pyramid pooling, and a simple decoder module. Both architectures proved to have a good track record on public benchmark datasets and have been recently applied to petrographic images (eg. Rubo *et al.*, 2019; Wang *et al.*, 2021). Many of the innovations included in both of them are targeted at detecting small scale details without losing the large scale characteristics of the objects. These convolutional neural networks are fairly complex in the amount of layers used and hence they have a good number of trainable parameters, meaning they are memory and GPU intensive. For this reason, there is usually a limit in the image resolution for the input layer, because if the input image is too big, the computer resources could run out during training. In our case after some trial and error, we used a 400 x 400 pixel input image that could be managed with the hardware resources available. This meant that the original image of 1600 x 1200 pixels had to be divided in 12 patches to be used for training and model prediction. During prediction, composing a mosaic of images predicted separately can also generate artifacts in the zones where they are joined together. Using a 400 x 400 input layer resulted in 1.9 million trainable parameters for Unet and 3.5 million for Deeplab V3+.

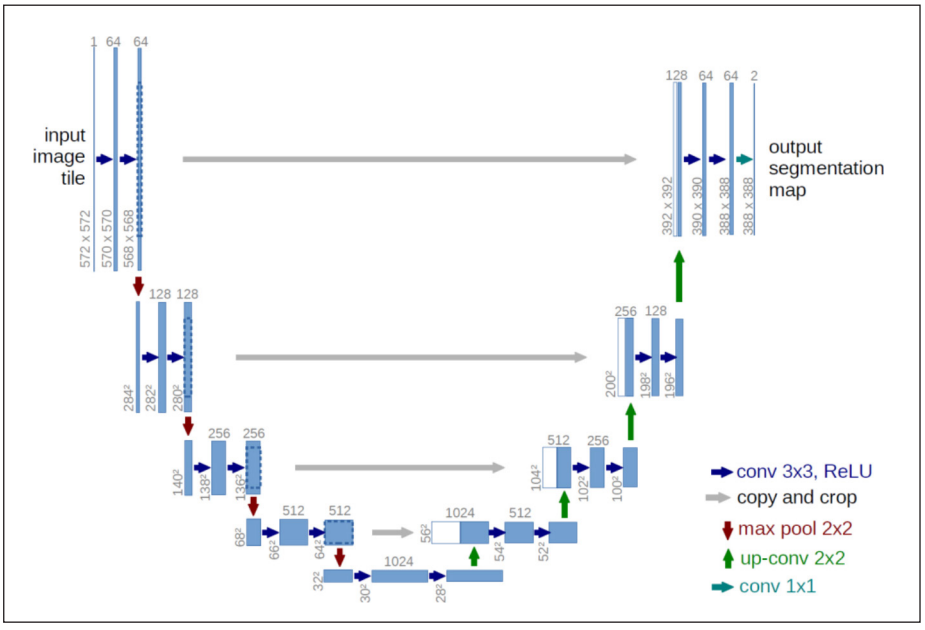

Figure 2. Unet architecture.

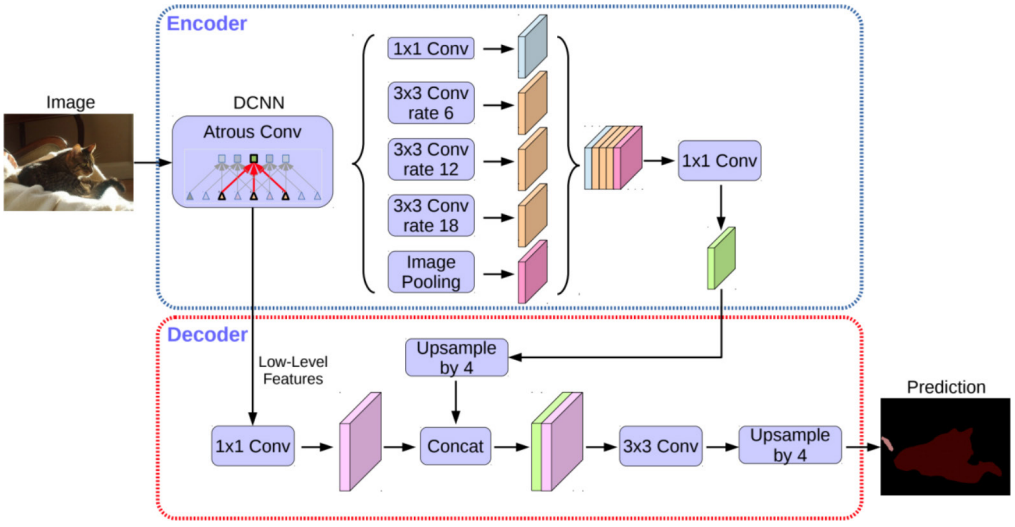

Figure 3. Deeplab V3+ architecture.

MODEL EVALUATION AND COMPARISON

In order to evaluate the generated models, an internal model consistency test and a cross validation process were carried out. In the former all the images were used both in training and evaluation. While in the later different models were trained, and in each one of them, three sets of images were left out of the training and used for evaluation. In this way, it was possible to primarily check if at least a trained model with all the images could identify minerals with a good result in the same images used for training. If this was not the case, then a cross validation test would surely give poor results. In addition a primary check allowed for the comparison between the two architectures used, and for the identification of significant differences in the results.

- **Internal consistency test**

The internal consistency test was intended to evaluate if it is possible to predict the classes of an image that was used in the training process. Thus, after some trial and error, we concluded that training for 100 epochs for the Unet architecture and 60 epochs for Deeplab gave the best results, avoiding overfitting at the same time. Training accuracy and loss is shown in Figure 4.

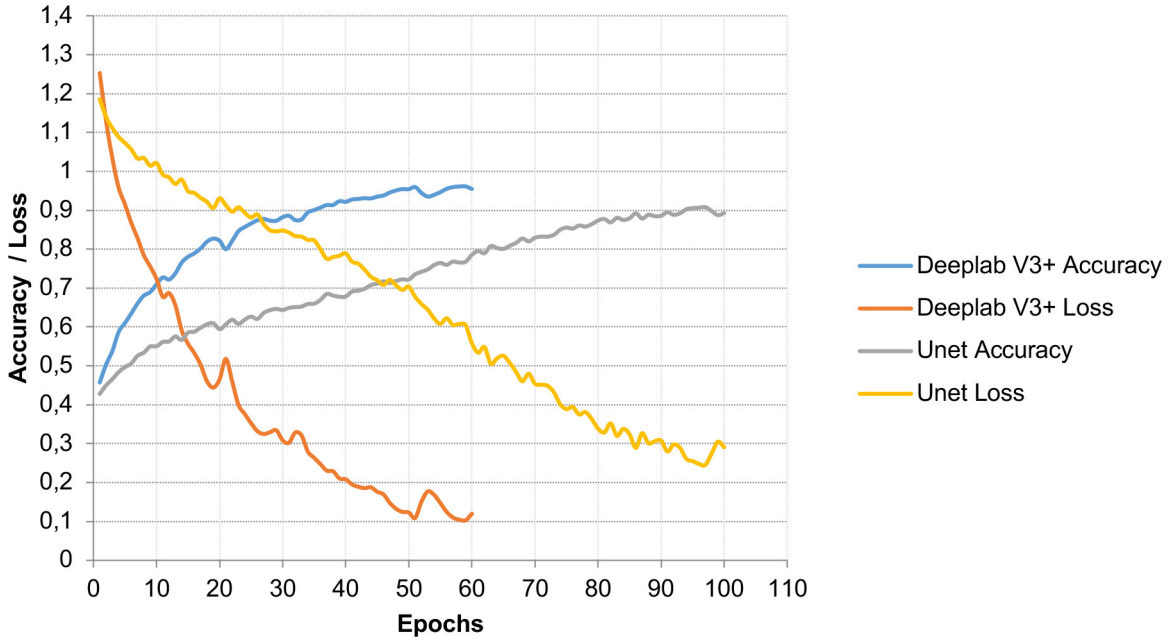

Figure 4. Training accuracy and loss for both architectures.

Once both models (Unet and Deeplab) were trained, all the images were used for prediction, and the average results are shown in Table 1.

Architecture	Train images	Epochs	Avg Accuracy	Folk Class. Accuracy	Avg Error Feldspar	Avg Error Lithic Fragments	Avg Error Quartz
Unet	27	100	81.9	85.2	8.0	5.6	7.0
Deeplab V3+	27	60	79.1	77.8	8.4	8.3	3.0

Table 1. Results of internal evaluation of each model.

The next stage in the model analysis consisted in linking the results to the Folk´s (1970) arenite ternary diagram (Figure 5) in order to test a preliminary classification. In this way an image was fed into the trained model, and the main components (F, L, and Q) percentages were obtained as output. The percentage of each mineral was calculated in relation to the total minerals detected, not counting the background pixels. Then, the percentages of each minerals from both the original mask and from the model prediction are entered into a Folk ternary plot in order to check if both data points are at least in the same class (region) of the Folk´s diagram. If it is the situation, the classification is counted as a success. Finally, the classification accuracy is calculated as the total of correctly classified samples divided by the total images used multiplied by 100.

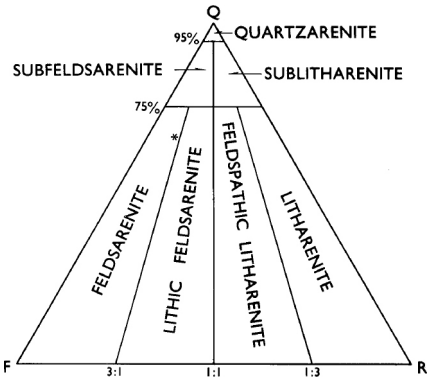

Figure 5. Folk classification ternary diagram (Folk *et al.* 1970).

The results show that Unet training converges slower than Deeplab, but has slightly better results in accuracy and Folk classification. The following figures show an example of the same image predicted by both Unet (Figure 6) and Deeplab (Figure 7). In both models this image resulted in an accuracy of around 80% which is near the average for both models.

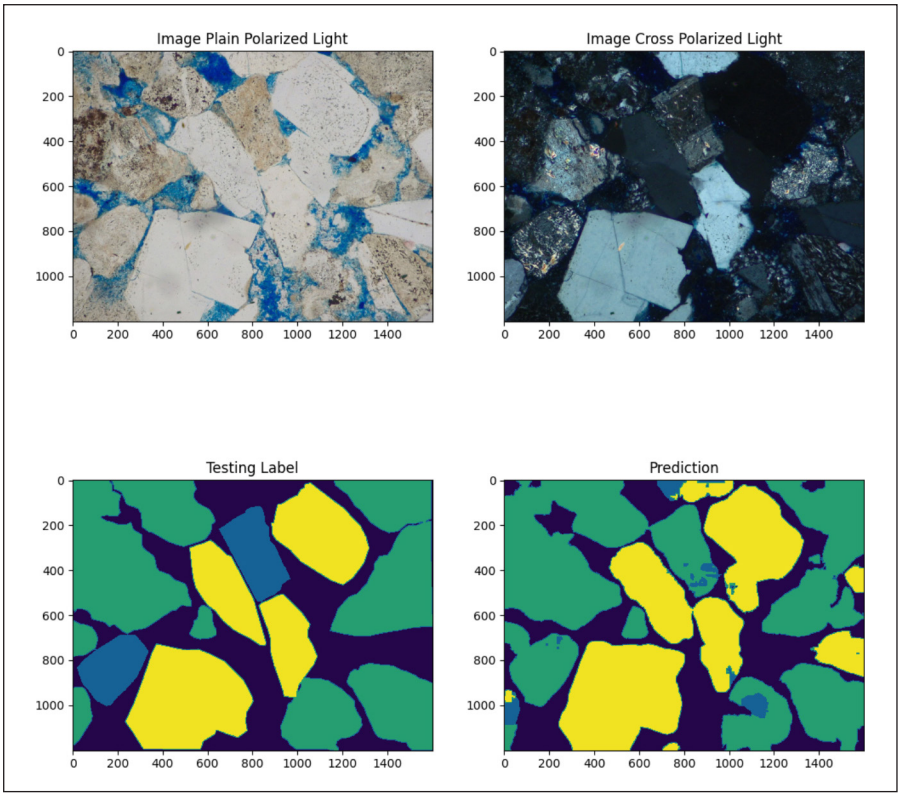

Figure 6. Prediction of Unet on one image (8b) used in the training process.

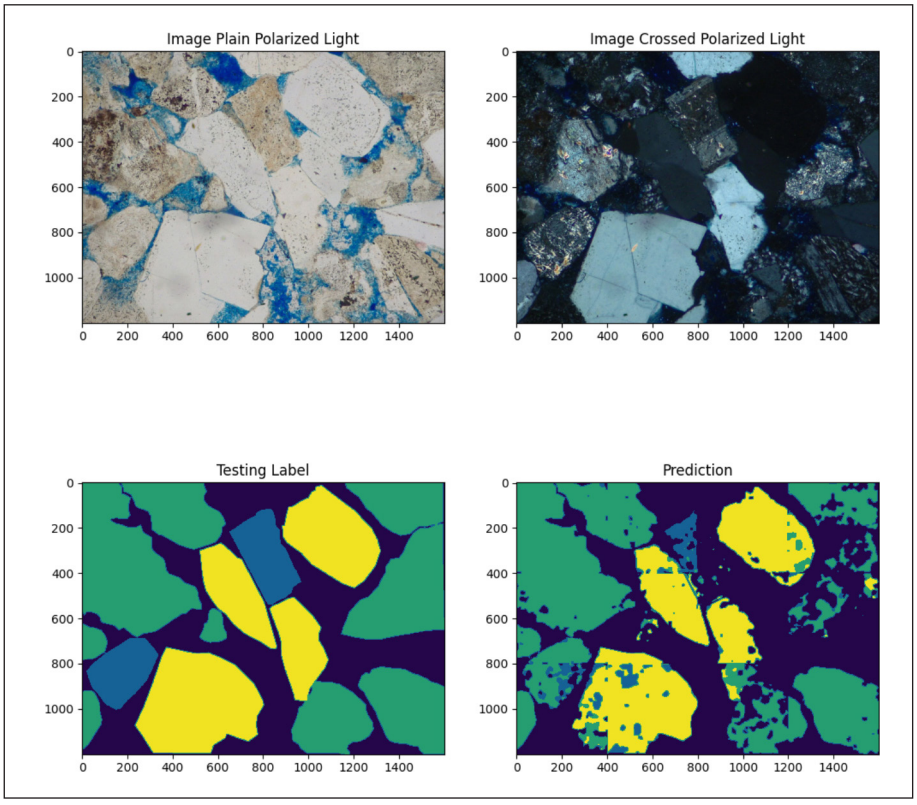

Figure 7. Prediction of Deeplab on one image (8b) used in the training process.

Cross Validation Tests

After not perfect but acceptable results in the initial consistency test (see Figures 6 and 7), we needed a more realistic way of testing the model accuracy with images that it has not seen before during training. To do this, and based on the annotated images we had available (twenty-seven set of images), a cross validation test was carried out. This process mainly consisted of training nine models. In each one of them, twenty-four sets of images were used for training, and three images (photomicrographs) were fed for testing (prediction). We rotated the images so all the images were used for testing in the nine models.

Then a prediction and accuracy evaluation was made with the three testing images with its corresponding model. The Table 2 shows the images used for testing in each model, whereas the summary of testing results of the cross validation test is shown in Table 3. A comparison of the accuracy for each architecture and model pair is shown in Figure 8. In addition Figure 9 and Figure 10 show examples of the same image with its prediction with each of the CNN architectures used. The test also shows that classification accuracy results are lower than in the internal validation, which is expected as the testing images were not part of the training set.

Model	Test Image 1	Test Image 2	Test Image 3
Cross_validation_model1	10a	11a	2a
Cross_validation_model2	2b	3a	3b
Cross_validation_model3	4a	4b	5a
Cross_validation_model4	5b	6a	6b
Cross_validation_model5	7a	7b	8a
Cross_validation_model6	8b	9a	9b
Cross_validation_model7	image006	image008	image039
Cross_validation_model8	image047	image053	image055
Cross_validation_model9	image057	image061	image065

Table 2. Cross validation images used in each model.

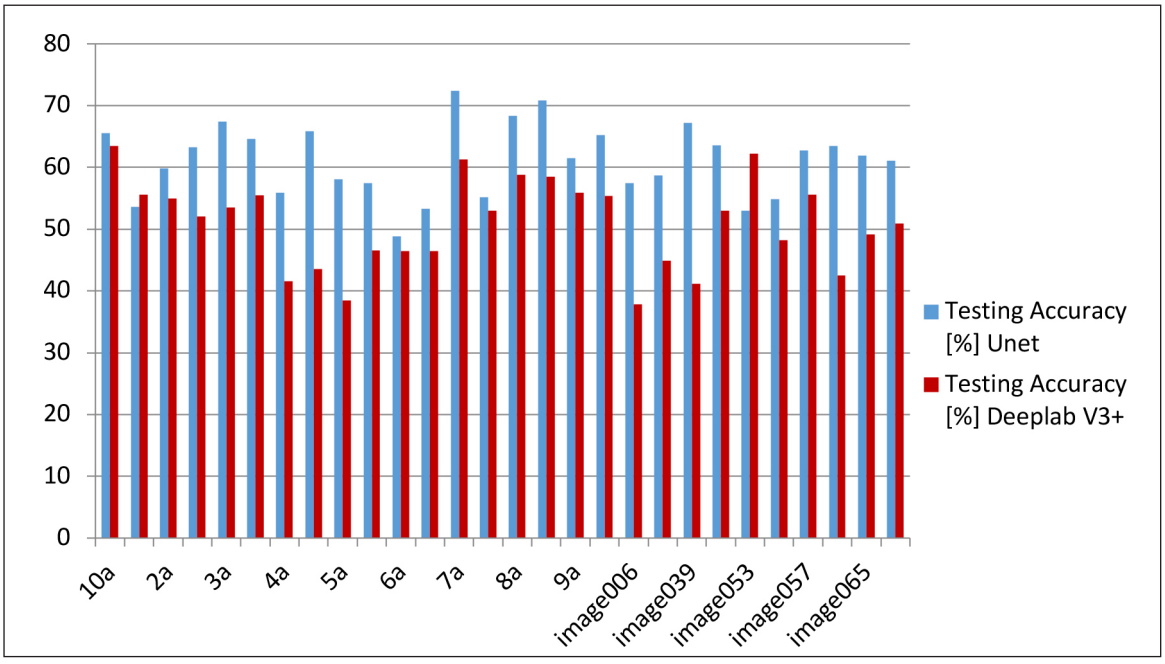

Figure 8. Testing accuracy for both architectures for each testing image.

Architecture	Train images	Epochs	Avg Accuracy [%]	Folk Class. Accuracy [%]	Avg Error Feldspar [%]	Avg Error Lithic Fragments [%]	Avg Error Quartz [%]
Unet	24	100	61.1	63.0	14.8	9.4	8.8
DeeplabV3+	24	60	50.9	66.7	14.6	13.0	11.4

Table 3. Cross validation summary

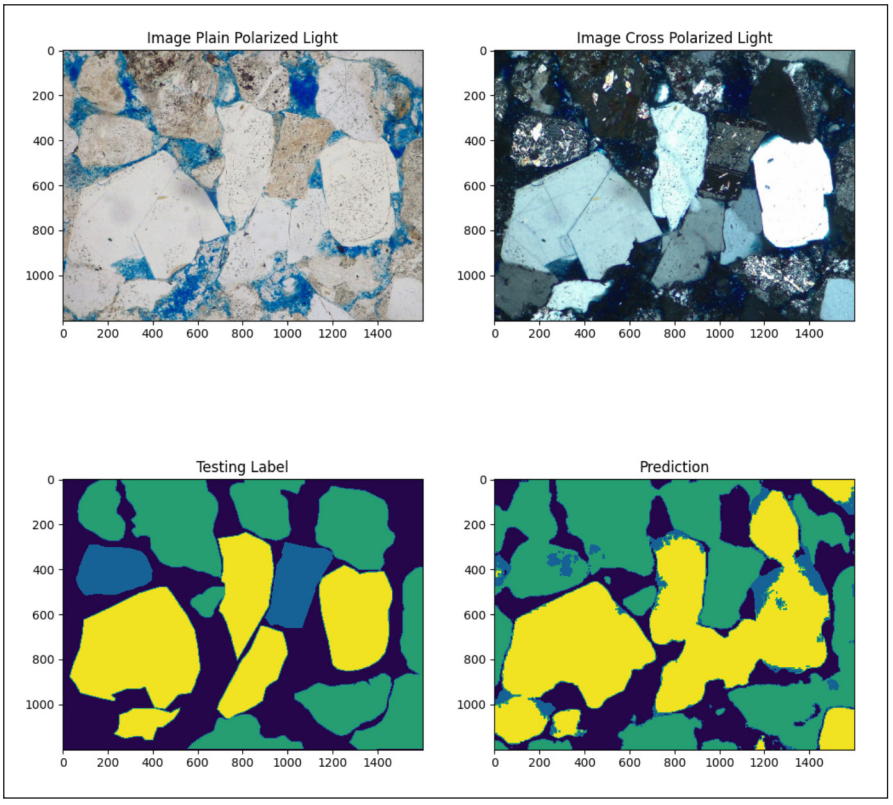

Figure 9. Unet prediction on image 8a with validation model 5 where the accuracy is near the average.

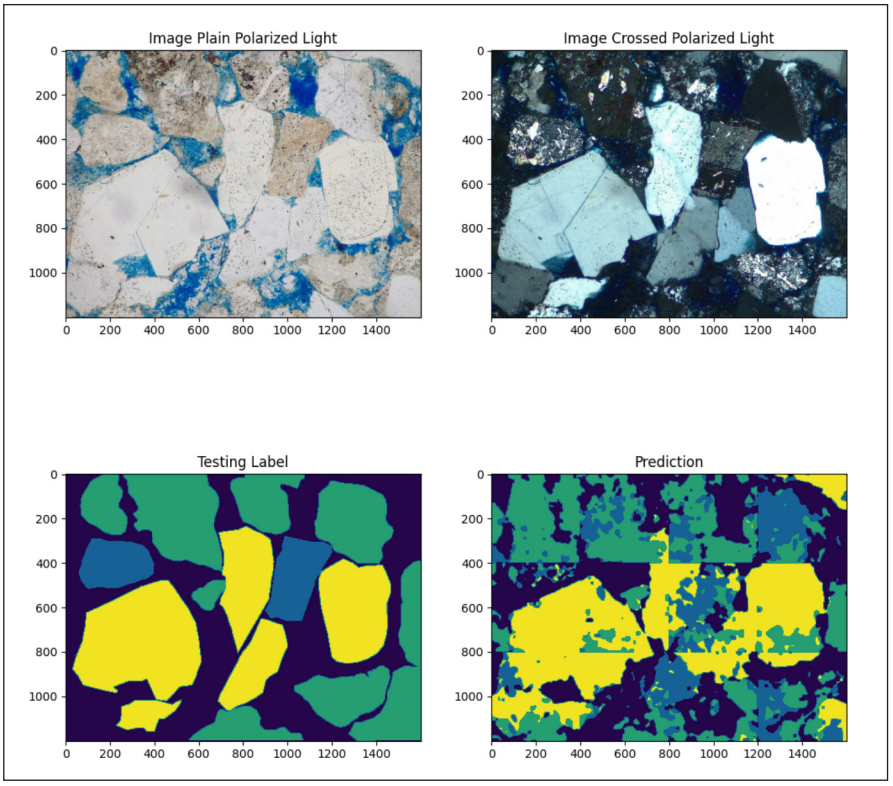

Figure 10 . Deeplab prediction on image 8a with validation model 5 where the accuracy is near the average.

DISCUSSION

Our results show that Unet had better results in all the tests compared to Deeplab. However the latter should not be discarded as a possible architecture, since it has performed well in similar petrographic segmentation tasks (3).

Testing accuracy (61% pixel accuracy for Unet and 50.9% for Deeplab) is low for an acceptable segmentation task and do not allow for these models to be used in a production environment. The low performance observed was somewhat to be expected considering the relatively low amount of images used in this work. However, the results are not bad enough to rule out this method as a way to quantify main components in sandstones. In fact, the models in some cases successfully detected clasts not annotated by the geologists in the mask.

In this regard it is possible to enhance the accuracy of these models by significantly increasing the amount of annotated images used for training and testing. Adding more cross polarized light images at different polarizer angles, multiplying the amount of channels of the input layer of the CNN, and performing modifications to the CNN architectures that could increase model segmentation performance. In addition, results can be enhanced if the image under analysis correspond or is related to the same source data from which the models were trained. This means that the models can be fed with specific data corresponding to certain geological area or litostratigraphical unit, which in turn might optimize the recognition of particular features of the related samples.

Overall the study shows the potential that these image methods have in assisting the geologist performing petrographical analysis. The models can be adequately trained, and also prepare them to analyze other features such as pores and possible cement type. However it is important consider for future research, once the performance of CNN models is increased, that the results should be compared to humans performing the same tasks. To be accepted, those results should fall between or at least closely approach the uncertainty levels of humans.

CONCLUSIONS

This processing method can assist the geoscientist performing the petrographic analysis, bringing preliminary qualitative and quantitative data of the analyzed rock. Under the expert eye, this information can be also refined; contributing to optimize the time used describing, characterizing and classifying clastic sedimentary rocks. In addition, the model can be substantially improved based on its training with new and supplementary data, and therefore refining the analytic method.

REFERENCES

Budennyy S., Pachezhertsev A., Bukharev A., Erofeev A, Mitrushkin D., 2017, Image Processing and Machine Learning Approaches for Petrographic Thin Section Analysis. SPE Russian Petroleum Technology Conference. https://doi.org/10.2118/187885-RU.

Chen LC., Zhu Y., Papandreou G., Schroff F., Hartwig A., 2018. Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation. https://arxiv.org/pdf/1802.02611.pdf

Folk R. L., Andrews P. B., Lewis D. B., 1970. Detrital sedimentary rock classification and nomenclature for use in New Zealand, New Zealand Journal of Geology and Geophysics, 13:4, 937-968, DOI: 10.1080/00288306.1970.10418211

Iwaszenko S., Róg L., 2021. Application of Deep Learning in Petrographic Coal Images Segmentation. https://www.mdpi.com/2075-163X/11/11/1265/pdf

Izadi H., Sadri J., Bayati M., 2017. An intelligent system for mineral identification in thin sections based on a cascade approach. Computers and Geosciences 99, 37-49.

Ładniak M, Młynarczuk M., 2015. Search of visually similar microscopic rock images. Comput. Geosci, 19:127-136. doi: 10.1007/s10596-014-9459-2

Ronneberger O., Fischer P., Brox T., 2015. U-Net: Convolutional Networks for Biomedical Image Segmentation. https://arxiv.org/abs/1505.04597

Rubo A. R., Carvalho Carneiro C., Fontana Michelon F., dos Santos Gioria R., 2019, Digital petrography: Mineralogy and porosity identification using machine learning algorithms in petrographic thin section images, Journal of Petroleum Science and Engineering 183, 1-14.

Wang Y., Bai X, Wu L, Zhang Y, Qu S., 2021. Identification of maceral groups in Chinese bituminous coals based on semantic segmentation models. Fuel 308, 1-12.