

# Mineral mapping tool using dual energy CT-scanner

Daniil Sherki, MSc-1, Petroleum engineering

January 27, 2023

Scientific advisors: Vladislav Alexeev, Denis Orlov, Ekaterina Muravleva,  
Dmitry Koroteev

## Abstract

Image rock segmentation is meaningful problem for digital rock  $\mu$ CT processing. Nowadays digital rock specialists use QEMSCAN technology which is time- and money-consuming method. However, Digital Petroleum specialists are able to automate processing  $\mu$ CT images of rocks using the convolution neural network (cNN) UNet, using  $\mu$ CT images in two different gas environments (air and xenon) as input data for Deep Learning model. The goal of this research is to try to use two different energy of CT-scanner to solve the problem of rock image segmentation without using xenon gas.

## 1 Introduction

Rock segmentation is a complex issue for digital rock  $\mu$ CT processing. Currently, digital petrophysicists are utilizing the QEMSCAN technology, which is both costly and time consuming by reason of human involvement.

To address this issue, Digital Petroleum researchers are exploring the application of a convolutional neural network (CNN) UNet to automate the segmentation of  $\mu$ CT images of rocks taken in two different gas environments - air and xenon - in [1].

There is hypothesis, that in this case can be used dual energy of CT-scanner image instead of images in air and xenon. It can be useful for many researchers, because in this case they don't need to use xenon. It will be enough to change the emitting energy of the tomograph.

This research's purposes is:

1. To investigate whether the case of double energies is applicable to the case of segmentation of core samples.
2. To compare this predictions with predictions on only one energy  $\mu$ CT images as input data.
3. To develop tool for 3D-visualizstion rock sample with  $\mu$ CT images.

## 2 Theoretical part

### 2.1 Basic terms

Basic terms for this research article:

**Augmentation** is a method of increasing training data with modification original training data copies (horizontal flip, vertical flip, rotation by some small angle and others).

**The batch size** is the amount of samples taken from the training dataset that are processed together before the model is updated. An epoch is a single pass through the entire dataset. The batch size must be a positive integer, and can't be bigger than the amount of training samples.

**Adam optimizer** is an optimization algorithm which combines the advantages of RMSProp and SGD with Momentum to adaptively adjust the learning rate. It is suitable for non-stationary objectives and problems that have noisy or sparse gradients. It's show himself well for segmentation image problem.

The weights update according to

$$\theta_t = \theta_{t-1} - \eta \frac{\hat{m}_t}{\sqrt{\hat{v}_t} + \epsilon}$$

$$\hat{m}_t = \frac{m_t}{1 - \beta_1^t}$$

$$\hat{v}_t = \frac{v_t}{1 - \beta_2^t}$$

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) g_t$$

$$v_t = \beta_2 v_{t-1} + (1 - \beta_2) g_t$$

where  $\eta$  is learning rate,  $\eta \approx 10^{-8}$  is small number to prevent dividing by zero,  $\beta_1 \approx 0.9$  and  $\beta_2 \approx 0.999$  are forgetting parameters.

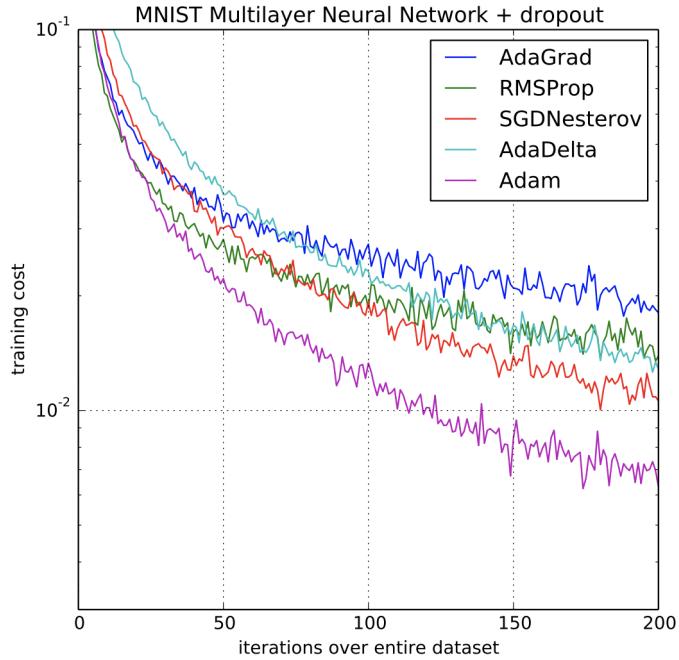


Figure 1: Optimizers comparing [2]

**Learning Rate Scheduler** is tool which allow to reduce the learning rate as the number of training epochs increases. Step decay is usually favored over other methods due to its straightforward interpretation of hyperparameters such as the fraction of decay and step timings in epochs. Additionally, step decay is successful in stabilizing the learning rate which helps the optimization process of stochastic gradient descent to quickly converge and achieve successful results.

**Argmax** returns the indices of the maximum values along an axis of array.

**Semantic segmentation** is a pixel-based classification problem, which identify each pixel as particular class. There are many neural networks and semantic algorithms which used for solve this problem such as U-net, Mask R-CNN as well as Feature Pyramid Network (FPN).

**U-net** is fully convolution neural network using for image segmentation problem. It is combination of encoder network and followed by decoder network. Encoders if convolution with  $3 \times 3$  kernel and ReLU-function. Technically, in this research U-net used such as convolution algorithm with black-bone ResNet-50.

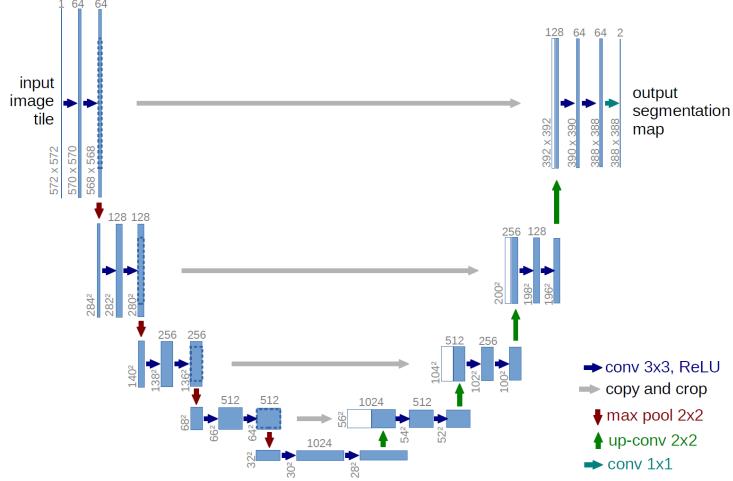


Figure 2: U-net architecture [4]

**ResNet-50** is a convolutional neural network that has 50 layers of depth. Researchers can download a pre-trained version of the network trained on more than a million images from the ImageNet database.

**Cross-entropy loss**, or log loss, measures the performance of a classification model whose output is a probability value between 0 and 1. Cross-entropy loss increases as the predicted probability diverges from the actual label. So predicting a probability of .012 when the actual observation label is 1 would be bad and result in a high loss value. A perfect model would have a log loss of 0.

$$\text{CE}(p_t) = -\log(p_t)$$

where  $p_t$  is a probability of ground truth class.

**Focal loss** is a option of cross-entropy loss function modification, which can control value of penalty with  $\gamma$ -modulating parameter with tuning  $\gamma \geq 0$ . It used in Deep Learning, because in same points Cross-Entropy has bigger value of derivartion as compared to cross-entropy loss function (figure 3).

$$\text{FL}(p_t) = -(1 - p_t)^\gamma \log(p_t)$$

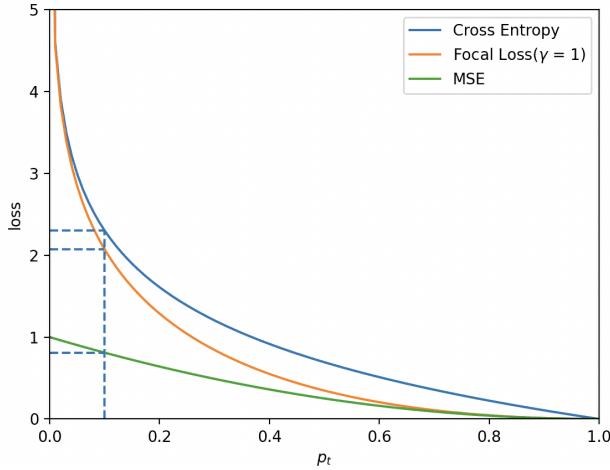


Figure 3: Loss-functions comparing

**Convolution** is the similar term in digital processing, but in Deep learning it is a mathematical operation that allows the merging of two sets of information. Convolution is applied to the input data to filter (kernel) the information and produce a feature map.

The **Intersection-Over-Union**(Jaccard Index) is an evaluation metric used to measure the accuracy of an object detector on a particular dataset and it is relation between area of overlap labeled image and predicted image and area of union of these two.

$$\text{IoU} = \frac{\text{Area of Overlap}}{\text{Area of Union}} = \frac{\text{TP}}{\text{TP} + \text{FP} + \text{FN}}$$

where TP - true positive, FP - false positive, FN - false negative.

**Dice metric**(Sørensen–Dice coefficient)

$$\text{Dice} = \frac{2 \times \text{Area of Overlap}}{\text{Total pixels combined}} = \frac{2 \times \text{TP}}{\text{TP} + \text{FP} + \text{TP} + \text{FN}}$$

## 2.2 QEMSCAN

QEMSCAN is the patent method for segmentation  $\mu$ CT scans into different minerals with certain software. This method combined two techniques. The first is SEM, the second is X-ray spectrometers (EDS) and the combined system measures backscattered electron and electron-induced secondary X-ray emission spectra.

Special software calculates mineral concentration in rock sample mineral composition and porosity value. QEMSCAN creates detailed statistic of mineral composition, color mask for each mineral and SEM images.

Manual pixel-by-pixel labeling will be very time-consuming operation. Thus, using QEMSCAN images as labels for neural network is good idea.

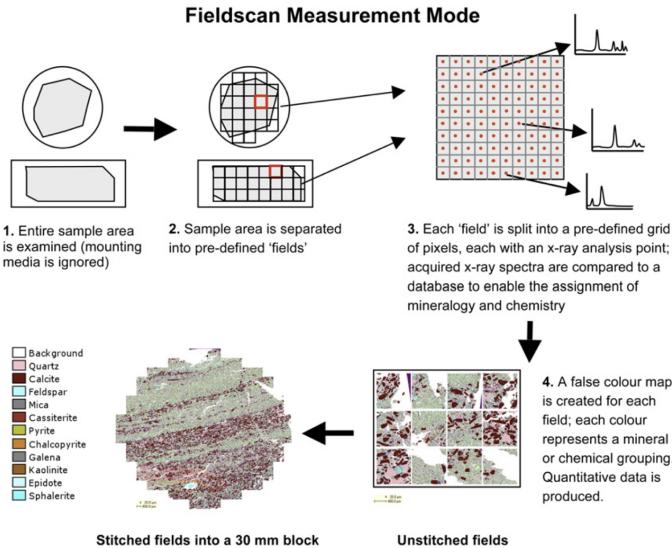


Figure 4: QEMSCAN algorithm [5]

### 2.3 U-net for air and xenon mineral mapping

U-Net has achieved excellent quality in many biomedical image segmentation tasks, including electron microscopy images. The encoding part corresponds to a typical convolutional network architecture. U-Net enhances SegNet by introducing skip-connections between the corresponding encoding and decoding layers through concatenation, which helps the neural network retain some high-level details. The result is higher overall pixel accuracy. It consists of repetitive encoder blocks, each consisting of two valid  $3 \times 3$  convolutions, followed by a rectified linear unit (ReLU) and a maximum  $2 \times 2$  merge operation (maxpooling) in steps of 2 for downsampling. Each encoder doubles the number of feature channels.

The decoder part of the U-network also consists of several decoder blocks. Each decoder consists of a  $2 \times 2$  upconversion layer that halves the number of feature channels and doubles the spatial resolution of each feature map, a concatenation with the corresponding trimmed feature map from the downsampling path, and two  $3 \times 3$  convolutions, each followed by a ReLU.

The cropping is necessary due to the loss of boundary pixels in each convolution. The last layer uses a  $1 \times 1$  convolution to map each 16-component feature vector to the desired number of classes. The U-Net architecture (fig. 2), is typically used for image segmentation-oriented networks.

The backbone is the approach used as part of coding in this paper. The long pass-through connections used in U-Net allow small and fine scale features to be preserved that would otherwise be lost in the decoding stages.

Digital Petroleum research have found that U-Net (with the following backbones: inceptionv3, efficientnetb4, inceptionresnetv2) and LinkNet (backbone: inceptionv3) performed better on this data.

It is now known that UNet+resnet-50 in architecture shows the best results

in rock sample  $\mu$ CT segmentation into different minerals.

This is results which Digital Petroleum researchers get with Unet cNN.

Table 1: Digital Petroleum researcher's results for mineral mapping with UNet

Model	Pores	Quartz	Albilite	Clay
Unet+inceptionv3	0.63	0.58	0.45	0.44

### 3 Application

#### 3.1 3D visualization of $\mu$ CT images

In this research and for practical purposes, there is necessity to visualize 3D-model of rock core sample. And for this purpose was selected special library for data visualizing on Python PyVista.

There is one reason why was selected PyVista, because it will be conveniently to visualize rock sample before, during and after pre-processing data in Jupyter Notebook, and after calculations.

The algorithm of this operation:

1. Collect many crops on one slice (if a number of crops 64 ( $8 \times 8$ ) with  $256 \times 256$ , slice will be collected with concatenate all row slice, and concatenate rows. The result will be  $2048 \times 2048$  pixels image).
2. Create UniformGrid with  $n\_slices \times 2048 \times 2048$  dimensioins.
3. Set values to points of this grid according to image arrays.
4. Visualize three-dimensional tensor with PyVista Plotter.
5. Add mesh.orthogonal\_slice module.

After these steps, it is possible to see 3D-model of core-sample:

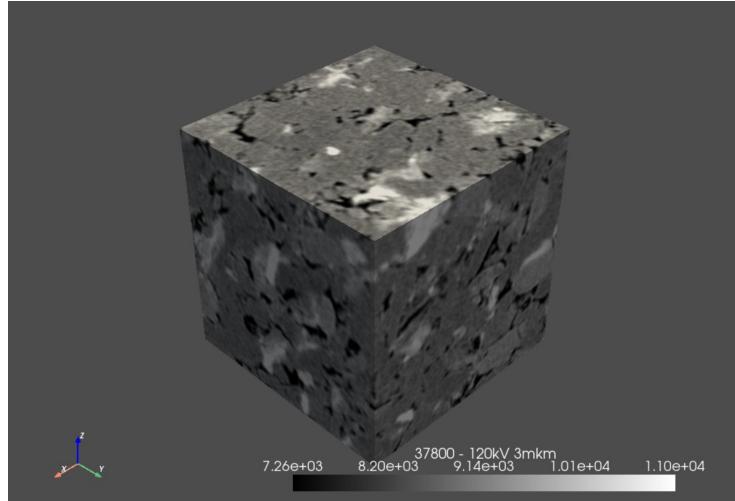


Figure 5: 3D-model of core sample in PyVista

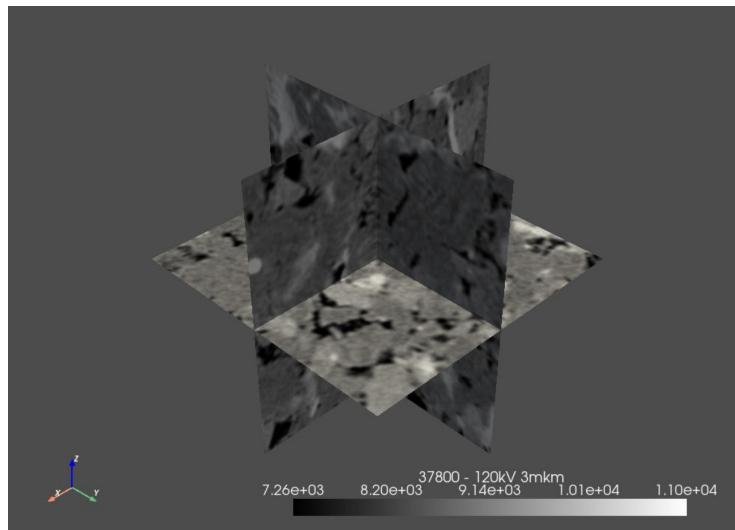


Figure 6: 3D slicing in PyVista

If you install `pyvtk` module in Jupyter Notebook, it is possible to move, rotate and scale 3D-model, and also move slices with `cpos` parametrs. With this function, it is possible to visualize any 3D-image, which can be represented as Numpy array (tensor).

### 3.2 cNN parameters

In air-xenon segmentation case, there is principal scheme, which contains several stages. This scheme with stage represented of figure 7.

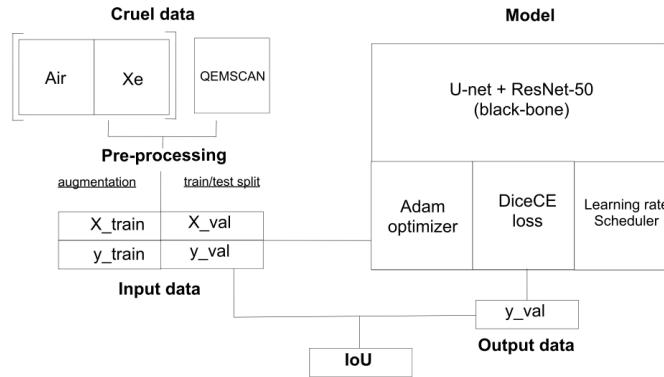


Figure 7: Principal scheme cNN for air-xe core mineral mapping

In this research, we have similar scheme (fig. 8).

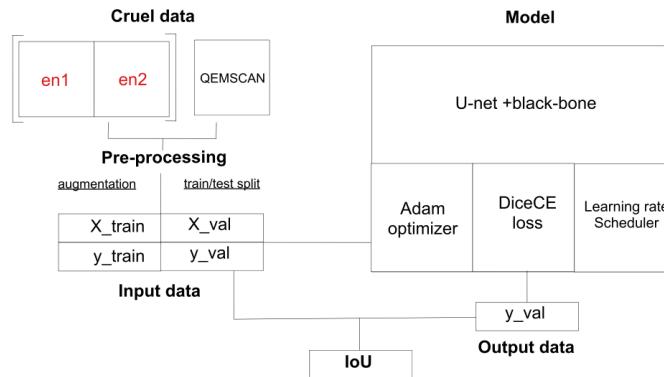


Figure 8: Principal scheme cNN for dual energy core mineral mapping in this research

As input data two different  $\mu$ CT images used with different value of  $\mu$ CT-scanner energy: 120 kV and 80 kV.

For our purposes we use crops  $256 \times 256$  pixel dimension from one slice 37800 sample of 3mkm  $\mu$ CT shooting and after test/train split operation there is:

- 21 samples for train;

- 9 samples for validation.

Heavy models like Unet have large number of parameters, and we need increase the number of train samples. For this purpose we use augmentation. Augmentation does not applied for validation dataset.

Augmentation included following operations:

- Horizontal flip with probabillity 0.5;
- Vertical flip with probabillity 0.5;
- Random rotate by 90 degrees with probabillity 0.5;
- Transpose with probabillity 0.5;
- Rotation by some angle(limit=30 degress) with probabillity 0.3.

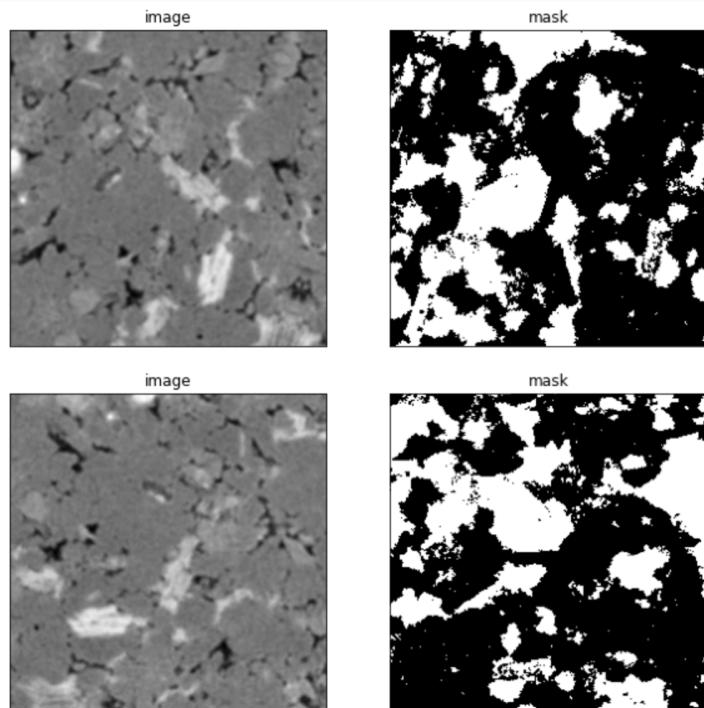


Figure 9: Augmentation example

These operations allow us to multiply the training dataset.

Weights initialization was according to mineral composition all train dataset.  
Adam optimizer was chosen as the best possible optimizer.

Dice Cross-Entropy loss was chosen as loss function for models training.

The Unet neural network is used in this study because it has shown remarkably well in the air-xenon case. The black-bones for this model will vary according to the following list (all models Imagenet pre-trained):

- ResNet-50;
- ResNet-18;
- Efficient-b0;
- Efficient-b4;
- Inception-v4.

Each model was trained on the same number of epochs (2500). With an increase or a small decrease in the number of epochs, no noticeable improvement was observed.

### 3.3 Results

The ResNet-50 neural network was the first to be trained.

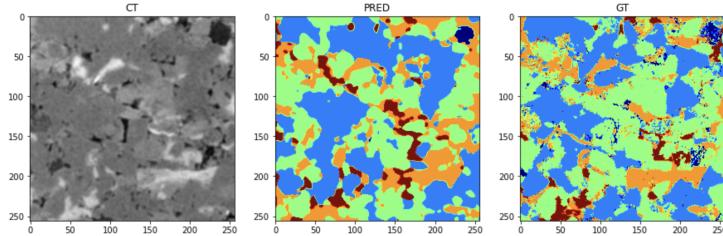


Figure 10: Unet+ResNet-50 results for dual energy case

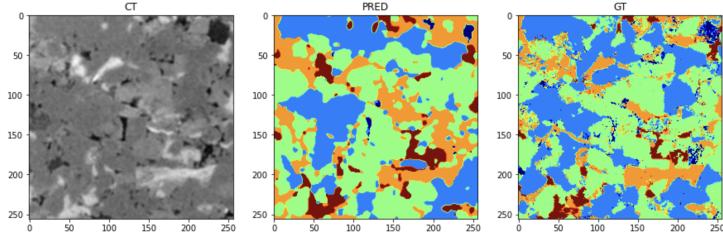


Figure 11: Unet+ResNet-50 results for only one energy case

The figures 10 and 11 show that in the case of this slice, the two-energy neural network is better at recognizing patterns, such as the porosity area in the upper right corner.

But if you look at the IoU value across all samples, you can see that the prediction values vary greatly from sample to sample for different classes.

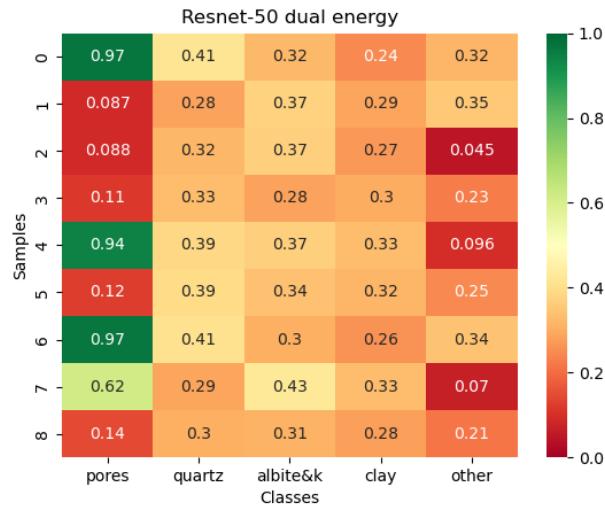


Figure 12: Unet+ResNet-50 heat map for dual energy case

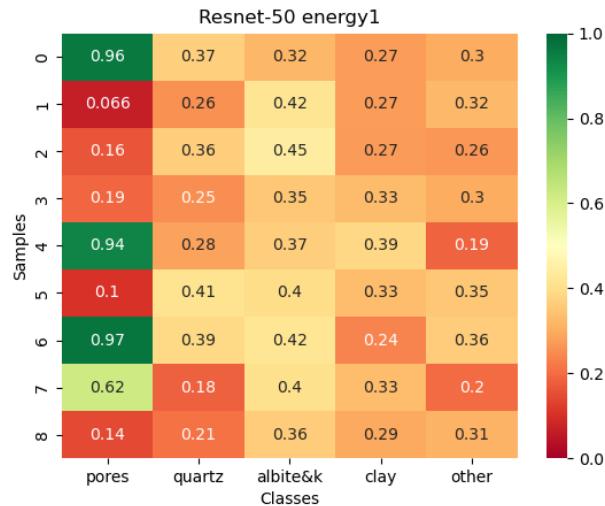


Figure 13: Unet+ResNet-50 heat map for only one energy case

If you look at the statistics for all samples on the bar graph (14), you can see that a neural network trained on only one energy value recognizes all classes except the second better than dual energy.

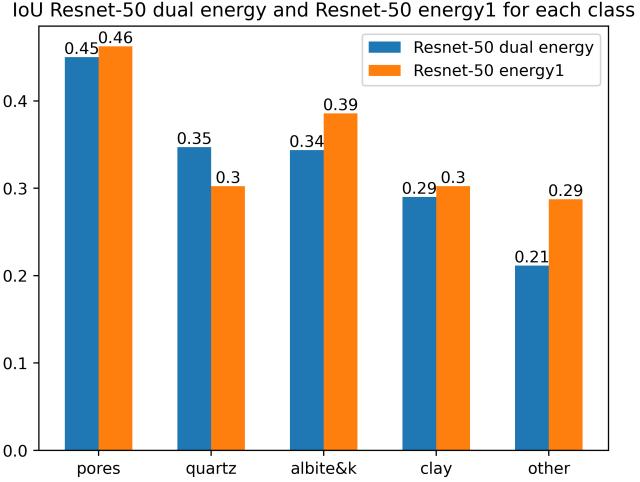


Figure 14: Comparing results for dual and only energy cases for ResNet-50

The Resnet-50 did not perform as well as expected. It was assumed that the neural network trained poorly on a huge number of parameters, because it is not enough the size of the dataset of a single slice. It was decided to try using neural networks with a much smaller number of parameters.

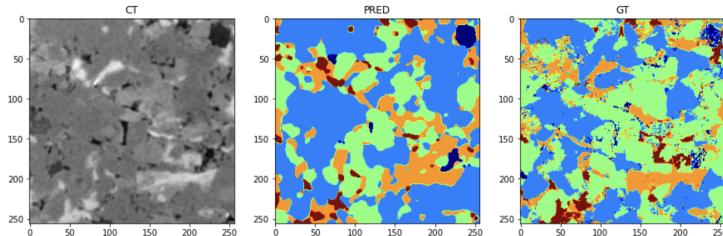


Figure 15: Unet+ResNet-18 results for dual energy case

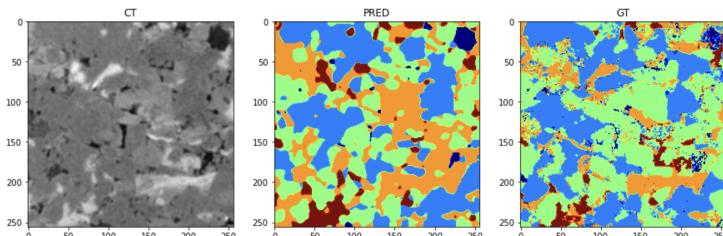


Figure 16: Unet+ResNet-18 results for only one energy case

You can see in the figure 16 that pattern recognition happened better (porosity in the upper right corner and "red" class in the lower left corner), but also

markedly increased the value of the false-positive predicted "orange" class. In this case, a neural network trained on two energies shows better performance for this slice. But by the total score for all classes and samples, it still loses to a neural network trained on a single energy (fig. 17).

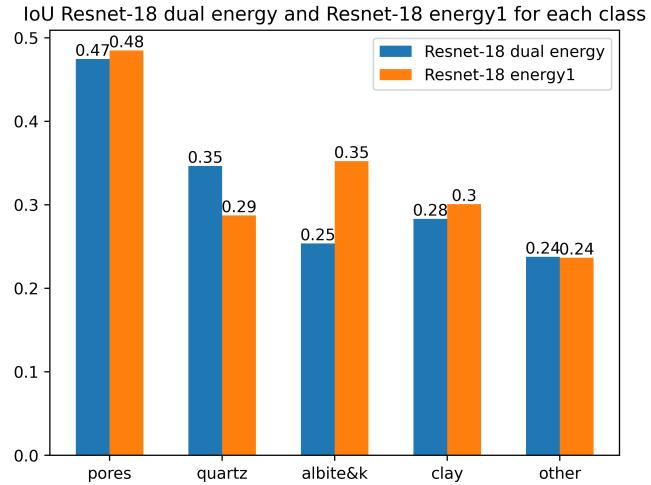


Figure 17: Comparing results for dual and only energy cases for ResNet-18

Then it was decided to try to train the neural network on two EfficientNet models: for 4 million parameters and 10 million.

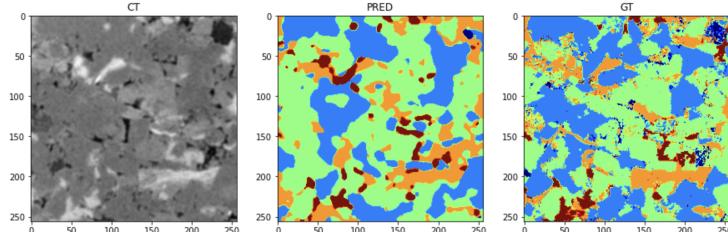


Figure 18: Unet+EfficientNet-b0 results for dual energy case

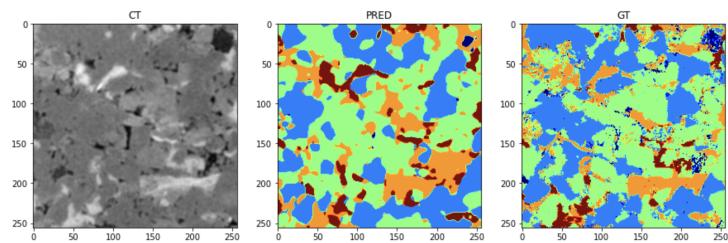


Figure 19: Unet+EfficientNet-b0 results for only one energy case

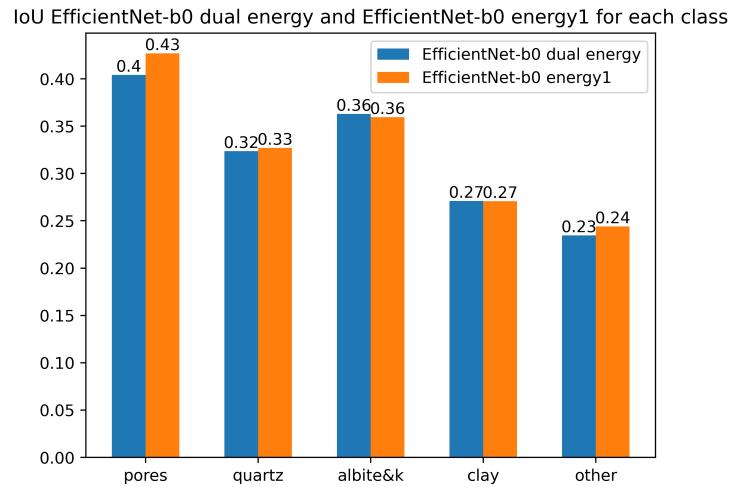


Figure 20: Comparing results for dual and only energy cases for EfficientNet-b0

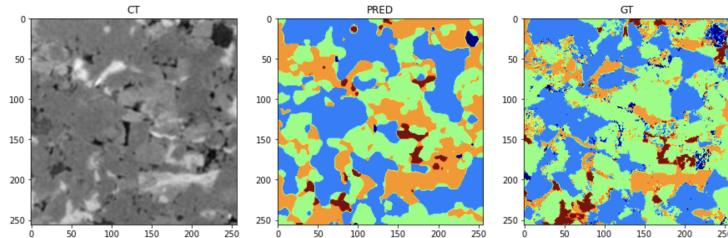


Figure 21: Unet+EfficientNet-b4 results for dual energy case

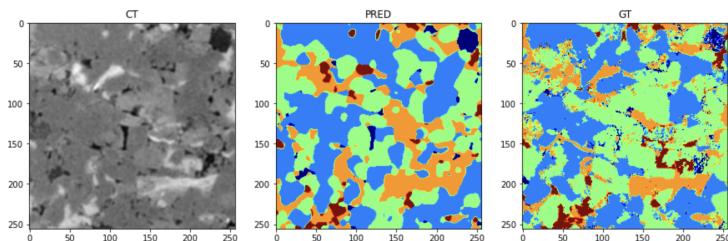


Figure 22: Unet+EfficientNet-b4 results for only one energy case

If evaluated visually, it seems that EfficientNet-b4 on only one energy looks more like the patterns seen on QEMSCAN.

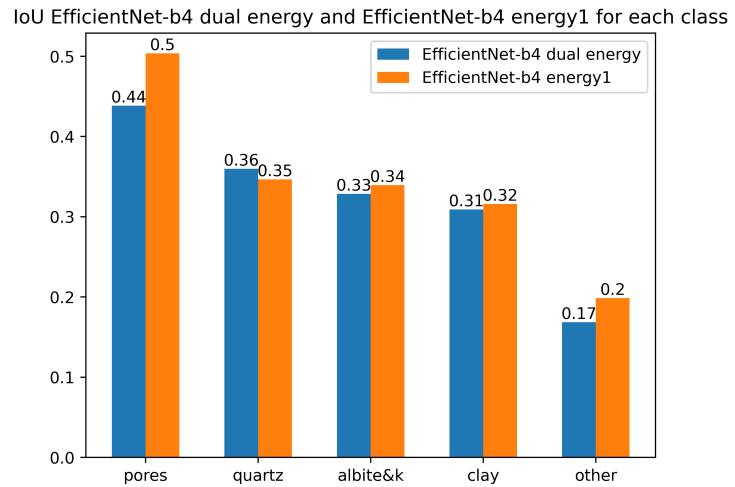


Figure 23: Comparing results for dual and only energy cases for EfficientNet-b4

This is confirmed by the score, which reached the highest value of 0.5 when recognizing class 0 (porosity).

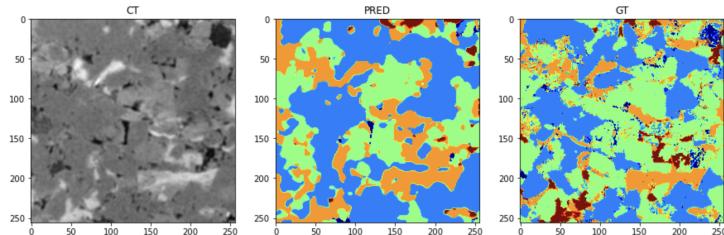


Figure 24: Unet+Inception-v4 results for dual energy case

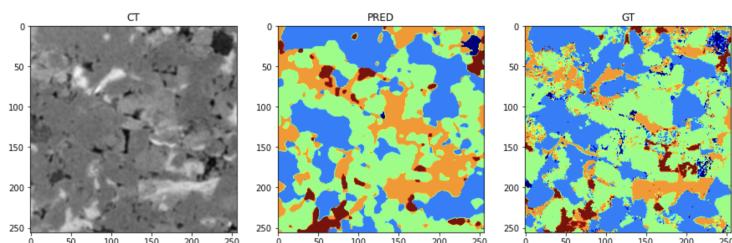


Figure 25: Unet+Inception-v4 results for only one energy case

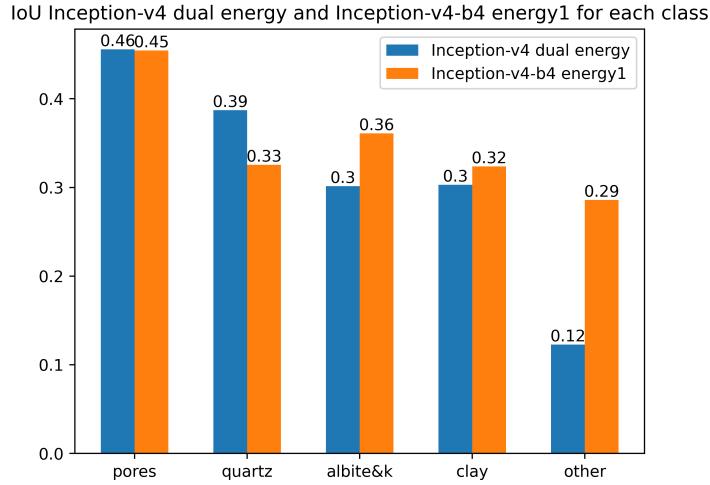


Figure 26: Comparing results for dual and only energy cases for Inception-v4

Thus, we can see a summary table with all the results of the study. It can be seen that two models can be considered to work well on a dataset of one slice. These are ResNet-18 trained on only one energy and EfficientNet-b4 trained on only one energy.

Table 2: Pivot table of result for different models

Model	Pores	Quartz	Albilite	Clay	Other
Dual ResNet-50	0.45	0.35	0.34	0.29	0.21
En1 ResNet-50	0.46	0.3	<b>0.39</b>	0.3	0.29
Dual ResNet-18	0.47	0.35	0.25	0.28	0.24
En1 ResNet-18	<b>0.48</b>	<b>0.29</b>	<b>0.35</b>	<b>0.3</b>	<b>0.24</b>
Dual EfficientNet-b0	0.4	0.32	0.36	0.27	0.23
En1 EfficientNet-b0	0.43	0.33	0.36	0.27	0.24
Dual EfficientNet-b4	0.44	0.36	0.33	0.31	0.17
En1 EfficientNet-b4	<b>0.5</b>	<b>0.35</b>	<b>0.34</b>	<b>0.32</b>	<b>0.2</b>
Dual Inception-v4	0.46	<b>0.39</b>	0.3	0.3	0.12
En1 Inception-v4	0.45	0.33	0.36	<b>0.32</b>	<b>0.29</b>

If you look at the problem globally, then the score of a model trained on only one energy and on two energies does not differ much. But on training on such a small dataset, training on one energy is noticeably better, but far from ideal.

In the case of ResNet-50 and ResNet-18, changing to a model with a smaller number of parameters managed to increase the score, but this logic does not work in the case of Efficient, where the change to a model with a larger number of parameters on the contrary led to a significant increase in the score.

In 5 out of 5 cases, the black-bones model trained on only one energy had a

bigger metric for class recognition than the model trained on dual energy. This can be explained by two reasons:

1. Dual energy gives no additional information to the model and only confuses it.
2. A small dataset does not allow a complex model with a large number of parameters to learn all the patterns of mineral recognition.

The author of the study is more biased toward option number 2 because the class of quartz was recognized significantly better in 4 of 5 cases (5 of 5 better or approximately the same) than the model trained on only one pattern. This may mean that it is easier to learn the quartz mineral pattern than the other minerals, and this additional information is given by the second energy input image. Perhaps it makes sense to train models not on two input energies, but on a larger number and with a much larger dataset.

## 4 Conclusions

1. was possible to successfully develop a 3D visualization tool for rock samples in Python.
2. The problem of image segmentation using  $\mu$ CT images at dual energies was investigated.
3. Ten neural networks were trained on different input data and on different black-bone models.
4. Was found that in the case of pattern recognition, not all models increase their score when moving from a more complex model to a simpler one (ResNet increase, Efficient does not increase).
5. In all cases, on a small training dataset from a single slice, a model trained on only one energy value recognizes classes better than a model trained on dual energy.
6. Some minerals, such as quartz, were found to be better recognized by the dual energy trained model.
7. Further investigation on a larger training dataset is required to confirm the hypothesis that dual energy models simply have not enough data to identify patterns of different minerals.

## References

- [1] V. V. Alekseev, D. M. Orlov, and D. A. Koroteev, “Multi-Mineral Segmentation of SEM Images Using Deep Learning Techniques,” in Day 3 Thu, October 14, 2021, Virtual, Oct. 2021, p. D031S017R004. doi: 10.2118/206526-MS.
- [2] D. P. Kingma and J. Ba, “Adam: A Method for Stochastic Optimization,” 2014, doi: 10.48550/ARXIV.1412.6980.
- [3] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, “Focal Loss for Dense Object Detection,” 2017, doi: 10.48550/ARXIV.1708.02002.
- [4] O. Ronneberger, P. Fischer, и T. Brox, «U-Net: Convolutional Networks for Biomedical Image Segmentation», 2015, doi: 10.48550/ARXIV.1505.04597.
- [5] Knappett, C., Pirrie, D., Power, M.R., Nikolakopoulou, I., Hilditch, J., Rollinson, G.K. 2005. Mineralogical analysis and provenancing of ancient ceramics using automated SEM-EDS analysis (QEMSCAN): A pilot study on LB I pottery from Akrotiri, Thera. Journal of Archaeological Science, in press doi:10.1016/j.jas.2010.08.022