Article

Integrating Machine/Deep Learning Methods and Filtering Techniques for Reliable Mineral Phase Segmentation of 3D X-ray Computed Tomography Images

Parisa Asadi and Lauren E. Beckingham *

Department of Civil & Environmental Engineering, Auburn University, Auburn, AL 36830, USA;
pza0029@auburn.edu
* Correspondence: leb0071@auburn.edu; Tel.: +1-(334)-844-6260

Abstract: X-ray CT imaging provides a 3D view of a sample and is a powerful tool for investigating the internal features of porous rock. Reliable phase segmentation in these images is highly necessary but, like any other digital rock imaging technique, is time-consuming, labor-intensive, and subjective. Combining 3D X-ray CT imaging with machine learning methods that can simultaneously consider several extracted features in addition to color attenuation, is a promising and powerful method for reliable phase segmentation. Machine learning-based phase segmentation of X-ray CT images enables faster data collection and interpretation than traditional methods. This study investigates the performance of several filtering techniques with three machine learning methods and a deep learning method to assess the potential for reliable feature extraction and pixel-level phase segmentation of X-ray CT images. Features were first extracted from images using well-known filters and from the second convolutional layer of the pre-trained VGG16 architecture. Then, K-means clustering, Random Forest, and Feed Forward Artificial Neural Network methods, as well as the modified U-Net model, were applied to the extracted input features. The models' performances were then compared and contrasted to determine the influence of the machine learning method and input features on reliable phase segmentation. The results showed considering more dimensionality has promising results and all classification algorithms result in high accuracy ranging from 0.87 to 0.94. Feature-based Random Forest demonstrated the best performance among the machine learning models, with an accuracy of 0.88 for Mancos and 0.94 for Marcellus. The U-Net model with the linear combination of focal and dice loss also performed well with an accuracy of 0.91 and 0.93 for Mancos and Marcellus, respectively. In general, considering more features provided promising and reliable segmentation results that are valuable for analyzing the composition of dense samples, such as shales, which are significant unconventional reservoirs in oil recovery.

Keywords: 3D X-ray computed tomography; U-Net convolutional neural network; feed-forward neural network; random forest; 3D imaging of shale samples; Mancos; Marcellus

1. Introduction

Shales are important unconventional oil reservoirs. In these low permeability formations, hydrocarbons can be extracted using hydraulic fracturing, often enhanced with the injection of CO_2 referred to as CO_2-enhanced oil recovery (CO_2-EOR) [1]. As hydrocarbon distribution can be correlated with the inorganic and organic constituents of the formation, understanding the mineralogy and mineral distribution in these samples is critical for assessing the gas capacity. Studying mineral distribution in these formations can also help with the prediction of fracture formation and hydrocarbon recovery as fractures may preferentially form in calcite [2] or clay-rich regions [2–4]. In CO_2-EOR systems, fracture aperture may be dynamic as minerals dissolve and precipitate following CO_2 injection. This process, however, is complex and not well understood where reactions may enhance

or reduce fracture permeability, influenced by the distribution of reactive minerals on the fracture surface [5,6].

X-ray computed tomography (X-ray CT) is a powerful means of imaging that can facilitate the 3D non-destructive characterization of geologic samples. While historically high-resolution X-ray CT imaging of samples was only possible using synchrotron sources, advancements in the development of laboratory and benchtop instruments have broadened access to high-resolution X-ray CT imaging in the form of X-ray mico- and nano-computed tomography (X-ray nano CT) instruments. These X-ray CT instruments can provide a three-dimensional (3D) depiction of an object with high resolution (up to 100 nanometers) with a relatively large field of view (FOV) compared to other 3D imaging techniques [7]. This scale of information is typically only widely available in the laboratory using 2D imaging approaches such as scanning electron microscopy (SEM). 3D X-ray CT imaging, however, offers an additional depth of information that is not available with 2D-based microscopy analysis [8]. This can facilitate the analysis of the 3D nature of the sample, such as pore size distribution and connectivity [9]. The use of 3D X-ray CT imaging instead of 2D imaging also eliminates stereological errors generated by conventional 2D microscopy analysis used for porous media samples, allowing more accurate analysis of the samples [8,10].

X-ray CT images consist of voxels of varying grayscale intensity correlated to the X-ray attenuation of the material. Variations in attenuation result due to differences in material properties including density and atomic number. Traditional image processing and segmentation rely on the similarity or intensity (or both) of the pixels to delineate the boundaries of the objects [11]. Segmentation of images by attenuation can facilitate quantitative image processing of sample properties and characteristics including porosity, surface area, and mineral volume fractions [12–14]. It should be noted, however, that not all mineral phases can be segmented using attenuation alone due to close or overlapping x-ray attenuation coefficients and partial volume effects [15–17]. In addition, beam hardening can make the grayscale attenuation of a given phase differ depending on the location of the phase within the sample [8], especially in the borders of the images. This is particularly noted in dense samples such as shale cores. These challenges make image segmentation, even into groups of minerals with similar attenuations, time-consuming, labor-intensive, and subjective. One practical alternative is utilizing some advanced techniques such as machine learning.

Machine learning methods utilize mathematical models to excavate nonlinear underlying patterns in a dataset, which helps a computer system make predictions or classifications on the dataset [10,18]. Within the petroleum industry and geosciences, machine learning models have been used in various applications, such as fluid transport analysis [19], rock typing [20,21], reservoir characterization [22], and multiscale imaging to quantify properties of shale source rocks [23], as well as phase segmentation [24] of SEM images of a shale. A few studies have even focused on segmenting different mineral phases within X-ray CT images using machine learning methods [8,10,15,25–27]. However, the effectiveness and reliability of machine learning methods to differentiate, quantify and extract features from X-ray CT images of shales are not well understood but could provide a valuable approach to image processing in the petroleum and geosciences field.

Machine learning-based phase segmentation makes it possible to simultaneously consider several extracted features in addition to the voxel attenuation for a more reliable phase segmentation. This ability would be valuable when several phases have similar/overlapping attenuation as well as in images with spatial variations in attenuation for a given phase. In these cases, additional features of phases may be key to individual phase segmentation [15,25]. While there exist works at the intersection of core imaging and machine learning, no study has yet investigated the potential improvement of phase segmentation in shale X-ray CT images using filtering techniques, as additional input, along with machine learning models to reliably segment different phases. This study investigates the performance of several filtering techniques with three machine learning methods and one deep learning method to assess the potential for reliable feature extraction

and pixel-level phase segmentation of a Marcellus and a Mancos shale. This study will help geologists to obtain the different distinguishable phases in 3D X-ray CT images to provide practical techniques for reliable phase segmentation. To our knowledge, it is the first time that the task of distinguishing mineral phases of 3D X-ray CT images is integrated through both pixel-level classifications using machine learning models along with filtering techniques and image segmentation method using a deep learning model on shale samples.

2. Materials and Methods

2.1. Mancos and Marcellus Shale Samples

Shale core samples 1″ in diameter and 2″ in length from the Marcellus and Mancos formations obtained from Kocurek Industries were used in this work. The Marcellus shale is an organic-rich shale formation in the northeastern US [28]. According to the U.S. Geological Survey (USGS) assessment, the Marcellus Shale contains about 84 trillion cubic feet and 3.4 billion barrels of undiscovered, technically recoverable natural gas and natural gas liquids, respectively. The Mancos shale formation is in the Midwestern US and is a major source of rock for oil and/or gas in the Rocky Mountain Region. It is also estimated that Mancos contains 66.3 trillion cubic feet of shale gas, 74 million barrels of shale oil, and 45 million barrels of natural gas liquids in addition to some unproved and undiscovered recoverable resources based on the U.S. Geological Survey [29].

The mineralogical analysis obtained by XRD analysis in [3] is listed in Table 1. The Marcellus shale is predominantly calcite, quartz, and pyrite. It should be noted that the clay minerals were not detected in XRD. This indicated their minor concentrations with respect to the major calcite phase. The Mancos formation has a larger variation in mineralogy and is 57.63% quartz with approximately equal clay (Muscovite and Kaolinite), carbonate (dolomite and calcite), and feldspar (microcline and albite) fractions, and minor pyrite.

Table 1. XRD composition analysis of the shale samples (Volume percentage of minerals).

Sample/ Mineral	Quartz	Pyrite	Calcite	Muscovite	Kaolinite	Microcline	Albite	Dolomite
Mancos	57.63	0.22	6.88	9.68	3.37	7.45	8.11	6.66
Marcellus	2.97	0.33	96.70					

2.2. Sample Preparation and Image Acquisition

Sub-samples for each formation were extracted from 1″ core sample for high-resolution X-ray CT imaging by Carl Zeiss Microscopy Customer Center Bay Area (Pleasanton). 0.8 mm × 1 mm sections were milled from the top 5mm of each core sample using a laser and mounted on pins (Figure 1). X-ray CT images of each sample were then collected using an Xradia 620 Versa (Zeiss Microscopy Customer Center Bay Area). The sample scan parameters are shown in Table 2 where the X-ray energy of 60 kV was utilized to provide better imaging of the less dense features of the sample [25]. Almost all parameters are the same for both samples, only the filter and exposure time are different. Projections were reconstructed using the Reconstructor Scout-and-Scan Control System software and exported as TXM and tiff format files. The resulting images are used in this study.
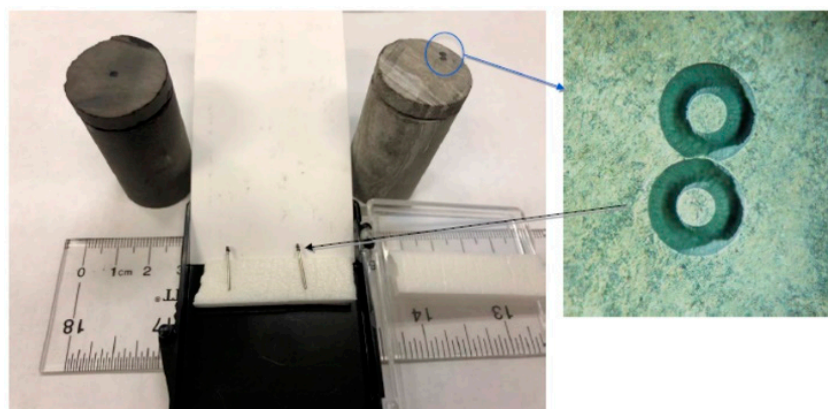
Figure 1. Marcellus and Mancos core samples were extracted and mounted on pins to scan (Carl Zeiss Microscopy Customer Center Bay Area).

Table 2. Experimental conditions of the X-ray CT scanning (Zeiss Microscopy Customer Center Bay Area).

Sample Name	Marcellus	Mancos
Voltage	60 kV	60 kV
Power	10 W	10 W
No. of projections	2201	2201
Spatial resolution	0.8 μm	0.8 μm
Bin	2	2
Objective	20×	20×
Filters	LE4	LE2
Exposure time	20 s	9 s

2.3. Classification Experiment

This study explored mineral phase segmentation in X-ray CT images of shale samples using machine/deep learning with filtering techniques to provide image features as additional inputs. Figure 2 shows the experimental design. The workflow includes defining input variables from cross-sectional slices of the original X-ray CT images, followed by machine learning and deep learning for phase segmentation. The machine learning section in the workflow of the study included three parts (three different sets of input images), each of which considered different extracted features as the input data for the machine learning methods. These stacked input variables were introduced to the Random Forest (RF), K-means clustering (K-means), Feed Forward Neural network (FNN), and the U-Net deep learning models to train and evaluate their computational performance and accuracy. Each machine learning model was trained and cross-validated for each sample using the 78,018,066-pixel data (81 × 994 × 969 pixels). For the deep learning method, data augmentation was also applied to increase the amount of data by augmenting images including flipping, zooming, shifting, and rotation. The associated images were cropped to 128 × 128-pixel size images to be used for the U-Net deep learning method. After data augmentation, 11,000 slices with 128 × 128-pixel size were provided.
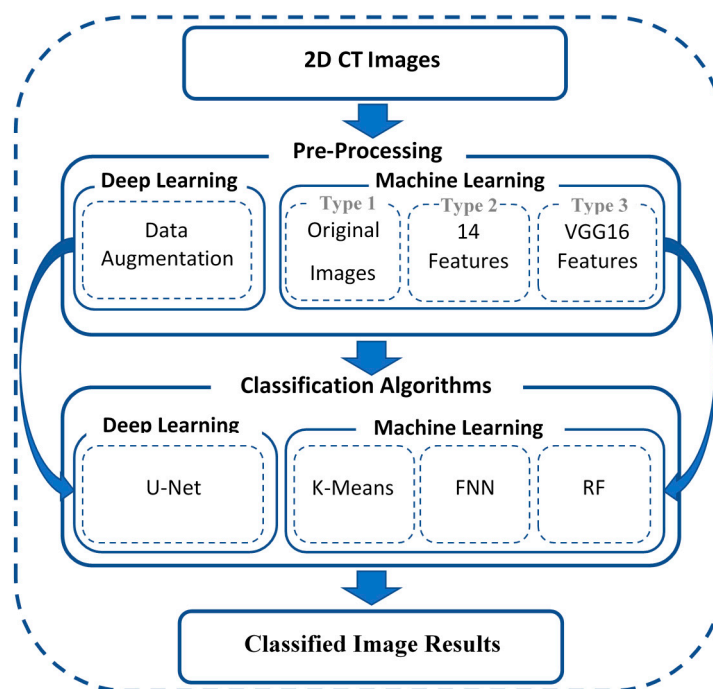
Figure 2. The (workflow) schematic of mineral classification using different filters as input variables and three machine learning approaches.

2.3.1. Image Processing and Feature Extraction

Mineral phases can have distinguishable features in images, such as texture and grain size. In feature-based classification, images can be represented in additional dimensions (i.e., extracted feature images) which helps to better explore the similarities and differences of different phases to classify phases in separate classes. In this study, different types of features were extracted from the 2D cross-sectional slices of X-ray CT images, and phase classification was performed leveraging these features in addition to attenuation values. Two different feature extraction strategies were utilized (Figure 2) to create feature maps, which are briefly discussed in the following sections.

Filtering Techniques

Filters were applied to both Mancos and Marcellus images using the OpenCV [30] and Scikit-image [31] filters in Python to provide feature maps representative of texture, grain size, edges, entropy, and abnormality, as well as color attenuation. Several well-known filters, including Gaussian, Median, Sobel (45 degrees, vertical and horizontal), and Gabor (Table 3), were used. Each filter convolutes the original image to perform specific tasks such as blurring (e.g., the Gaussian filter), edge detection, or texture extraction (e.g., the Gabor and Sobel filters). For example, the Gaussian filter convolutes the image using specific standard deviations (σ), creating feature maps in different scales. Filtered images with higher standard deviation (more blurred) lose details of smaller grains while those with smaller standard deviations preserve more features from the smaller grains. The difference in the two images provides new information that can be also used as input variables to the machine learning models. In this work, edge features and textures were extracted using Sobel and Gabor filters (horizontal and vertical), while blobs and corners were extracted with a difference of Gaussians, the determinant of the Hessian matrix, and Laplacian filters.

Table 3. Filters used to extract features and their associated kernel size and standard deviations (SD).

Filter Name	Kernel Size	SD	Task	Abbreviation
Median	3×3	-	denoising and integrating phases	-
Gaussian	3×3	1	denoising and integrating phases	-
Difference of Gaussians 1-10	3×3	1, 10	Detecting smaller features	DoG 1-10
Difference of Gaussians 1-5	3×3	1, 5	Detecting smaller features	DoG 1-5
Difference of Gaussians 2-5	3×3	2, 5	Detecting smaller features	DoG 2-5
Difference of Gaussians 1-2	3×3	1, 2	Detecting smaller features	DoG 1-2
Sobel (45 degrees)	3×3	-	Detecting edges and textures	Sobel-45
Sobel (Vertical)	3×3	-	Detecting edges and textures	Sobel-V
Sobel (Horizontal)	3×3	-	Detecting edges and textures	Sobel-H
Laplace	3×3	-	Detecting blobs and edges	-
Difference of Hessians	3×3	3, 5	Detecting blobs and corners	DoH
Gabor (Horizontal)	3×3	-	Detecting textures	Gabor-H
Gabor (Vertical)	3×3	-	Detecting textures	Gabor-V

Extracting Filters from a Deep Convolutional Layer

One promising technique for feature extraction is applying pre-trained filters obtained from a trained deep learning model. In this study, the trained filters obtained from the second convolutional layer of the first block in the visual geometry group (VGG16) network [32] were applied to extract image features. The VGG16 convolutional neural network is widely used in a wide variety of fields because of its high generalization capability, simple structure, and accuracy where is it among the top-5 of models tested on ImageNet, a dataset of over 14 million images belonging to 1000 classes [32]. As a result, the VGG16 network has learned a rich array of feature representations for a wide range of images. Here, features were extracted from X-ray CT images using the second convolutional layer of the first block of VGG16 with 64 filters. This layer was selected since it is the deepest layer of VGG16 which its output feature images will have the same size as the original input CT images. As a result, same-size feature images were provided. In this study, the filters of this block were applied to each X-ray CT slice to extract features for machine learning models.

2.3.2. Image Segmentation and Labeling

Segmentation refers to the grouping of pixels into several classes by identification and isolation of pixels that have the same features into a single category. The most common segmentation method is based on the histogram analysis of grayscale intensities of the pixels in the image which provides the distribution of the grayscale level of each pixel in the image. In X-ray CT images, this provides a relative identification of the different phases contained in the image. Figure 3 shows the histograms of gray intensities for a 2D slice from the Mancos and Marcellus datasets. The two peaks indicate a clear threshold value for separating two classes, which in this case are the background (left peak which has zero intensity) and the sample (right peak). Further segmentation and phase labeling were carried out manually based on knowledge of phases in the sample (their attenuation grayscale, texture, and size characteristics) and characteristics of these phases from processed SEM, BSE, and EDS images in prior work [3] and XRD data (Table 1). Figure 4 shows an example of labeled images and their corresponding colors and phases (Furthermore, reader are referred to 3D view videos (S1 to S6) of both samples in the Supplementary Material which show the samples and segmented mineral phases). For example, in the Marcellus sample, phases were considered as grains (mostly quartz), matrix (which was mostly calcite and clays), pyrite, organic matter, and background. Labels were assigned by manually segmenting, correcting, and post-processing of 2D slices. In total, 81 2D slices (994×969-pixel images, i.e., 78,018,066 pixels in total) from the X-ray CT image were segmented, in which, 60 labeled X-ray CT images were used to train the classifiers and the 21 images used to evaluate the performance of trained models.
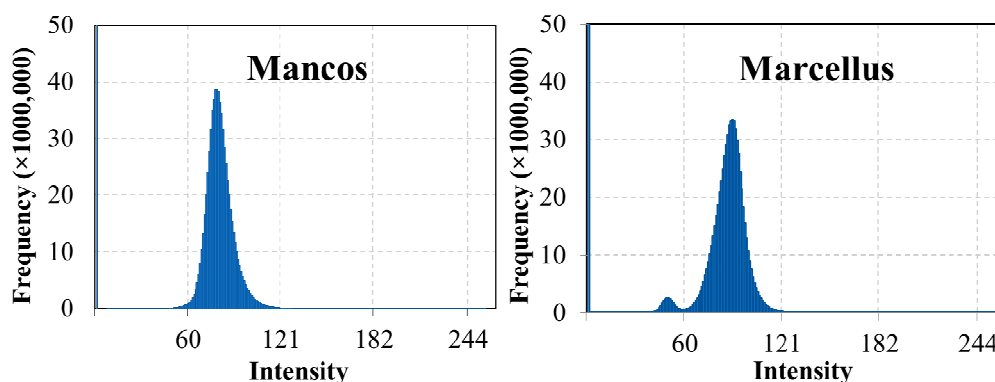
Figure 3. Representative histograms for 2D slices from Mancos and Marcellus X-ray CT images.
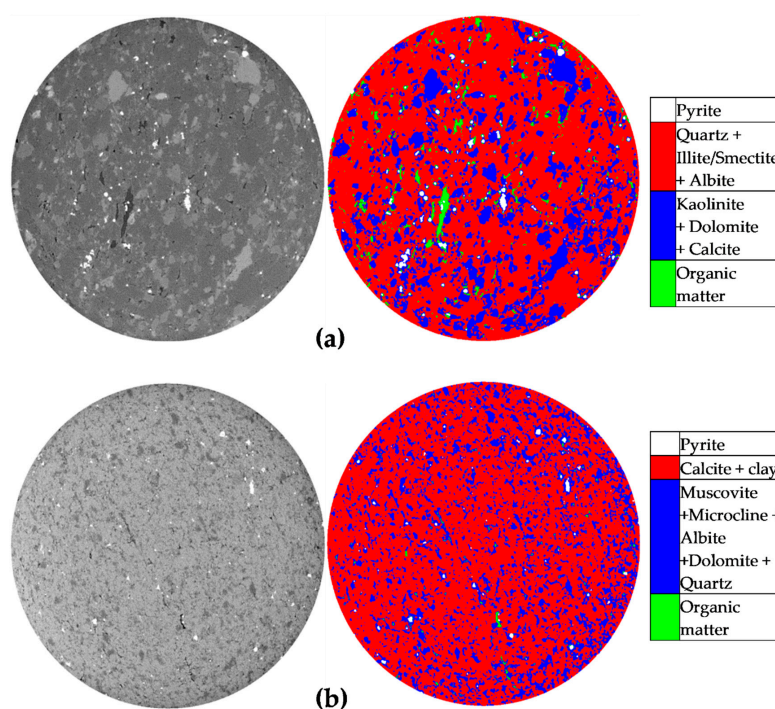

Figure 4. Labeled images and their corresponding phases and colors. (**a**) Mancos sample; (**b**) Marcellus sample.

2.3.3. Machine Learning

Machine learning algorithms rely on sets of features to train the classification models. Once trained, models can be used to identify features in an unknown dataset and group them into respective classes. Machine learning methods in general fall into unsupervised and supervised categories. Here, three classification machine learning algorithms are implemented. Of these, k-means is unsupervised, while RF and Feed Forward Artificial Neuron Network are non-linear supervised classification algorithms. The theories and parameters for the considered algorithms are briefly described in this section.

K-Means Clustering

K-means clustering partitions a collection of data into a k number group of data to make respective clusters. It then calculates the k centroid and assigns each point to the cluster with the nearest centroid from the respective data point [33]. There are several ways of calculating the similarities between pixels, in this study we used Euclidean distance to define the most similar centroid. Once the grouping is done it recalculates the new centroid for each cluster and a new Euclidean distance between each center and each data point to

assign the points in the cluster with minimum Euclidean distance. As such, K-means is an iterative algorithm in which it minimizes the sum of distances from each object to its cluster centroid [33].

Here, K-means clustering was implemented using the Scikit-learn library [34] in python. The K was considered 3 to 9 to find the best (k = 5) number of clusters. In addition, the epoch varied from 100 to 1000, the initialization method was set to "K-means++" to speed up the algorithm convergence. The "n-initint" was set to 10 and the best results were kept, "n-jobs" was set to -1 for parallel computation, and the algorithm was set to "Elkan".

Random Forest

Random Forest (RF) [35] is an ensemble learning method for classification and regression consisting of multiple individual trees. In the training step, observations are randomly sampled from the training dataset to build an individual classification/regression tree and the best feature to split a node is selected within a randomly picked subset of features to further add randomness to the system. This process is repeated for a certain number of trees (ntree) and different randomly selected sets of input features (mtry). The overall goal of the process is to decrease the variance and improve the training accuracy. Then, the predictions are made for the remaining observations based on a weighted vote across all trees. The hierarchical structure of RF makes it an appealing method for phase segmentation where the data are largely imbalanced (i.e., some phases have more frequency than minority classes).

In this study, the RF model was implemented using the "Random Forest" Scikit-learn library [34]. In the first run, a forest of 100 random trees was created (i.e., ntree = 100); each tree had unlimited depth and was grown without pruning. At each node, m input variables (estimated as the square root of the numbers of predictors, i.e., mtry $\approx$ 4 for type2 input variables and mtry $\approx$ 8 for type3) were randomly selected among all input variables to training the model. In this study, only two main hyperparameters, ntree and mtry, were fine-tuned using random search, and the depth of the tree remained as the default settings in the RF classifier in the Scikit-learn library. Then the random search [36] was applied to fine-tune the hyperparameters of RF (i.e., mtry and ntree) by grid searching between the number of trees in the range of 100 to 700 trees and the number of randomly selected input variables at each split (i.e., mtry) in the rage of 2 to 8, resulting in the final forest with ntree = 200 and mtry = 2 in each node for type2 variable inputs and ntree = 300 and mtry = 2 for type3 variable inputs.

Feed Forward Neural Network

Feed Forward Neural Networks can learn and model non-linear and complex relationships, which makes them a compelling approach for phase segmentation since high non-linearity can exist among input features and the target classes. A network consists of an input layer, hidden layers, and a layer of output neurons with multiple activated perceptrons stacked together and associated weights and biases. Adjacent layers are fully connected, and results obtained from each layer are fed into the next layer through a non-linear transformation called an activation function. Neurons of the input layer receive the input features, process them, and pass the processed information through the hidden layers to the last layer. The output of the last layer adjusts based on the predefined loss function using backward propagation by changing the weights and biases. Thus, the network trains by adjusting the weights to predict the correct class label of the given inputs for the classification task.

In this study, we used a shallow Feed Forward Neural Network as it was easier to be trained and fine-tuned on a small dataset. Based on the recommendations of previous studies [37,38], several structures were randomly selected and tested, including one or two hidden layers with a various number of neurons (i.e., 2×14 and 2×64, for type 2 and type 3 input variables, respectively), and their corresponding performance compared. For the type 1 input variable, the best number of hidden neurons was 32. Finally, a three-layer

neural network with a single hidden layer consisting of a two × number of input neurons was selected as the hidden units.

TensorFlow library [39] was used to develop the Feed Forward Neural networks. For this study, the input layer consisted of N neurons, which corresponds to the number of input features, and the output layer included M neurons, which is the number of segmented phases in each sample, with a SoftMax activation function. A single hidden layer consisting of K neurons was selected as the hidden units. For example, the input layer consisted of 14 neurons for type 2 variable inputs and 64 for type 2 variable inputs, each corresponding to one of the selected input variables, which was connected to the hidden layer through a rectified linear unit (ReLU) activation function. In addition, each output neuron represented a distinct phase identified by having maximum probability obtained from a SoftMax activation function. The Categorical cross-entropy and focal loss and dice functions were considered as loss functions to adjust weights and biases, and Adam (adaptive moment estimation) was selected as the optimizer. We also used early-stopping callbacks to avoid overfitting. Fine-tuning was carried out using a fine-tuning function in the TensorFlow library where the numbers of epochs varied from 100 to 1000 and the learning rate (LR) from 10^{-6} to 10^0 (the best LR was 0.0001).

2.3.4. Deep Learning Model (U-Net)

U-Net is a fully connected Convolutional Neural Network (CNN) method that was initially developed for biomedical image segmentation by [32]. It is comprised of two main sections: an encoder and a decoder. The encoder contains several blocks which take an input and apply two convolutional kernels of 3 × 3, followed by a rectified linear transformation (ReLU) and a max-pooling operation with a stride of 2 × 2. The reduced spatial information with increased feature channels in the encoder path allows the network to learn the complex structure effectively due to propagating context information to the higher dimension. The decoder section consists of several expansion blocks, each of which passes the input to two convolutional layers followed by a 2 × 2 up-sampling layer. To maintain symmetry, the channels in the CNN layers are cut into half after applying each block. In the decoder, the size of the image gradually increases, and the depth gradually decreases. Finally, the last layer feature maps pass through a 1 × 1 CNN layer with M feature maps where M is the same as the number of classes desired.

Here, 2D slices from the original X-ray CT images and their corresponding segmented maps were used to train the network. The focal (equation 1) and dice loss (equation 2) functions were applied, along with categorical cross-entropy as a loss function to adjusted weights and biases, and their performance to training a multi-class classifier model was compared. The focal loss function is designed to address the class imbalance by downweighting easy classes such that their contribution to the total loss is small even if their number is large. Thus, it focuses on training a sparse set of hard classes.

$$\text{Focal Loss}(p_t) = -\alpha_t (1 - p_t)^\gamma \, log(p_t) \tag{1}$$

where p_t is the probability of a given class, $\gamma > 0$, and when $\gamma = 1$ focal loss works like cross-entropy loss function, and α range from [0, 1]. In this study, $\gamma = 2$ had the best performance and α was set to 1. Dice loss based on [40,41] has also been adapted as a loss function:

$$\text{Dice Loss}(Y; \, Y_p) = 1 - \frac{(2 \times Y \times Y_p + 1)}{(Y + Y_p + 1)} \tag{2}$$

where Y is labeled data and Y_p is predicted values. Adam (adaptive moment estimation) was selected as the optimizer, and similar to Feed Forward Neural network, early-stopping was used to prevent overfitting. In addition, fine-tuning was applied using the fine-tuning function in the TensorFlow library where the number of epochs varied from 100 to 1000 (epoch = 500 was the best one), and the learning rate (LR) from 10^{-6} to 10^0 (LR = 0.0001

was the best one). Note that the datasets were normalized into an array with an interval of [0, 1] so that the models can perform faster.

3. Model Evaluation

To evaluate the prediction performance and the trained models, the multiclass version of the F1-score, Intersection over Union or overall goodness-of-fit (IOU), and overall accuracy of each method were determined and compared. Model evaluation was carried out using the test image set against the ground truth images.

The F1-score based on [40,41] is identified as

$$\text{F1} = \frac{2 \times Precision \times Recall}{(Precision + Recall)} \tag{3}$$

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

$$\text{Recall} = \frac{TP}{(TP + FN)} \tag{5}$$

where *TP* is true positive prediction, *FP* is false positive prediction, and *FN* is false-negative prediction. To calculate the multiclass version of the F1-score, each class was considered individually as the positive class, with the other classes considered as negative classes, to get the binary F1-score for all classes. Then, the multiclass F1-score was calculated by averaging the binary F1-scores of all classes.

IOU evaluates the overlap of the predicted test image set and the ground truth images using the following equation:

$$\text{IOU} = \frac{Area\ of\ overlap}{Area\ of\ union} \tag{6}$$

In addition, Overall accuracy was calculated based on the total number of true predicted values over the total number of pixels.

4. Results and Discussion

4.1. Comparison of the Model's Performance and the Prediction Results

This section compares and contrasts the performance of all models. Figures 5 and 6 show the test performance of pixel-wise phase segmentation of all the models with 21 stacked 2D image slices (994 × 969 pixels) that were randomly chosen from the 3D X-ray CT image stack. All three machine learning methods (RF, FNN, and K-means) were trained and tested on the original dataset (OI), 14 extracted features (14F), and 64 extracted features (VGG16F). The latter was obtained from pre-trained filters of the convolutional layer of VGG16 architecture.
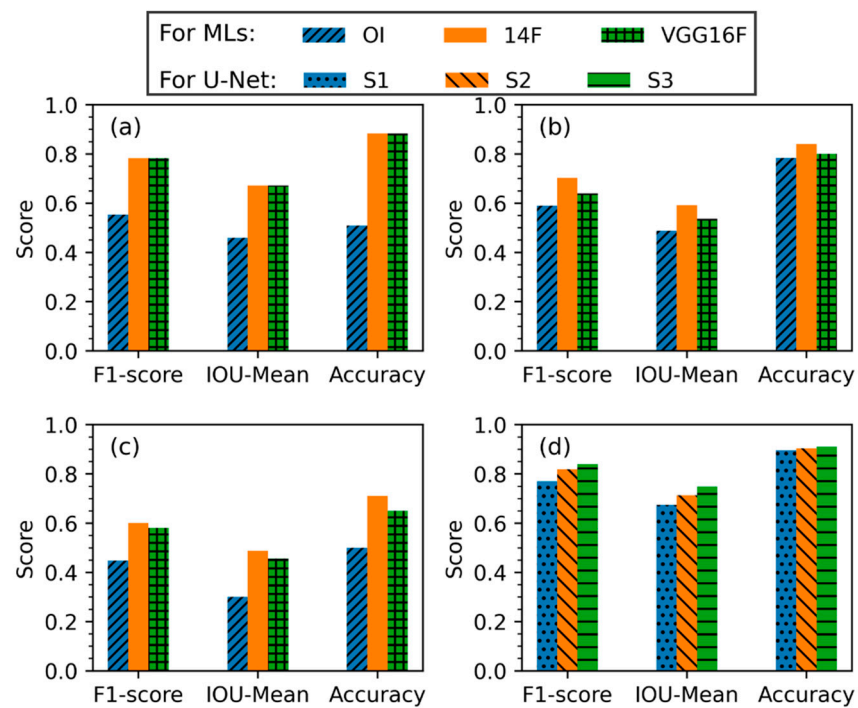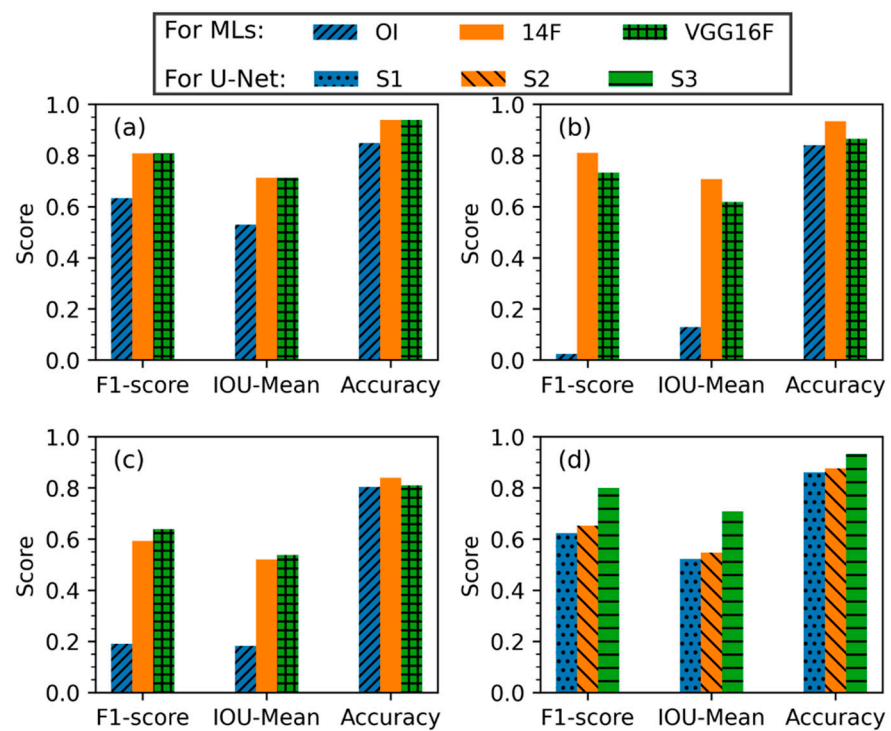
Figure 5. Overall performance of different machine learning methods and U-Net deep learning method on perdition of different phases in the Mancos dataset; (**a**) Overall performance of the RF model, (**b**) Overall performance of the FNN model, (**c**) Overall performance of the K-means model, (**d**) Overall performance of the U-Net model.

Figure 6. Overall performance of different machine learning methods and U-Net deep learning method on perdition of different phases in Marcellus dataset. (**a**) Overall performance of the RF model, (**b**) Overall performance of the FNN model, (**c**) Overall performance of the K-means model, (**d**) Overall performance of the U-Net model.

In addition, the U-Net model was trained with different loss functions (i.e., categorical cross-entropy, focal and dice losses). Focal and dice losses were added to improve the prediction of minority phases in addition to preserving the accuracy of the majority classes. In the first scenario (S1) for the U-Net model, loss and metric were set to the defaults of the original U-Net method (i.e., categorical cross-entropy as loss function and accuracy as a metric). For the second scenario (S2), the metric was set to mean IOU and F1-scores, and the loss function was set to focal loss. The linear combination of focal and dice loss (dice loss + 2 × focal loss) was set as a loss function for the third scenario (S3), and the metric was similar to S2. To select the best coefficient for focal and dice losses in the third scenario (S3), several coefficients were implemented using random search. The linear combination of loss functions (coefficient = 1 for dice loss and coefficient = 2 for focal loss) showed the highest IOU, F1-score, and accuracy. In addition, the class weights for dice loss were obtained based on the fraction of each class with respect to the total number of pixels, which reduced the influence of the imbalance dataset on the models' performance.

As shown in Figures 5 and 6, using only the original image (OI) as input, the accuracy of the machine learning models was less than that for feature-based machine learning image segmentation (i.e., 14F and VGG16F) for both Mancos (Figure 5) and Marcellus (Figure 6) datasets. Considering IOU and F1-score, which includes a weighted prediction precision for each class, reveals that using only the original image (OI) for phase segmentation is not reliable and including more input features might result in a more robust prediction. Figure 6 also indicates all machine learning models achieved a decent accuracy with 14F inputs outperforming the OI and VGG16F.

This higher performance for 14F can be attributed to the fact that many filters were examined visually, and the top 14 filters were selected for the segmentation task. These results demonstrate the importance of feature engineering and also selecting the top features for the machine learning classification tasks. The obtained results are in good agreement with the previous study conducted by [15] on X-ray images of several sandstone samples which showed the accuracy of machine learning methods was largely affected by feature vector selection and was improved by utilizing more features obtained from some filtering techniques.

Among machine learning models implemented on both datasets (Figures 5 and 6), the RF model outperformed other methods in terms of accuracy, IOU, and F1-scores. On the other hand, K-means achieved the lowest performance on this dataset since it is an unsupervised model which may not perform well on noisy data due to spatial variations in attenuation of a given phase, especially near the borders of the X-ray CT images. The U-Net (U-Net, S3) model achieved the most accurate results in general. This finding is in line with the previous study [24] that showed RF and U-Net methods had the best performance for a mineral segmentation task on SEM images of a shale sample compared to other machine learning models.

In addition, the results for the U-Net deep learning method using different metric and loss functions (S1, S2, and S3) are promising that are in good agreement with previous works using focal loss and dice losses [42,43]. The higher performance of the deep learning method compared to machine learning methods is that the deep learning pixel segmentation better identifies patterns in complex image datasets. It eventually uses these patterns to perform classification. Another reason is that it has a dynamic (online) process of learning and training for adjusting its feature weights and biases. In addition, compared to Feed Forward Neural network, U-Net utilizes deep layers which helps it to project data to a higher dimension (more features). As a result, this image segmentation technique is a more powerful tool for partitioning challenging image datasets, which may include a variety of artifacts and noises, as well as grayscale similarities which previously are difficult to segment by thresholding or any other traditional approaches.

In general, the results obtained from the comparison of different methods (Figures 5 and 6) show the U-Net model with the combination of focal and dice loss (S3) and the RF with 14 features (14F) outperform other methods in terms of accuracy (0.88 for RF and 0.91 for

U-Net on the Mancos dataset, and 0.94 for RF and 0.93 for U-Net on Marcellus dataset), IOU (0.67 for RF and 0.75 for U-Net on the Mancos dataset, and 0.71 for RF and 0.71 for U-Net on Marcellus dataset), and F1-score (0.78 for RF and 0.84 for U-Net on the Mancos dataset, and 0.81 for RF and 0.80 for U-Net on Marcellus dataset) on both test samples. This is in good agreement with previous studies [11,24]. In the study conducted by [11], the U-Net method that was applied on SEM images of Duvernay Shale samples showed an IOU of 0.9 indicating the power of the U-Net deep learning method for phase segmentation of core shale images over other machine learning methods.

Figures 7 and 8 show IOU scores for the perdition of different phases for the Mancos dataset and Marcellus datasets. In Figure 7, '0' is organic matter, '1' is the Kaolinite + Dolomite + Calcite group, '2' is background, '3' is Quartz + Illite/Smectite + Albite group, '4' is pyrite. In Figure 8, '0' is organic matter, '1' is the Muscovite + Microcline + Albite +Dolomite + Quartz group, '2' is background, '3' is Calcite + clay group, 4 is pyrite. The results show that major phases such as grains (mostly quartz), matrix (which was mostly calcite and clays), pyrite, and background were correctly detected in almost all cases. However, organic matter (class '1') was not predicted accurately due to misclassification with the background. Both background and organic matter classes share similar comprising features including black grayscale value with no specific pattern or texture.
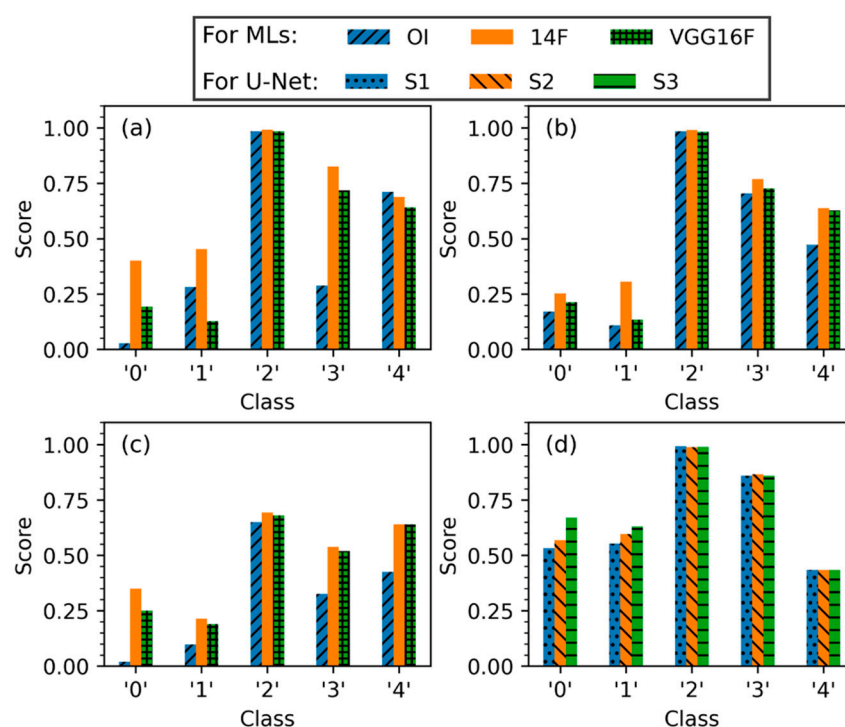

Figure 7. IOU score for the perdition of different phases for the Mancos dataset ('0' is organic matter, '1' is the 'Kaolinite + Dolomite + Calcite' group, '2' is background, '3' is 'Quartz + Illite/Smectite + Albite' group, '4' is pyrite). (**a**) IOU score for the RF model, (**b**) IOU score for the FNN model, (**c**) IOU score for the K-means model, (**d**) IOU score for the U-Net model.

As shown in Figures 7 and 8, considering extracted features (e.g., 14F and VGG16F) in addition to color attenuation improved samples' IOU score in almost all classes. The increase in IOU was as low as 4% or as high as 45%. This indicates the power of feature-based machine learning which is particularly valuable for a small dataset where applying deep learning methods and trainable filtering techniques is difficult or even impossible.
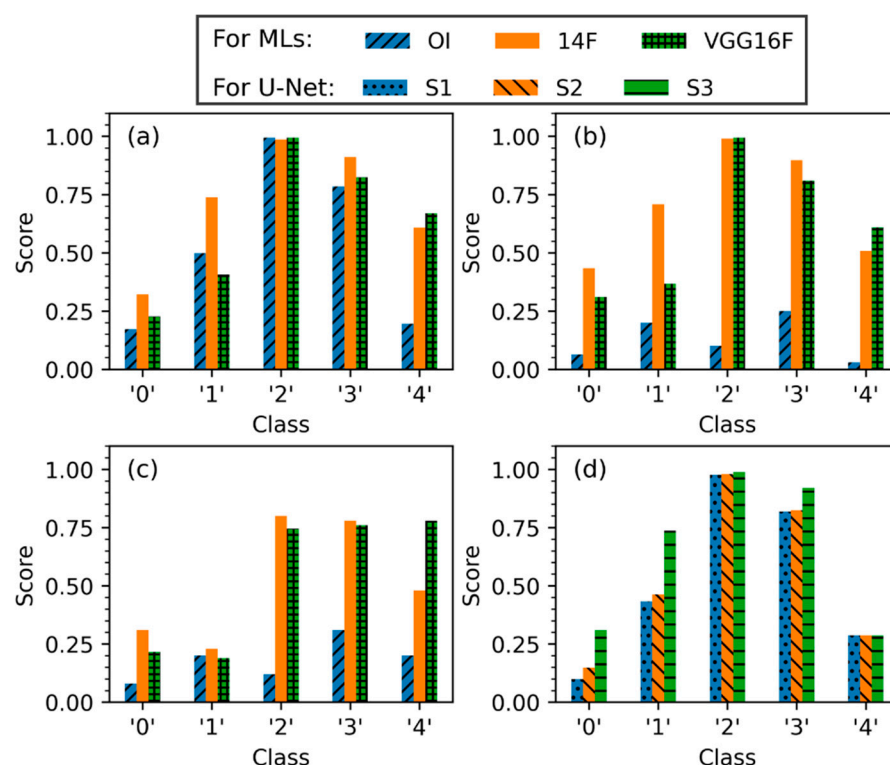
Figure 8. IOU score for the perdition of different phases for Marcellus dataset ('0' is organic matter, '1' is the 'Muscovite + Microcline + Albite +Dolomite + Quartz' group, '2' is background, '3' is 'Calcite + clay' group, '4' is pyrite). (**a**) IOU score for the RF model, (**b**) IOU score for the FNN model, (**c**) IOU score for the K-means model, (**d**) IOU score for the U-Net model.

Compared to the RF model with the original image (OI) and 14 features (14F), the predicted images by the U-Net model exhibited a better performance (less scatteredness, artifacts, and misclassification) in terms of prediction of different classes, especially for isolated small particles. It is because the U-Net model takes the spatial information of the input data along with attenuation (grayscale) information. The results are in good agreement with [11,24].

4.2. Applying Trained Machine Learning and U-Net Models on an Unseen Sample

In this section, an example of applying machine learning and U-Net deep learning models for phase segmentation of unseen X-ray CT images is discussed and visualization is provided. A total of 963,186 pixels were fed into the trained models to visually compare the results. The ground truth (labeled data) and predicted images are shown in Figure 9 (Marcellus) and Figure 10 (Mancos), respectively. In both figures, the label data is shown along with the output from the models. Only the top-performing methods are selected and shown for brevity, for example, for RF and FNN, the results with original images (OI) as input as well as VGG16 filters (VGG16F) as input are shown while for U-Net the scenario 3 (S3) is presented. As shown in these figures, the difference between simple and feature-based semantic segmentation is well demonstrated on visualization of pixel-wise semantic segmentation obtained from the RF and FNN models (Figures 9 and 10). It is clear that adding extra features to machine learning models improved the prediction results.
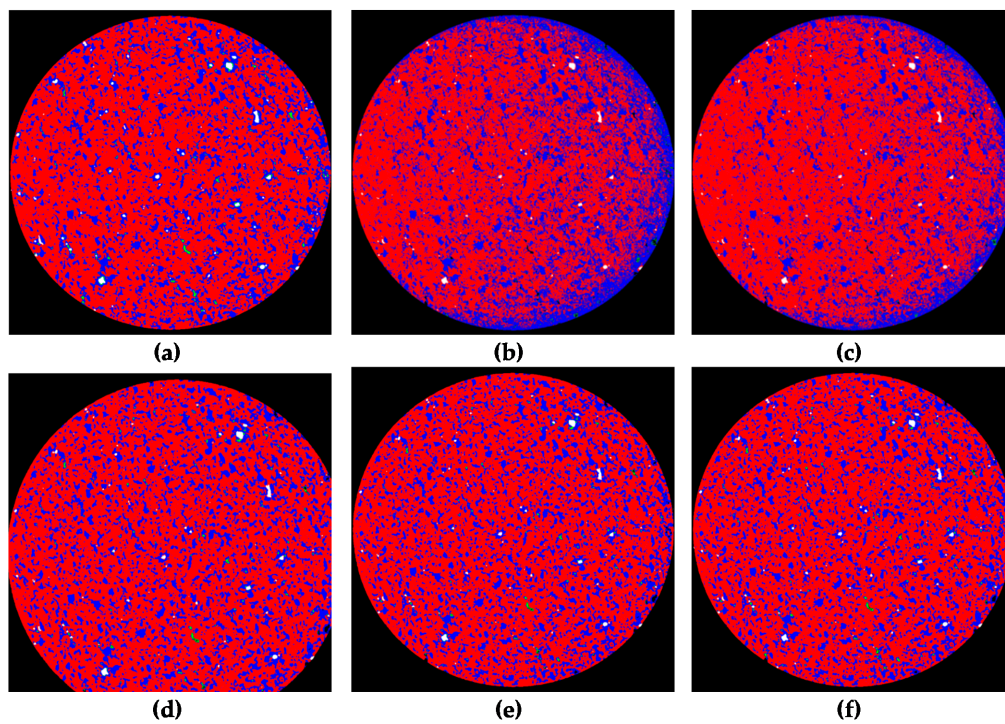
Figure 9. Output phase segmented images using ML methods with filtering techniques and deep learning method for Marcellus sample. (**a**) label image, (**b**) FNN model using original image as input (OI), (**c**) RF model using original image as input (OI), (**d**) U-Net considering focal and dice loss as the loss function (S3), (**e**) FNN model using 14 extracted features as input (14F), (**f**) RF model using 14 extracted features as input (14F).
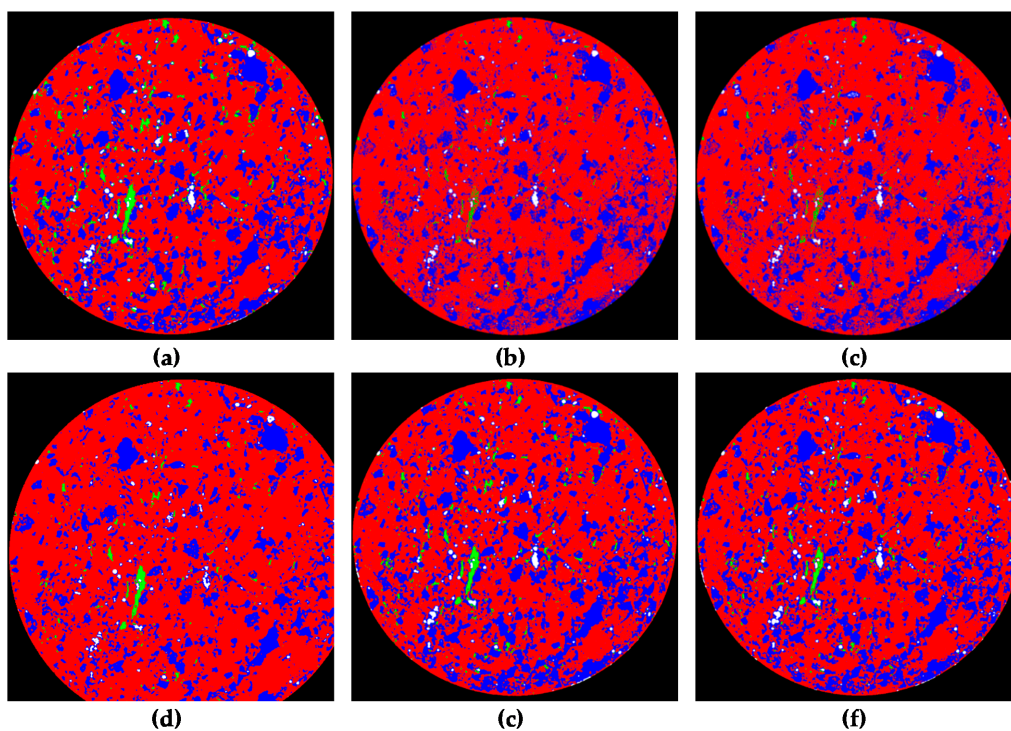

Figure 10. Output phase segmented images using machine learning methods with filtering techniques and deep learning method for the Mancos sample. (**a**) label image, (**b**) FNN model using original image as input (OI), (**c**) RF model using original image as input (OI), (**d**) U-Net considering focal and dice loss as the loss function (S3), (**e**) FNN model using 14 extracted features as input (14F), (**f**) RF model using 14 extracted features as input (14F).

In addition, results show that the machine learning with different filtering techniques removed artifacts (such as the blue ring in Figure 9) and improved prediction near the borders in both samples. As such, the special variation in color attenuation of a given phase which prevents a correct phase segmentation, especially on the edges of X-ray CT images, is not a prohibitive issue and can be addressed. The result of U-Net with Focal and dice loss showed the best segmentation among all methods which improved the reliability of prediction of minority classes and removed noise from the predicted segmented images. This is in good agreement with previous studies [42,43]. In general, a comparison of predicted results in both samples shows that filtering techniques and the deep learning method improve the predation of borders of each phase (edges).

4.3. Core Segmentation and Analysis

The X-ray CT image dataset of the Mancos and Marcellus core samples were processed and segmented using the best-trained model (U-Net, S3) with python programming language. Visualization of segmented pyrite, matrix, grains, organic matter, and background within the cores are shown in Figure 11 for both samples. As shown, the most prevalent component of the Marcellus sample is calcite (red color) which is confirmed by XRD data. In Mancos, the quartz-dominated matrix is the most dominant phase that also agreed with the XRD data.
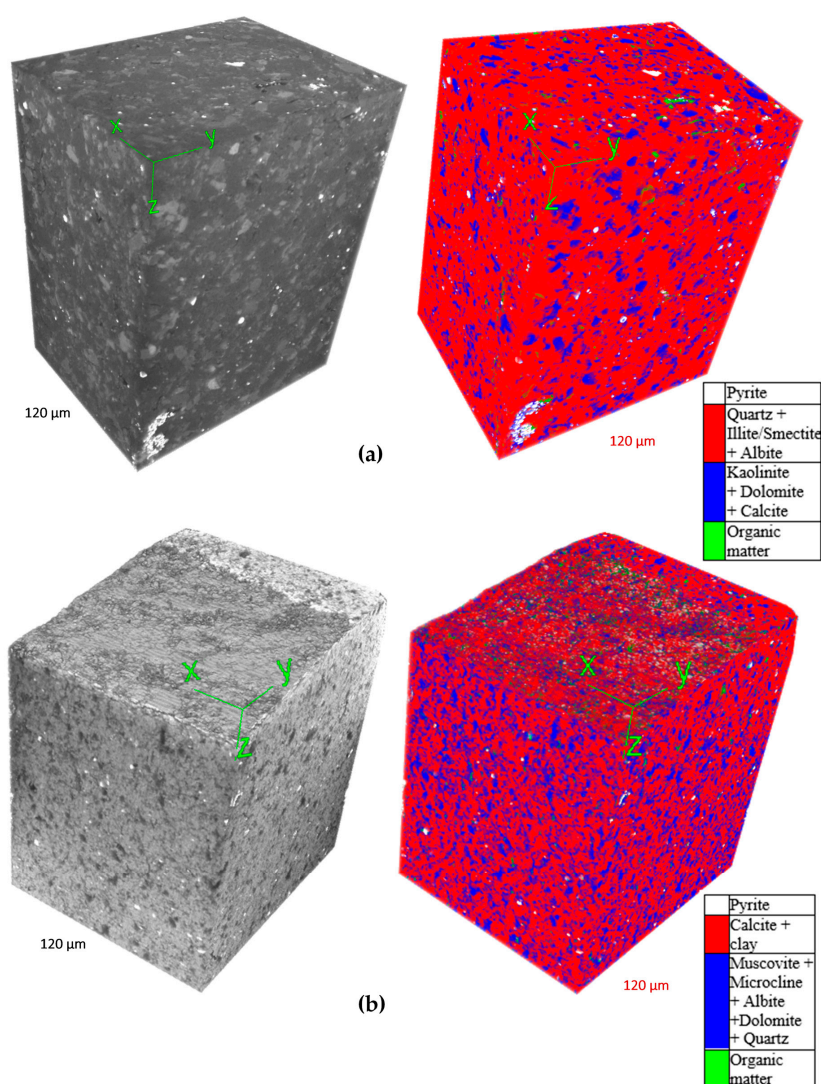

Figure 11. The 3D view of segmented core samples; (**a**) Mancos and (**b**) Marcellus samples.

Segmented images may be used to extract useful data such as grain size distributions or mineral volume fractions. Here, the grain size distribution analysis of granular minerals was performed to determine the relative size of the most abundant granular mineral phase (blue color in both samples) presented within both samples (Figure 12). The ImageJ plugin was used to calculate the area, size, and spatial information of the segmented blue phase within predicted images. Then, each grain diameter was calculated to plot the diameter distribution.

Figure 12. Grain size distribution of granular minerals within samples imaged with X-ray CT; (**a**) the distribution for Mancos sample, (**b**) the distribution for Marcellus sample.

Figure 12 depicts the grain size distribution of granular minerals (blue color in both samples) within both Mancos and Marcellus images and the predicted ones obtained from the RF and U-Net (S3) models. The analysis of grain distribution confirms the previous data and suggests although the predicted grain distribution of both methods is similar to the true grain distribution, the U-Net is more similar to ground truth data. It is particularly more evident for the Mancos sample.

4.4. Feature Importance

The RF model, which had the best performance among all machine learning methods, was used to determine the most important variables for pixel-wise phase segmentation

based on the "mean decrease in Gini Impurity". This metric was chosen due to its robustness in ranking the variable importance.

The rank of each variable at each dataset was obtained using Gini Impurity and the final variable importance ranking was calculated by averaging the results over both samples. Figure 13 shows the relative rank of variable importance based on a decrease in Gini Impurity for the classification task. The higher the number, the more important the filter. Higher numbers show a higher contribution to successful classification. As shown in the Figure, the Median filter had the highest importance followed by the original image and difference of the Gaussian filter. The feature importance results for VGG16F extracted features showed "filter 32" has the highest contribution to the classification tasks. VGG16 has 64 filters which are named from 1 to 64. In this figure, the most contributing filters are ranked and named from F-4 to F-32.
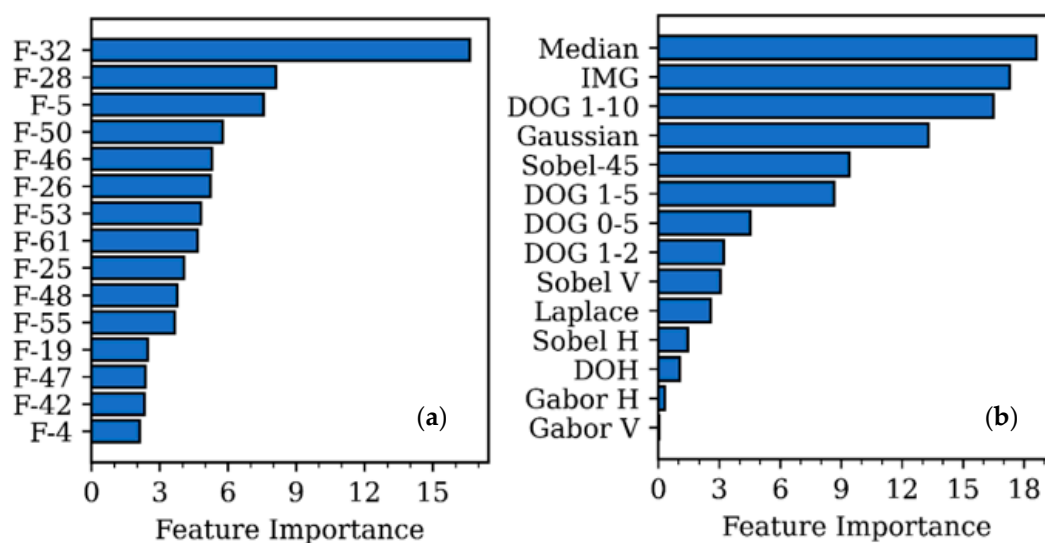

Figure 13. The rank of the extracted features is based on their contribution to improving segmentation (vertical variables are filter names, for example, F-32 is the 32nd filter of the second convolutional layer of VGG16, and horizontal numbers are the average rank of filters called feature importance value). (**a**) the rank for the VGG16F extracted features; (**b**) the rank for the 14F extracted features.

5. Conclusions

This research explores the feasibility of using machine learning approaches with feature extraction techniques for pixel-level phase segmentation of shales in 3D X-ray CT images. Once segmented, the categorized data could be used to retrieve useful information such as grain size distributions.

Based on the results for two different datasets, RF had the best accuracy among all applied machine learning methods due to its capability to handle imbalanced datasets and data scarcity. The feature-based RF model (14F-RF and VGG16F-RF) improved the segmentation results significantly for both samples since filtering techniques helped to find additional features to reliably segment different phases. Feature importance analysis showed the Median and Gaussian filters had the highest contribution in phase segmentation due to removing unwanted noise and providing more integrated phases.

The results from U-Net showed even higher performance compared to RF. Considering all three methods of evaluations (i.e., F1-score, IOU, and accuracy), the U-Net method has a better performance in predicting each class compared to all other methods which provide a more reliable phase segmentation in different sample types.

It was also shown that the loss function plays an essential role in determining the model performance in both Marcellus and Mancos samples. The results showed that,

for the complex objectives of mineral phase segmentation, it is not efficient to train the model only based on a universal loss function such as categorical cross-entropy which just monitors the overall loss since the majority class can directly affect it. In fact, for highly imbalanced segmentation, focal and dice losses, which are focus-based loss functions, work better as it minimizes the error based on each class as well as overall error. As a result, a minority class is less likely to be overwhelmed by a majority class.

Overall, it was shown that the U-Net deep learning model can outperform machine learning models for mineral phase segmentation and is the recommended method when a large dataset is available. This study will help geologists to obtain the different distinguishable phases in 3D X-ray CT images to provide practical techniques for reliable phase segmentation. The future work would be further distinguishing discrete mineral phases by training ML/DL methods using SEM images of a surface of a given sample as ground truth for the X-ray CT mineral segmentation task. The trained model then can be utilized to segment the complete stack of X-ray CT images into individual mineral phases.

Supplementary Materials: The following are available online at https://github.com/Parisa-Asadi/Integrating-machine-deep-learning-and-filtering-techniques-for-reliable-mineral-phase-segmentation/blob/main/README.md, Video: S1: A 3D View of Mancos Shale (View 1); S2: A 3D view of Mancos Shale (View 2); S3: A 3D view of Phase Segmented Mancos Shale; S4: A 3D view of Marcellus Shale (View 1); S5: A 3D view of Marcellus Shale (View 2); S6: A 3D view of Phase Segmented Marcellus Shale.

References

1. Yang, H.; Zhao, Y.; Zhang, X.; Liu, G.; Du, X.; Shang, D.; Yu, Y.; Chen, J.; Wang, H.; Tu, H. Supercritical CO2 fracturing with different drilling depths in shale. *Energy Sources Part A Recovery Util. Environ. Eff.* **2021**, 1–20. [CrossRef]
2. Yakaboylu, G.A.; Gupta, N.; Sabolsky, E.M.; Mishra, B. Mineralogical characterization and strain analysis of the Marcellus shales. *Int. J. Rock Mech. Min. Sci.* **2020**, *130*, 104345. [CrossRef]
3. Brunhoeber, O.M.; Anovitz, L.M.; Beckingham, L.E. Role of mineralogy in controlling fracture formation. *ACS Earth Space Chem.* **2021**. under review.
4. Yoon, H.; Ingraham, M.D.; Grigg, J.; Rosandick, B.; Mozley, P.; Rinehart, A.; Mook, W.M.; Dewers, T. Impact of Depositional and Diagenetic Heterogeneity on Multiscale Mechanical Behavior of Mancos Shale. *Am. Assoc. Pet. Geol. Mex. Utah USA* **2019**, 121–148. [CrossRef]
5. Spokas, K.; Fang, Y.; Fitts, J.P.; Peters, C.A.; Elsworth, D. Collapse of reacted fracture surface decreases permeability and frictional strength. *J. Geophys. Res. Solid Earth* **2019**, *124*, 12799–12811. [CrossRef]
6. Iloejesi, C.O.; Beckingham, L.E. Assessment of geochemical limitations to utilizing CO_2 as a cushion gas in compressed energy storage systems. *Environ. Eng. Sci.* **2021**, *38*, 115–126. [CrossRef] [PubMed]
7. Goral, J.; Andrew, M.; Olson, T.; Deo, M. Correlative core-to pore-scale imaging of shales. *Mar. Pet. Geol.* **2020**, *111*, 886–904. [CrossRef]

8. Guntoro, P.I.; Ghorbani, Y.; Koch, P.H.; Rosenkranz, J. X-ray microcomputed tomography (μCT) for mineral characterization: A review of data analysis methods. *Minerals* **2019**, *9*, 183. [CrossRef]

9. Bensinger, J.; Beckingham, L.E. CO_2 storage in the Paluxy formation at the Kemper County CO_2 storage complex: Pore network properties and simulated reactive permeability evolution. *Int. J. Greenh. Gas. Control* **2020**, *93*, 102887. [CrossRef]

10. Guntoro, P.I.; Tiu, G.; Ghorbani, Y.; Lund, C.; Rosenkranz, J. Application of machine learning techniques in mineral phase segmentation for X-ray microcomputed tomography (μCT) data. *Miner. Eng.* **2019**, *142*, 105882. [CrossRef]

11. Chen, Z.; Liu, X.; Yang, J.; Little, E.; Zhou, Y. Deep learning-based method for SEM image segmentation in mineral characterization, an example from Duvernay Shale samples in Western Canada Sedimentary Basin. *Comput. Geosci.* **2020**, *138*, 104450. [CrossRef]

12. Landrot, G.; Ajo-Franklin, J.B.; Yang, L.; Cabrini, S.; Steefel, C.I. Measurement of accessible reactive surface area in a sandstone, with application to CO2 mineralization. *Chem. Geol.* **2012**, *318*, 113–125. [CrossRef]

13. Qin, F.; Beckingham, L.E. Impact of image resolution on quantification of mineral abundances and accessible surface areas. *Chem. Geol.* **2019**, *523*, 31–41. [CrossRef]

14. Anjikar, I.S.; Wales, S.; Beckingham, L.E. Fused Filament Fabrication 3-D Printing of Reactive Porous Media. *Geophys. Res. Lett.* **2020**, *47*, e2020GL087665. [CrossRef]

15. Chauhan, S.; Rühaak, W.; Anbergen, H.; Kabdenov, A.; Freise, M.; Wille, T.; Sass, I. Phase segmentation of X-ray computer tomography rock images using machine learning techniques: An accuracy and performance study. *Solid Earth* **2016**, *7*, 1125–1139. [CrossRef]

16. Cnudde, V.; Boone, M.N. High-resolution X-ray computed tomography in geosciences: A review of the current technology and applications. *Earth Sci. Rev.* **2013**, *123*, 1–17. [CrossRef]

17. Wildenschild, D.; Sheppard, A.P. X-ray imaging and analysis techniques for quantifying pore-scale structure and processes in subsurface porous medium systems. *Adv. Water Resour.* **2013**, *51*, 217–246. [CrossRef]

18. Suthaharan, S. Machine learning models and algorithms for big data classification. *Integr. Ser. Inf. Syst.* **2016**, *36*, 1–12.

19. Ulker, E.; Sorgun, M. Comparison of computational intelligence models for cuttings transport in horizontal and deviated wells. *J. Pet. Sci. Eng.* **2016**, *146*, 832–837. [CrossRef]

20. Gupta, I.; Rai, C.; Sondergeld, C.; Devegowda, D. Rock Typing in Eagle Ford, Barnett, and Woodford Formations. *SPE Reservoir Eval. Eng.* **2018**, *21*, 654–670. [CrossRef]

21. Gupta, I.; Rai, C.; Sondergeld, C.; Devegowda, D. Rock typing in the upper Devonian-lower Mississippian woodford shale formation, Oklahoma, USA. *Interpretation* **2018**, *6*, SC55–SC66. [CrossRef]

22. Saikia, P.; Baruah, R.D.; Singh, S.K.; Chaudhuri, P.K. Artificial Neural Networks in the domain of reservoir characterization: A review from shallow to deep models. *Comput. Geosci.* **2020**, *135*, 104357. [CrossRef]

23. Semnani, S.J.; Borja, R.I. Quantifying the heterogeneity of shale through statistical combination of imaging across scales. *Acta Geotech.* **2017**, *12*, 1193–1205. [CrossRef]

24. Li, C.; Wang, D.; Kong, L. Application of Machine Learning Techniques in Mineral Classification for Scanning Electron Microscopy-Energy Dispersive X-Ray Spectroscopy (SEM-EDS) Images. *J. Pet. Sci. Eng.* **2021**, *200*, 108178. [CrossRef]

25. Anderson, T.I.; Vega, B.; Kovscek, A.R. Multimodal imaging and machine learning to enhance microscope images of shale. *Comput. Geosci.* **2020**, *145*, 104593. [CrossRef]

26. Karimpouli, S.; Tahmasebi, P. Segmentation of digital rock images using deep convolutional autoencoder networks. *Comput. Geosci.* **2019**, *126*, 142–150. [CrossRef]

27. Da Wang, Y.; Shabaninejad, M.; Armstrong, R.T.; Mostaghimi, P. Physical accuracy of deep neural networks for 2d and 3d multi-mineral segmentation of rock micro-CT images. *arXiv* **2020**, arXiv:2002.05322.

28. Gihm, Y.S.; Hwang, I.G.; Kim, H.T.; Lee, H.S.; Lee, D.S. Geological Characteristics and Development Strategy of the Marcellus Shale. *J. Korean Soc. Mineral. Energy Resour. Eng.* **2011**, *48*, 371–382.

29. Hawkins, S.J.; Charpentier, R.R.; Schenk, C.J.; Leathers-Miller, H.M.; Klett, T.R.; Brownfield, M.E.; Finn, T.M.; Gaswirth, S.B.; Marra, K.R.; Le, P.A.; et al. *Assessment of Continuous (Unconventional) Oil and Gas Resources in the Late Cretaceous Mancos Shale of the Piceance Basin, Uinta-Piceance Province, Colorado and Utah*; United States Geological Survey: Reston, VA, USA, 2016; pp. 2016–3030.

30. Bradski, G.; Kaehler, A. *Learning OpenCV: Computer Vision with the OpenCV Library*; O'Reilly Media, Inc.: Newton, MA, USA, 2008.

31. Van der Walt, S.; Schönberger, J.L.; Nunez-Iglesias, J.; Boulogne, F.; Warner, J.D.; Yager, N.; Gouillart, E.; Yu, T. scikit-image: Image processing in Python. *Peer J.* **2014**, *2*, e453. [CrossRef]

32. Ronneberger, O.; Fischer, P.; Brox, T. U-net: Convolutional networks for biomedical image segmentation. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Munich, Germany, 5–9 October 2015; pp. 234–241.

33. Dhanachandra, N.; Manglem, K.; Chanu, Y.J. Image segmentation using K-means clustering algorithm and subtractive clustering algorithm. *Procedia Comput. Sci.* **2015**, *54*, 764–771. [CrossRef]

34. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Duchesnay, E. Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.

35. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]

36. Feurer, M.; Springenberg, J.; Hutter, F. Initializing bayesian hyperparameter optimization via meta-learning. In Proceedings of the AAAI Conference on Artificial Intelligence, Austin, TX, USA, 25–30 January 2015; Volume 29.

37. Heaton, J. Chapter 12: Pruning and Model Selection. In *AIFH, Volume 3: Deep Learning and Neural Networks*; Heaton Research, Inc.: St. Louis, MO, USA, 2015; pp. 195–204.
38. Hart, E.; Sim, K.; Kamimura, K.; Meredieu, C.; Guyon, D.; Gardiner, B. Use of machine learning techniques to model wind damage to forests. *Agric. For. Meteorol.* **2019**, *265*, 16–29. [CrossRef]
39. Abadi, M.; Barham, P.; Chen, J.; Chen, Z.; Davis, A.; Dean, J.; Devin, M.; Ghemawat, S.; Irving, G.; Isard, M.; et al. Tensorflow: A system for large-scale machine learning. In Proceedings of the 12th USENIX Symposium on Operating Systems Design and Implementation, Savannah, GA, USA, 2–4 November 2016; pp. 265–283.
40. Dice, L.R. Measures of the amount of ecologic association between species. *Ecology* **1945**, *26*, 297–302. [CrossRef]
41. Sorensen, T.A. A method of establishing groups of equal amplitude in plant sociology based on similarity of species content and its application to analyses of the vegetation on Danish commons. *Biol. Skar.* **1948**, *5*, 1–34.
42. Yeung, M.; Sala, E.; Schönlieb, C.B.; Rundo, L. Unified Focal loss: Generalising Dice and cross entropy-based losses to handle class imbalanced medical image segmentation. *arXiv* **2021**, arXiv:2102.04525.
43. Wang, L.; Wang, C.; Sun, Z.; Chen, S. An Improved Dice Loss for Pneumothorax Segmentation by Mining the Information of Negative Areas. *IEEE Access* **2020**, *8*, 167939–167949. [CrossRef]