Journal of Engineering Research and Reports

Volume 26, Issue 11, Page 190-202, 2024; Article no.JERR.125433
ISSN: 2582-2926

A Computer Aided System for Reservoir Rock Classification

Abdunaser O. Susi [a*], Elhusain S. Saad [b], Ahmed M. Daoub [a] and Rajab N. Abuajila [a]

[a] Petroleum Engineering Department, Collage of Engineering, Misurata University, Libya.
[b] Electrical and Electronics Engineering Department, Collage of Engineering, Misurata University, Libya.

Authors' contributions

This work was carried out in collaboration among all authors. All authors read and approved the final manuscript.

Article Information

DOI: https://doi.org/10.9734/jerr/2024/v26i111324

Open Peer Review History:
This journal follows the Advanced Open Peer Review policy. Identity of the Reviewers, Editor(s) and additional Reviewers, peer review comments, different versions of the manuscript, comments of the editors, etc are available here: https://www.sdiarticle5.com/review-history/125433

Original Research Article

Received: 03/09/2024
Accepted: 05/11/2024
Published: 08/11/2024

ABSTRACT

This study is mainly aimed to know the reservoir rocks among core specimens that are obtained during the stage of drilling of exploration oil and gas wells by the processing of the image digitally using a Computer-Aided System.

This computerized method has been conducted using two programming languages; Python and Matlab. The Python program is used to segment rock images and classify the extracted features, whereas the MATLAB language has been used to extract intensity features from segmented images. The database used in this project has been collected from two sources, one of them from the Sarier formation (the biggest Libyan oil field) while the second source was from the internet. The images used in this study are sedimentary rocks images, which include sandstone, limestone, and dolomite.

**Corresponding author: Email: naser_o_72@eng.misuratau.edu.ly;*

Cite as: *Susi, Abdunaser O., Elhusain S. Saad, Ahmed M. Daoub, and Rajab N. Abuajila. 2024. "A Computer Aided System for Reservoir Rock Classification". Journal of Engineering Research and Reports 26 (11):190-202. https://doi.org/10.9734/jerr/2024/v26i111324.*

As a result, this study has reached that the system, which was used here is capable of classifying rocks and can be used to classify oil rocks or any other type of rocks. Also, it can be used near the well due to certain conditions such as the absence of experts due to the lack of enough database, the study has only introduced the preliminary results as proof of the concept.

Keywords: Drilling; exploration; computer-aided; sedimentary; rock; image; sarier.

1. INTRODUCTION

There are many definitions of rocks, according to (Ehlers and Blatt, 1997). He defines rocks as stuff that the earth is made of. A more acceptable scientific definition of rocks is that; a rock is a naturally occurring solid cohesive aggregate of one or more mineral or mineral materials. Rocks are widely classified into three groups as it shown in the Fig. 1; based on their formation processes (Mibei 2014).

Those three major rock types are Igneous rock; are those that have formed by the cooling and crystallization of magma, either at the Earth's surface or within the crust (University of Auckland 2005). The second one is Sedimentary rocks, which have formed when eroded particles of other rocks have been deposited (on the ocean floor, stream/lake beds, etc.) and compacted, or by the precipitation of minerals/mineralogist from water (University of Auckland 2005). The third one is Metamorphic rocks, which are those that have formed when existing rocks have undergone pressure and/or temperature changes so that their original mineralogy has been changed (University of Auckland 2005).

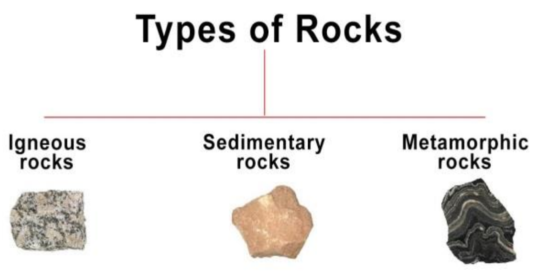

Fig. 1. The types of rock

1.1 Sedimentary Rocks

Sedimentary rocks are formed by the deposition of material at the Earth's surface and (or) within bodies of water. Sedimentation is the collective name for processes that cause mineral and/or organic particles (detritus) to settle and accumulate or minerals to precipitate from a solution. Sediments can be detrital, chemical or organic. Detrital sediments are mechanically eroded from pre-existing rocks. Chemical sediments on the other hand are fluid precipitates or evaporates deposited in various environments. Sedimentary rocks are important in regard to resources like limestone deposits, coal and oil. They are also important geologically in interpretation of earth's history (Carlson et al, 2009). Below is the descriptions of some common sedimentary rocks, which include; limestone, sandstone, and dolomite.

Types of sedimentary rocks:

Limestone: Limestone is a sedimentary rock as it appears in the Fig. 2; composed primarily of calcite, a calcium carbonate mineral with a chemical composition of CaCo3. It usually forms in clear, calm, warm, shallow marine waters (King n.d.). Limestone is usually a biological sedimentary rock, forming from the accumulation of shell, coral, algal, fecal, and other organic debris. It can also form by chemical sedimentary processes, such as the precipitation of calcium carbonate from lake or ocean water (King n.d.).

Fig. 2. Limestone rock

Dolomite: Dolomite, also known as "dolostone" and "dolomite rock," is a sedimentary rock as it appears in the Fig. 3; composed primarily of the mineral dolomite, CaMg (Co3)2. Dolomite is found in sedimentary basins worldwide. It is thought to be formed by the post depositional alteration of lime mud and limestone by magnesium-rich groundwater (King n.d.).

Dolomite and limestone are very similar rocks. They share the same color ranges of white-to-gray and white-to-light brown (although other colors such as red, green, and black are possible). They are approximately the same hardness, and they are both soluble in dilute hydrochloric acid. They are both crushed and cut for use as construction materials and used for their ability to neutralize acids (King n.d.).

Fig. 3. Dolomite rock

1.2 Sandstone

Sandstone is a sedimentary rock composed of sand-size grains of mineral, rock, or organic material. It also contains a cementing material that binds the sand grains together and may contain a matrix of silt- or clay-size particles that occupy the spaces between the sand grains.

Sandstone is one of the most common types of sedimentary rock, and it is found in sedimentary basins throughout the world. Deposits of sand that eventually form sandstone are delivered to the basin by rivers, but may also be delivered by the action of waves or wind. Some sand grains might be organic particles, such as sand and shell debris produced within the basin (King n.d.).

Fig. 4. Sandstone Rock

2. TYPES OF SAND GRAINS

The grains in a sandstone can be composed of mineral, rock, or organic materials. Which and in what percentage depends upon their source and how they were altered during transport.

Mineral grains in sandstones are usually quartz. Sometimes the quartz content of these sands can be very high - up to 90% or more. These are sands that have been worked and reworked by wind or water and are said to be "mature." Other sands can contain significant amounts of feldspar, and if they came from a source rock with a significant quartz content, they are said to be "immature" (King n.d.).

3. PETROLEUM GEOLOGY

The study of origin, occurrence, movement, accumulation, and exploration of hydrocarbon fuels. It refers to the specific set of geological disciplines that are applied to the search for hydrocarbons (oil exploration) (https://en.wikipedia.org/wiki/Petroleum_geology)

4. MAJOR SUB DISCIPLINES IN PETROLEUM GEOLOGY

Several major sub disciplines exist in petroleum geology specifically to study the seven key elements discussed above (https://en.wikipedia.org/wiki/Petroleum_geology)

Exploration stage: Although a basin analysis is usually part of the first study a company conducts prior to moving into an area for future exploration, it is also sometimes conducted during the exploration phase. Exploration geology comprises all the activities and studies necessary for finding new hydrocarbon occurrence. Usually seismic (or 3D Seismic) studies are shot, and old exploration data (seismic lines, well logs, reports) are used to expand upon the new studies. Sometimes gravity and magnetic studies are conducted, and oil seeps and spills are mapped to find potential areas for hydrocarbon occurrences. As soon as a significant hydrocarbon occurrence is found by an exploration- or wildcat-well the appraisal stage starts (https://en.wikipedia.org/wiki/Petroleum_geology)

a. Source rock analysis

Maturation of source rocks depends strongly on temperature, such that the majority of oil generation occurs in the 60° to 120 °C range. Gas generation starts at similar temperatures, but may continue up beyond this range, perhaps as high as 200°C. In order to determine the likelihood of oil/gas generation, therefore, the thermal history of the source rock must be calculated. This is performed with a combination of geochemical analysis of the source rock (to determine the type of kerogens present and their maturation characteristics) and basin modelling methods, such as back-stripping, to model the thermal gradient in the sedimentary column (https://en.wikipedia.org/wiki/Petroleum_geology)

b. Basin Analysis

A full scale basin analysis is usually carried out prior to defining leads and prospects for future drilling. This study tackles the petroleum system and studies source rock (presence and quality); burial history; maturation (timing and volumes); migration and focus; and potential regional seals and major reservoir units (that define carrier beds). All these elements are used to investigate where potential hydrocarbons might migrate towards. Traps and potential leads and prospects are then defined in the area that is likely to have received hydrocarbons (https://en.wikipedia.org/wiki/Petroleum_geology).

c. Reservoir analysis

The existence of a reservoir rock (typically, sandstones and fractured limestones) is determined through a combination of regional studies (i.e. analysis of other wells in the area), stratigraphy and sedimentology (to quantify the pattern and extent of sedimentation) and seismic interpretation. Once a possible hydrocarbon reservoir is identified, the key physical characteristics of a reservoir that are of interest to a hydrocarbon explorationist are its bulk rock volume, net-to-gross ratio, porosity and permeability (https://en.wikipedia.org/wiki/Petroleum_geology)

d. Appraisal Stage

The Appraisal stage is used to delineate the extent of the discovery. Hydrocarbon reservoir properties, connectivity, hydrocarbon type and gas-oil and oil-water contacts are determined to calculate potential recoverable volumes. This is usually done by drilling more appraisal wells around the initial exploration well. Production tests may also give insight in reservoir pressures and connectivity. Geochemical and petro physical analysis gives information on the type (viscosity, chemistry, API, carbon content, etc.) of the hydrocarbon and the nature of the reservoir (porosity, permeability, etc.) (https://en.wikipedia.org/wiki/Petroleum_geology)

Production stage: After a hydrocarbon occurrence has been discovered and appraisal has indicated it is a commercial find the production stage is initiated. This stage focuses on extracting the hydrocarbons in a controlled way (without damaging the formation, within commercial favorable volumes, etc.). Production wells are drilled and completed in strategic positions. 3D seismic is usually available by this stage to target wells precisely for optimal recovery. Sometimes enhanced recovery (steam injection, pumps, etc.) is used to extract more hydrocarbons or to redevelop abandoned fields (https://en.wikipedia.org/wiki/Petroleum_geology).

5. MATERIALS AND METHODS

5.1 Materials

Digital image: A digital image is an image composed of picture elements, also known as pixels as it appears in the Fig. 5, each with finite, discrete quantities of numeric representation for its intensity or gray level that is an output from its two-dimensional functions fed as input by its spatial coordinates denoted with x, y on the x-axis and y-axis, respectively. Depending on whether the image resolution is fixed, it may be of vector or raster type (Gonzalez et al. 2018).

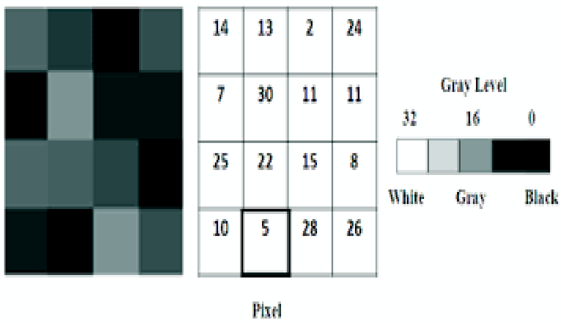

Fig. 5. Digital image

Image processing and rock classification: In recent studies in 2019, there is research for a transformative learning method for automatic identification of sandstone microscopic images by shedding light and feature identification images(https://doi.org/10.1016/j.cageo.2017.03.007).

Classification of sandstone microscopic images is an essential task in geology, and the classical method is either subjective or time-consuming. Computer aided automatic classification has been proved useful, but it seldom considers the situation where sandstone images are collected from separated regions. In this paper, we provide a method called Festra, which uses transfer learning to handle the problem of interregional sandstone microscopic image classification (Frucci et al. 2008). The method contains two parts: one is feature selection, which aims to screen out features having great difference between the regions, the other is instance transfer using an enhanced Trada Boost, whose object is to mitigate the difference among thin section images collected from the regions (Frucci et al. 2008). Experiments are conducted based on the sandstone images taken from four regions in Tibet to study the performance of Festra. The experimental results have proved both effectiveness and validity of Festra, which provides competitive prediction performance on all the four regions, with few target instances labeled suitable for the field use (Frucci et al. 2008).

And another recent study showing the classification of rocks from the field image spots It is analyzed using a deep convolutional neural network, which consists of seven layers.

Deep learning is receiving significant research attention for pattern recognition and machine learning. Its application here has effectively identified rock types from images captured in the field. This paper proposes an accurate approach for identifying rock types in the field based on image analysis using deep convolutional neural networks. The proposed approach can identify six common rock types with an overall classification accuracy of 97.96%, thus outperforming other established deep-learning models and a linear model (Frucci et al. 2008).

The results show that the proposed approach based on deep learning represents an improvement in intelligent rock-type identification and solves several difficulties facing the automated identification of rock types in the field (Frucci et al. 2008).

Related image processing:

Techniques: In this section we will try to explain three digital i.mage processing techniques involved in our work to classify rock images. These techniques are image segmentation, feature extraction, and feature classification.

Image segmentation using:

Active Contour Models (ACM): The original image of the sample, Fig. 6a. And Image segmentation, Fig. 6b is the process of partitioning a digital image into multiple image segments, also known as image regions or image objects (sets of pixels). The goal of segmentation is to simplify and/or change the representation of an image into something that is more meaningful and easier to analyze. Image segmentation is typically used to locate objects and boundaries as in Figure (2.2) objects (lines, curves, etc.) in images.

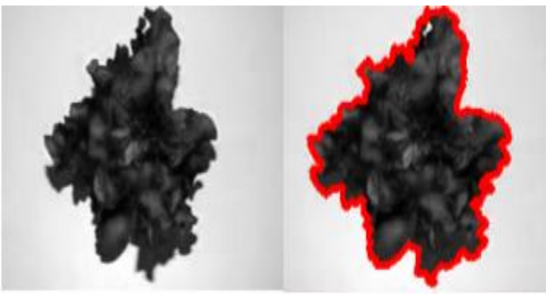

Fig. 6.

a) Original Image [9] **b) Final boundary of object [9]**

More precisely, image segmentation is the process of assigning a label to every pixel in an image such that pixels with the same label share certain characteristics (Kass et al. 1988).

Active Contour Models (ACM) or snakes, which were used in this study are deformable contours that have been used in many image analysis applications, including the image-based tracking of rigid and nonridged objects.

In I their basic forms, the mathematical formulation draws from the theory of optimal approximation involving a functional. A traditional snake is a parametric contour embedded in the image plane. The contour is represented as a parametric curve.

ACM also known as snake or deformable contour as shown in figure.

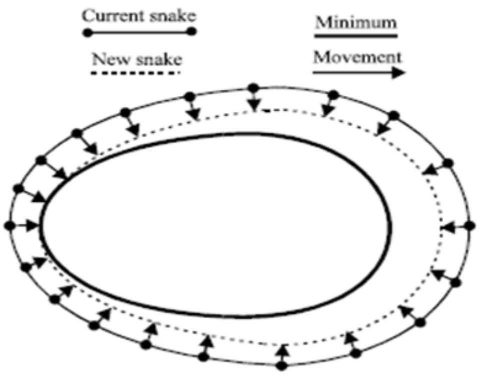

Fig. 7. Active contour or snake (https://developers.google.com/machine-learning/crash-course/classification/roc-and-auc#roc-curve)

The description of the deformable contour model can be explained by using energy functional.

The energy functional: The idea of the deformable contours that there is an energy function associated with each contour shape. The image contour can be detected by getting the minimum energy function which is a sum of several terms that corresponding to a force work on the contour. Assume, a contour C = C(S) where S is the arc length. The energy function is:

$$\mathcal{E} = \int (\alpha (s)\, E_{cont} + \beta (s)\, E_{curv} + \gamma (s)\, E_{image})\, ds \qquad (1)$$

Where the parameters α, β and γ control the relative influence of the corresponding energy term (Hardie et al. 2008).

Each term in the energy function serves a different purpose. The terms E_{cont} and E_{curv} encourage continuity and smoothness of the deformable contour or works as internal energy, respectively while E_{image} accounts for edge detection, dragging the contour toward the closest image edge or works as external energy (Hardie et al. 2008).

i. Continuity Term: The continuity is:

$$E_{cont} = \left| \frac{dc}{ds} \right|^2 \qquad (2)$$

The contour c is replaced by a chain of N image points, P1……PN, So that:

$$\mathbf{E_{cont}} = (\overline{\boldsymbol{d}} - \left| \boldsymbol{P_i} - \boldsymbol{P_{i-1}} \right|)^2 \qquad (3)$$

Where $\bar{d}$ is the average distance between (P_i, P_{i-1}), and when $\left|\boldsymbol{P_i} - \boldsymbol{P_{i-1}}\right| >> \bar{d}$ it can be approximated to:

$$\mathbf{E_{cont}} = \left| \boldsymbol{P_i} - \boldsymbol{P_{i-1}} \right|^2 \qquad (4)$$

ii. Smoothness Term:

The goal of the smoothness term is to avoid oscillations of the deformable contour, The goal of the smoothness term is to avoid oscillations of the deformable contour, so it smoothed the curvature by the second derivative of the contour:

$$E_{cont} = \left| P_{i-1} - 2P_i + P_{i+1} \right|^2 \qquad (5)$$

iii. Edge attraction Term:

Also, it is called the external force which is achieved by the spatial gradient of the intensity image I, computed at each snake point:

$$E_{image} = -\left| \nabla I \right|^2$$

E_{image} becomes very small (negative) whenever the norm of the spatial gradient is large (near images edges), so it makes " small and attracting the snake towards image contours, and it unlike E_{cont} and E_{curv} depends only on the contour not on its derivative (Hardie et al. 2008).

Feature extraction: In this section, we will try to introduce the concept of feature extraction in image processing. These features can be extracted from the area of interest which here might be a segmented object. These features can be divided in general into three categories: Intensity, gradient, and geometric features. In this project, we have been studying some intensity features.

Intensity Features are features describing intensity values of area of interest. In this project, we have considered twelve intensity features. Some of the density features that were used in this project are shown in the Table (1) (Rock Classification from Field Image Patches Analyzed Using a Deep Convolutional Neural Network).

The maximum, minimum and average of intensity values found within the rock shape are among these extracted features (Rock Classification from Field Image Patches Analyzed Using a Deep Convolutional Neural Network).

These are some of the definitions, equations, and density features that have been used:

I. Mean, $\overline{X}$:

Is the average score and this property is calculated by the following equation:

$$\bar{X} = \frac{\sum_{i=1}^{n} x_i}{n}, \tag{6}$$

Where: X_i: is each individual data point in the set.
n: is the total number of data points.

II. Standard Deviation (s):

Standard deviation is the square root of the variance, this property is calculated by the following equation:

$$s = \left[\frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n-1}\right]^{\frac{1}{2}} \tag{7}$$

Where:

$\overline{X}$: is the Mean
X_i: is each individual data point in the set.
n: is the total number of data points.

III. Contrast:

The Contrast of the square of the standard deviation, Contrast = (standard deviation)² and in symbols:

$$\text{Variance } (\sigma^2) = \sigma \times \sigma \tag{8}$$

IV. Skew, SK**:**

The Skew is the fourth central torque divided by four standard deviations.

$$Sk = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^3}{s^3(n-1)} \tag{9}$$

Where:

S: Standard Deviation
$\overline{X}$: is the Mean
X_i: is each individual data point in the set.
n: is the total number of data points.

V. Kurtosis:

The kurtosis is the third centripetal moment divided by the cube of the standard deviation.

$$Sk = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^4}{s^4(n-1)} \tag{10}$$

VI. Moment of order eq. (6 to 10):

Moments are of great importance in the study of curves of frequency distributions in terms of the degree of deviation.

Its flatness and its contribution to the generation of some measures of central tendency and some measures of dispersion.

The central moments of order K are calculated relative to the arithmetic mean and are denoted by the symbol (M_k), and one of the following relationships is known:

The K-rank equation for the raw data:

$$\boldsymbol{M}_k = \frac{\sum(x_i - \bar{x})^k}{n} \tag{11}$$

Table 1. Showing features intensity (Rock Classification from Field Image Patches Analyzed Using a Deep Convolutional Neural Network)

No.	Features Intensity
1	Mean
2	Contrast
3	Standard deviation
4	Skew
5	Kurtosis
6	Moment 1
7	Moment 2
8	Moment 3
9	Moment 4
10	Moment 5
11	Moment 6
12	Moment 7

Classification: The classification of objects in images is one of the most important branches in computer vision techniques. One of these common classifiers is Support Vector Machine which has been selected as our classifier in this project.

Support Vector Machine (SVM): Supportive Vector Machine Algorithm is one of the controlled machine learning algorithms that is used to analyze data for statistical classification or regression analysis. The algorithm starts from the sorted data, often the encoding is using only two categories, the data either belongs to the first category or to the second category, this kind of classifiers called a binary classifier (Haweil 2021).

Supporting vectors is the best hyperplane equation that tries to separate the two categories as explained in Fig. 8.

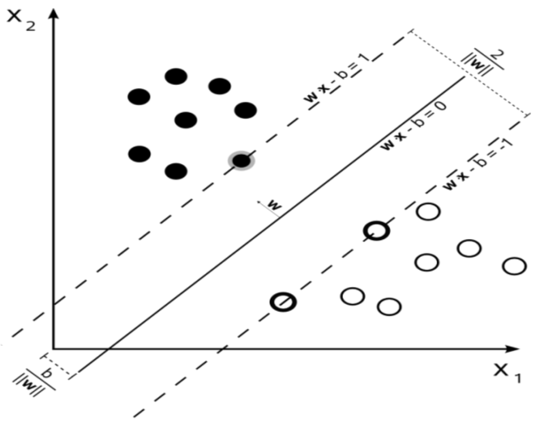

Fig. 8. Supportive vector machine

With a normalized or standardized dataset, these hyperplanes can be described by the equations

$$w^T x - b = 0 \tag{12}$$

The parameter $\frac{b}{|w|}$ determines the offset of the hyperplane from the origin along the normal vector w And $w^T x - b = 1$ (Anything on or above this boundary is of one class, with label -1)

And $w^T x - b = -1$ (Anything on or below this boundary is of the other class, with label -1) Geometrically, the distance between these two hyperplanes is $\frac{2}{|w|}$, so to maximize the distance between the planes one wants to minimize $|w|$ (Haweil 2021).

Measure the performance of CAS Using Roc curve: An ROC curve (receiver operating characteristic curve) is a graph showing the performance of a classification model at all classification thresholds. This curve plots two parameters (Kaggle. n.d.).

- True Positive Rate.
- False Positive Rate.

True Positive Rate (**TPR**) is a synonym for recall and is therefore defined as follows:

$$TPR = \frac{TP}{TP + FN} \tag{13}$$

False Positive Rate (**FPR**) is defined as follows:

$$FPR = \frac{FP}{FP + TN} \tag{14}$$

Table 2. Highlights the terminology related to true and false positive and negative (Kaggle. n.d.)

True positive (TP)	A test result that correctly indicates the presence of a condition or characteristic.
True negative (TN)	A test result that correctly indicates the absence of a condition or characteristic.
False positive (FP)	A test result which wrongly indicates that a particular condition or attribute is present.
False negative (FN)	A test result which wrongly indicates that a particular condition or attribute is absent.

An ROC curve plots TPR vs. FPR at different classification thresholds. Lowering the classification threshold classifies more items as positive, thus increasing both False Positives and True Positives. The following figure shows a typical ROC curve (https://en.wikipedia.org/wiki/Confusion_matrix).

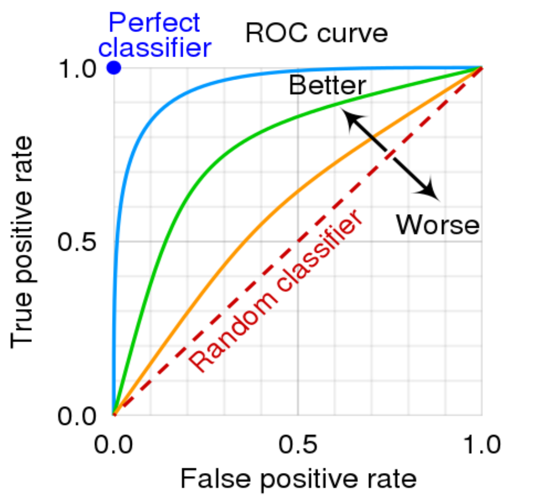

Fig. 9. Curve plots TPR vs. FPR [15]

as the best possible ROC curve, as it ranks all positives above all negatives. It has an AUC of 1.0.

AUC: Area Under the ROC Curve stands for "Area under the ROC Curve." That is, AUC measures the entire two-dimensional area underneath the entire ROC curve (think integral calculus) from (0,0) to (1,1) (https://en.wikipedia.org/wiki/Support-vector_machine**)**.

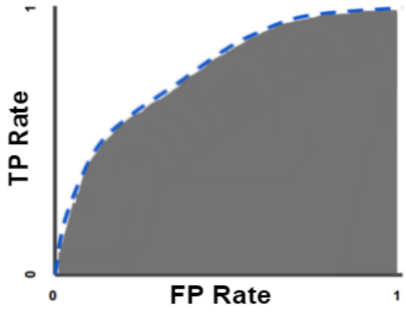

Fig. 10. AUC ROC Curve (https://en.wikipedia.org/wiki/Support-vector_machine)

5.2 Methods

In this chapter, the stages of the rock identification system will be explained in this chapter. The system has been implemented using two programming languages Python and Matlab. by specialized expert who ran the used programs to classify the rocks.

The python program used to segment rock images and classify the extracted features. The MATLAB language has been used to extract intensity features from segmented images.

Data image: The database used in this project has been collected from two sources, one of them is from the Sarier formation by Geological Instructor, while the second database source is from the internet (Stockman 2011).

The database used in this project is not enough because in order to get solid results from the system, we need huge number of data (at least 1000 images) while the total number of collected images is just about 50 images. Most of images used in this project is sedimentary rocks images, which include sandstone, limestone and dolomite. The following figures represent samples of these images:

Sandstone:

Fig. 11 Sandstone (Stockman 2011)

Limestone:

Fig. 12. Limestone (Stockman 2011)

Dolomite:

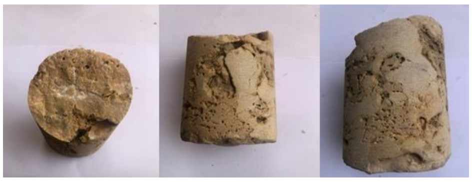

Fig. 13. Dolomite (Hardie et al. 2008)

Block diagram of Computer Aided System (CAS) for Rock Identification: In this section, we will introduce the system that was used to classify the rock type. The Fig. 14 shows the block diagram of a computer system that identifies the rock image by extracting the features from segmented rock area and classifies them.

The study tries to introduce the implemented system by explaining each block of this figure.

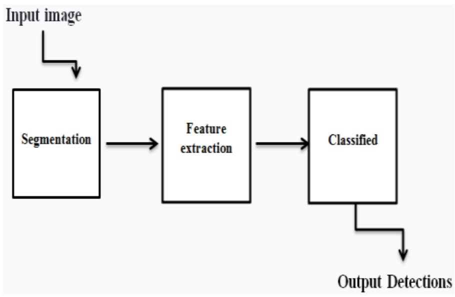

Fig. 14. Block Diagram of the implemented system used in the classification of the rocks

6. RESULTS AND DISCUSSION

Segmentation: The Input image of the segmentation is a grayscale image of the rock, and the output is the binary image which is called the mask of the rock (e.g. 1 for rock pixels and zeros for pixels outside the rock boundary). The following figures show the input and output of the segmentation stage. An image of Sandstone was entered into the ACM method, and the output from it is the boundaries of the rock body as in Figs. 15,16.

Image segmentation is a fundamental component of computer vision technologies and algorithms, playing a crucial role in improving the efficiency and effectiveness of image analysis (Kass et al. 1988).

In this study, we utilize one of the most effective image segmentation techniques known as Active Contours, or "snakes." The Active Contour model is a segmentation method that separates pixels of interest by using energy forces and constraints for further processing and analysis (https://developers.google.com/machine-learning/crash-course/classification/roc-and-auc#AUC). It operates by minimizing an energy function that combines two components: one derived from the image and the other related to the shape of the contour, including its length and smoothness. The energy minimization occurs implicitly with respect to shape and explicitly with respect to image features (Frucci 2008).

The primary application of Active Contours in image processing is to define smooth, closed contours around regions of interest in an image. This method is particularly useful for identifying and delineating uneven or complex shapes within images (Frucci 2008). Usually, the energy function is composed of:

1. Internal Energy: This term ensures that the contour remains smooth and adheres to some regularity constraints. It usually consists of two components: Stretching Energy: Penalizes the length of the curve, encouraging it to remain close to its initial shape. Bending Energy: Penalizes changes in curvature, promoting smoothness.
2. External Energy: This term pulls the contour towards features in the image that are of interest, such as edges or other prominent structures. It measures how well the curve fits the desired features, guiding it to the relevant structures based on the image data.

Finally, the resulting energy is the total of these internal and external energies, and the contour evolves to minimize this total energy, balancing smoothness with fitting the image features (Hardie et al. 2008).

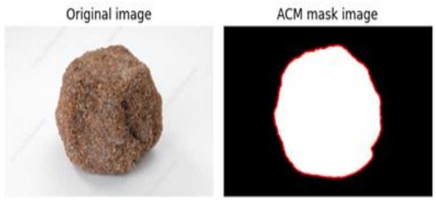

Fig. 15. An example of ACM mask for selected Sandston image

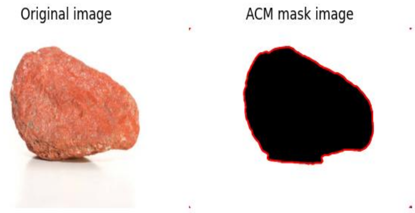

Fig. 16. an example of ACM mask for selected Sandston image

An image has been entered from Limestone for segmentation by ACM method, and the output from it is the boundaries of the rock body as Figs. 17,18.

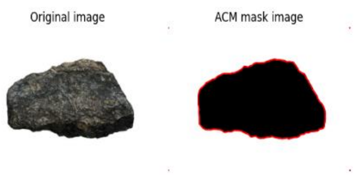

Fig. 17. An example of ACM mask for selected Limestone image

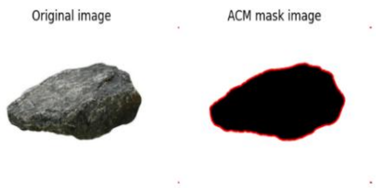

Fig. 18. An example of ACM mask for selected Limestone image

It is clear from the previous figures that the ACM technique work very well to get the right boundary of the rocks. Because the description of the rock using the feature extraction technique depends on the segmentation output, if the area of interest is segmented accurately, then the features will be extracted exactly.

Features extraction: Next to the segmentation stage is the feature extraction stage. The input of this stage is just the area of interest (the segmented rock) and the output is a vector consisting of teen intensity features.

E.g. : V= [Mean, Contrast, Standard deviation, Skew, Kurtosis, Moment, eq. (6 to 11)].

It extracts a vector from each image. These extracted features are computed from the area surrounded by the boundary using the ACM algorithm. Then data are stored in a file with labeling.

Training data: To distinguish rocks and properly separate the taxa, the training data must be of high accuracy, and the size of the database must be sufficient to obtain the best classification results.

Testing Data: This experiment needs some data for testing because of the need to measure the performance of the whole system. After dividing the data, the study needs to train the classifier with training data. The classification algorithm is a supervised learning technique that is used to identify the category of new observations on the basis of training data. In classification, a program learns from the given dataset or observations and then classifies new observation into a number classes or groups, The program learns from the given dataset or observations and then classifies new observation into a number of classes or groups, such as Yes or No, 0 or 1, sandstone or No because it uses a binary classifier.

The extracted features will be arranged as an array with columns and rows:

$$\mathbf{V} = \begin{bmatrix} a_0 & a_1 & a_2 & a_3.... & a_n \\ b_0 & b_1 & b_2 & b_3.... & b_n \\ c_0 & c_1 & c_2 & c_3.... & c_n \\ : & : & : & : & : \\ n_0 & n_1 & n_2 & n_3.... & n_n \end{bmatrix}_{12\times n}$$

Where each column of V matrix $[a_i \ b_i \ \ c_i ... n_i]_{1\times 12}^{T}$ is the intensity feature vector of i^{th} image (so the number of V rows represents the number of intensity features used in this study) and number of columns represent the number of segmented images (number of features). Then we need to label each column of V to train the classifier $L = [1 \quad 1 \quad 0 \ ....0]_{1\times n}$ which refers to the label of each column of the V matrix, its value is assigned during classification either 0 or 1 depending on which kind of rock we are interested in classifying (for example we are interested in classifying sandstone rock, thus each sandstone image will assign **1** and all other types of rocks will take **0**), we follow this procedure because our classifier is a binary classifier. Two-thirds of the data has been assigned for training data while one-third has been assigned for testing.

It applies the training data to SVM and gets the weights of the classifier which tries to separate, for example, sandstone from all other types of rocks. A similar way has been applied to other types of rocks.

Measure the performance of the system: Here, the study will measure the performance of the implemented system using the ROC curve, Also the study apply the testing data to the classifier and get the values of true detection and false detection. By changing the weights of the SVM classifier, and get another equation that try to classify the sandstone from all other types of rock, and for this weights, and it can get another values of true and false probability of detection. Thus, the study can get more values of probability of true and false detection by changing the weights of the classifier. Then it can plot these values of probability of false and true detection to get the ROC curve.

Because the data is not enough for classification as it is mentioned early. So, the study needs a large number of rocks, about 1000 images, thus the following ROC figure does not give us the right performance, but an approximate form to show the end of the performance work as proof of concept. If this study applied more data, then the performance will be more accurate.

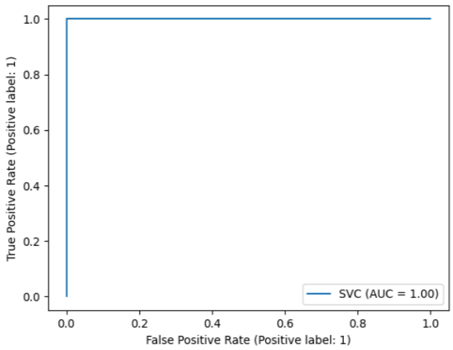

Fig. 19. ROC curve to classify the rock using SVM classifier

7. CONCLUSIONS

From all the above; results and discussion; the study made the following conclusions:

1. This system can be used near the well under certain conditions such as the absence of experts or labs.
2. The similarity of the structure of rocks is anywhere in the world, so the program can be used anywhere with any rock dataset.

8. RECOMMENDATIONS

In this study, it can make the following recommendations:

1. It is recommended that provide a huge amount of datasets (Rock samples) to get an accurate classification and better performance.
2. The input image can be enhanced by using some local contrast techniques which give better segmentation boundaries.
3. Use other types and methods of image segmentation.
4. Collect more efficient features which can better classification.

DISCLAIMER (ARTIFICIAL INTELLIGENCE)

Author(s) hereby declare that NO generative AI technologies such as Large Language Models (ChatGPT, COPILOT, etc) and text-to-image generators have been used during writing or editing of this manuscript.

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES

(n.d.). *Rock classification from field image patches analyzed using a deep convolutional neural network*. Retrieved from [URL]

Frucci, M., & Di Baja, G. S. (2008). From segmentation to binarization of gray-level images. *Journal of Pattern Recognition Research, 3*(1), 1–13.

Gonzalez, R., Woods, R. E., & Eddins, S. L. (2018). *Digital image processing* (3rd ed.). Pearson. ISBN 978-0-13-335672-4.

Hardie, R. C., Rogers, S. K., Wilson, T., & Rogers, A. (2008). Performance analysis of a new computer aided detection system for identifying lung nodules on chest radiographs. *Medical Image Analysis, 12*(3), 240–258.

Hardie, R. C., Rogers, S. K., Wilson, T., & Rogers, A. (2008). Performance analysis of a new computer aided detection system for identifying lung nodules on chest radiographs. *Medical Image Analysis, 12*(3), 240–258.

Haweil, H. (2021). A group of photos from Sariyer field, Libya. Misurata University, Petroleum Engineering Department, College of Engineering.

https://developers.google.com/machine-learning/crash-course/classification/roc-and-auc#roc-curve

https://developers.google.com/machine-learning/crash-course/classification/roc-and-auc#AUC

https://doi.org/10.1016/j.cageo.2017.03.007

https://en.wikipedia.org/wiki/Confusion_matrix

https://en.wikipedia.org/wiki/Petroleum_geology

https://en.wikipedia.org/wiki/Support-vector_machine

Kaggle. (n.d.). *Rock classification*. Retrieved from https://www.kaggle.com/salmaneunus/rock-classification

Kass, M., Witkin, A., & Terzopoulos, D. (1988). Snakes: Active contour models. *International Journal of Computer Vision, 1*(4), 321–331.

King, H. M. (n.d.). *Photos of common clastic, chemical, and organic sedimentary rock types*. Retrieved from https://www.geology.com

Mibei, G. (2014). *Introduction to types and classification of rocks*. Geothermal Development Company 1374011.

Scialert. (n.d.). *Fulltext article*. Retrieved from https://scialert.net/fulltext/?

Stockman, G., & Shapiro, L. G. (2001). *Computer vision*. Prentice Hall PTR.

University of Auckland. (2005). *Rocks and minerals: Rock types*. Retrieved from https://flexiblelearning.auckland.ac.nz/rocks_minerals/rocks/index.html

Peer-review history:
The peer review history for this paper can be accessed here:
https://www.sdiarticle5.com/review-history/125433