Seismic Structures Classification Using Novel Features from Seismic Images

Ghadah Alhabib
Department of Computer Science, Prince Mohammad Bin Fahd University, Khobar, Saudi Arabia.
202000093@pmu.edu.sa

Jaafar Alghazo
Department of Electrical and Computer Engineering, Virginia Military Institute, Lexington, VA, USA.
alghazojm@vmi.edu

Ghazanfar Latif
Computer Science Department, Prince Mohammad Bin Fahd University, Khobar, Saudi Arabia.
glatif@pmu.edu.sa

Ghassen Ben Brahim
Department of Computer Science, Prince Mohammad Bin Fahd University, Khobar, Saudi Arabia.
gbrahim@pmu.edu.sa

***Abstract*— Seismic facies can be used as novel features to classify different classes of seismic structures. Classification of seismic structure is beneficial for mineralogy, grain size approximation, the permeability of deposition units, and the identification of areas of interest. To extract features of seismic images, the following extraction methods were used: Discrete Wavelet Transform Features, Discrete Cosine Transform Features, Discrete Fourier Transform Features, and Gabor Features. The classification methods being considered are Support Vector Machine (SVM), Random Forest (RF), Fast Decision Trees (FDT), and Naïve Bayes (NB). The proposed study uses the LANDMASS database, composed of two datasets, LANDMASS-1, with 17,667 images, and LANDMASS-2, with 4,000 images. The datasets contain seismic images of four different classes of seismic structures; Chaotic, Fault, Horizon, and Salt Dome. The outcome of this study proves that the combination of Forest Tree classification method and the Discrete Cosine Transform Features extraction method achieved the highest accuracy, which was around 94.17% – higher than that achieved considering similar methods reported in the extant literature.**

Keywords*— *Seismic Facies, Seismic Features, Seismic Classification, Seismic Structures.

I. Introduction

Seismic structure classification refers to the interpretation and examination of seismic facies from seismic reflector information. The fundamental characteristics used to distinguish seismic facies are interval velocity, frequency, amplitude, lateral continuity, and bedform internal and external configuration [1]. The classification of structures supports the identification of areas of interest and the approximation of grain size, mineralogy, and permeability of deposition units. Since 1950, high-resolution seismic images produced from seismic reflection imaging have been greatly used in oil, gas, and water exploration [2]. Seismic imaging is the set of methods that project an intense sound source into the ground, receive it back through geophones using observed seismograms as inputs, and then record it to examine subsurface conditions and distinguish high concentrations of contamination. Recordings of signals are then processed into images of the geologic structure. The resulting digital model of the subsurface may be used to detect preferential flow paths, examine the placement and screening of wells, identify dense contaminants, and select a remediation technology.

There are two different types of seismic images, reflected waves and refracted waves [3]. Reflected waves move downward, hit a geologic boundary, and bounce back to the surface. On the contrary, refracted waves turn at a layer of a rock or soil surface and pass it before traveling back to the surface. They present more subsurface detail but may be challenging to interpret when multiple echoes exit.

Seismic facies classification is interpreted and automated using machine learning algorithms and the interpretation of seismic data is achieved through fault interpretation, horizon interpretation, chaotic interpretation, and salt dome interpretation [4]. The datasets used in this research include these four distinct seismic structure classes [5]. A horizon is a thin bed with a characteristic fossil or lithology content or a bedding surface with a marked change in lithology within a sequence of volcanic rocks or sedimentary. It represents a stratigraphic surface, either lithostratigraphic or chronostratigraphic. Examples include volcanic eruptions, tsunamis, meteorite impacts, and ice ages. When seismic data is interpreted, horizons are the change in seismic velocity and density across a boundary between two layers of rock, called reflectors, which are identified on individual profiles. A horizon is a matrix of samples to be drawn on a map and saved in a 3-column ASCII file.

A chaotic horizon is an unstratified reflection geometry of high amplitude reflections that possess discontinuousness with other reflectors in a single unit. Unstratified facies are not deposited in layers. It has a higher bulk density, shear strength, and resistivity.

A fault is a fracture or a group of fractures between two sections of rocks. Faults permit the relative movement of two blocks and may vary in length from kilometers to millimeters. A fault is divided into a fault core indicating the displacement of the rocks and a surrounding damage zone. Fault zone identification greatly depends on wave frequency. Zones with low wave frequency are broad, offering minimum information about facies distribution. However, zones with high wave frequencies characterize the fault volume and its internal structure from seismic attributes.

A salt dome is a type of structural dome created through diapirism, in which salt or some type of evaporite minerals is thrust into overlying rocks. Salt domes can form structures like salt plugs, which are salt sheets that intrude from the top of the dome, and salt welds, which appear when an extensive amount of salt restricts the growth of a dome and the contacts above and below merge. Seismic reflection identifies the stark density contrast between the salt and its surrounding sediment. Salt domes are depressed blocks of crust surrounded by parallel normal faults that can be neighbored by reverse faults.

This work attempts to tackle the problem of classification of various classes of seismic structures, which brings many benefits in various research domains such as mineralogy, grain size approximation, the permeability of deposition units, the identification of areas of interest, etc. The different classes of seismic classes that are considered are: Chaotic, Fault, Horizon, and Salt Dome. Different machine learning techniques will be applied on existing LANDMASS dataset following the extraction of novel features using the Discrete Wavelet Transform, Discrete Cosine Transform, Discrete Fourier Transform, and Gabor Features.

The rest of the paper is organized as follows. First, we present a review of most of the existing work related to the application of machine learning techniques on seismic imaging classification. Then, the used approach in terms of dataset description, feature extraction techniques being used, and classification algorithms being applied are detailed. Section 4, discusses the achieved performance along with results analysis. Finally, in Section 5 conclusions are drawn.

II. Recent Studies

In [6], the author proposes using seismic imaging to locate salt bodies and salt deposits. The training dataset includes 101x101 pixels 4000 seismic image patches and segmentation masks, and the testing dataset includes 18,000 seismic images. The proposed method consists of a convolutional neural network (CNN) used for semantic segmentation and the proposed architecture of the network is based on the U-Net approach combined with ResNet, DenseNet, and U-Net model and consists of 5 convolutional layers with a kernel size of 3x3. In [7], the authors propose using image analysis techniques like the histogram of oriented gradients (HOG) for seismic image segmentation and geological analysis. The methodology identifies HOG features, derives statistical parameters related to the texture attributes, and then integrates the selected images to separate the targeted section. It also considers a new hybrid texture attribute based on HOG parameters. It relies on two geometrically synthetic models and two field seismic data examples of salt body and mud diapirs. The salt body showcases a chaotic pattern with the unsharp boundary and the mud diapirs present a mild chaotic pattern with thin marginal flows and interdigitated boundary. The method can separate distinct seismic patterns and geological objects with any shape.

In [8], the authors propose using machine learning algorithms and advanced CNN image processing techniques to interpret faults and salt domes for hydrocarbon exploration. In [9], the authors use a machine learning based novel multiscale attention convolutional neural network (MACNN) for fault detection on seismic images for improved geologic fidelity. The methodology uses a multiscale spatial-channel attention mechanism to join and improve encoder feature maps of distinct spatial resolutions and to produce higher resolution and higher fidelity fault maps. The developed architecture allows MACNN to improve the contextual information fixed in the maps. MACNN proved to have higher fidelity fault and higher resolution maps compared to conventional CNNs.

In [10], the authors propose a method that uses CNN to identify fault zones from 3D seismic images using a 2-step method: training and prediction. In the training step, the CNN model accepts seismic images and identifies fault or non-fault points. Afterward, in the prediction step, the model computes faults probabilities at every position in the image and does not necessitate precomputed attributes to extract the faults. The datasets used are synthetic and field datasets. In [11], the authors suggest using seismic attributes to detect salt bodies from migrated images through seismic facies classification and fault detection. The study develops a novel deep learning method based on Se-ResNet and U-Net. The model was examined using K-Fold cross-validation. In addition, the dataset is composed of 101x101 pixel images and masks which are later transformed into NumPy Array for computation. The training of these images was divided into two steps: the first 300 epochs model was trained using Dice loss and BCE and the second 300 epochs were trained using the Lovasz Loss function.

In [12], researchers develop PDCNN, which is a patch-based denoising method using deep CNN for seismic images. It heavily relies on patch clustering and joint denoising. The methodology clusters overlapping patches of noisy photos into K classes and thus, each class contains image patches with close noise levels. Then, a model from a set of trained CNN models is designated for each class. The dataset includes both synthetic and field seismic images. In [13], the authors recommend using an end-to-end DNN seismic inversion network (SeisInvNets) with novel components to interpret seismic data. The spatially aligned feature maps are used to build velocity models from the enhanced seismic traces. To test the methodology, a dataset with 12,000 pairs of seismic data and a velocity model are used.

In [14], the authors propose using CNN deep learning to extract Paleokarst Collapse features in 3D seismic images. The workflow produces 3D training pairs including synthetic seismic images and label images of the collapsed paleokarst features simulated in the seismic images. The methodology identifies 3D paleokarst systems and computes their geometric parameters. In [15], the authors propose using a deep CNN workflow for fault recognition of 3D seismic images. The study open-sources the Thebe dataset and fault labels. It uses edge detection networks to detect faults and identify areas of high fault probability. The workflow uses image processing methods, the fault discretization algorithm, and PaleoScan for fault analysis.

In the next section, the proposed methodology is detailed. This includes description of the used datasets, feature extraction techniques being used, and classification algorithms being applied.

III. Methodology

The proposed methodology makes use of a dataset composed of seismic images with 4 different classes, which are: Horizons, Salt Domes, Chaotic Horizons, and Fault. These images act as an input to 4 feature extraction methods, namely: Discrete Wavelet Transform (DWT) Features, Discrete Cosine Transform (DCT) Features, and Discrete Fourier Transform (DFT) Features, to extract the features which will be used during the classification phase. Four classification methods are considered, which are: Support Vector Machine, Random Forest, Fast Decision Trees, and Naïve Bayes. In the proposed approach, performance results (in terms of accuracy, precision, recall, and F1 measure) for various combinations of feature extraction methods and classification algorithms are collected and compared. The best possible combination determines the outcome of the paper.

A. Expiremental Dataset

This research is conducted on a dataset obtained from the School of Electrical and Computer Engineering (ECE) at the Georgia Institute of Technology. Georgia Tech and KFUPM have collaborated to produce datasets at the Center for Energy and Geo-Processing (CeGP) for mineral research purposes. The following database is named LANDMASS [15] (LArge North-Sea Dataset of Migrated Aggregated Seismic Structures) and was extracted from the North Sea F3 block under the Creative Commons license (CC BY-SA 3.0) using a specific distance measure. This database contains two datasets, LANDMASS-1 and LANDMASS-2. The first dataset comprises 17,667 images of size 99x99 pixels and a normalized value range of [-1, 1] with 9,385 Horizon patches, 1,891 Salt Dome patches, 5,140 chaotic patches, and 1,251 Fault patches. The second dataset comprises 4,000 images of size 150x300 pixels and a normalized value range of [0,1] with 1000 patches for each class.

The dataset contains four classes of seismic structures. The images of these four different classes were verified and photos with outliers or more than one structure were eliminated. The examined classes are the following.

1. Horizons: It mostly includes smooth horizons and may include sigmoidal-like structures.
2. Salt Domes: Salt dome structures.
3. Chaotic Horizons: Horizons with a chaotic-like or noisy appearance.
4. Faults: Images/patches contain one or more faults with varying significance.

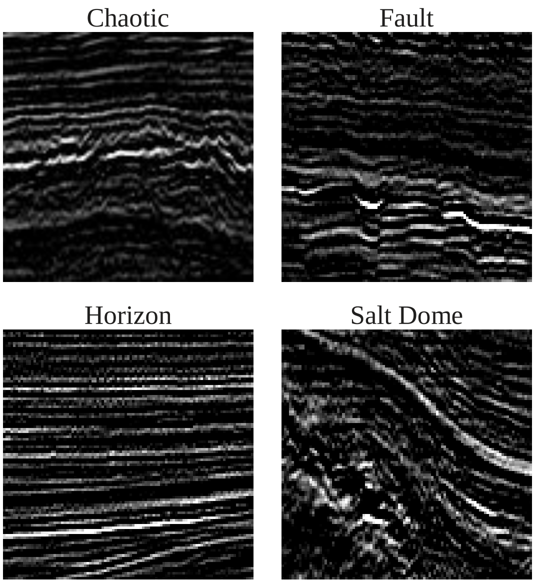

Fig. 1. Sample images of different Seismic Classes

B. Proposed Features

Feature extraction refers to the transformation of raw data into relevant numerical features that can be processed. These methods use specialized deep networks and algorithms to extract features from images and signals. They represent the desired sections of an image as a compact feature vector. Feature extraction involves the reduction of redundant data from a dataset that may lead to challenges while analyzing and examining images with many variables and cause the classification algorithm to overfit to training samples.

The proposed features includes Discrete Wavelet Transform (DWT), Discrete Cosine Transform (DCT), Discrete Fourier Transform (DFT) and Gabor Features.

a) Discrete Wavelet Transform Features: DWT is a linear transformation method in which the wavelet variance decomposes a signal or variance into several sets of wavelet basis functions [16-17]. Each component is a time series of coefficients specifying the time evolution of the signal in the corresponding sinusoid with a specific frequency band. The methodology extract features from time series to construct a classification model. The strength of each set in the decomposition indicates the extent of variability between adjacently located averages. The method relies on wavelet variance/wavelet spectrum to eventually produce an analysis of variance. It relies on the spectral analysis technique that uses the Fourier-based spectral density function.

The algorithm used to calculate the 1-D DWT functions on a vector x of length 2n, where n is an integer, produces a transformed vector w of equal length. Vector x is filtered with a discrete-time, low-pass filter of a specific length at intervals of two. These values are then stored in the first eight elements of w. Afterward, vector x is filtered with a discrete-time, high-pass filter of a given length at intervals of two. These values are again stored in the last eight elements of w. The image divides into 4 sub-bands, low-resolution image, horizontal, vertical, and diagonal. It can then be divided into several levels. The method has high computational efficiency due to its exceptional localization properties. Furthermore, the images it produces can be reconstructed and returned to their original state. It is also suitable for spectral analysis/time-frequency analysis and signal denoising. Some of its disadvantages include the lack of directional selectivity. It is also not suitable for pure stationary analysis and is computationally intensive for fine analysis. Additionally, the method has a less efficient discretization.

b) Discrete Cosine Transform Features: This method divides the image into spectral sub-bands of differing importance with respect to the image's visual quality [18]. It is a set of basic functions which takes an array of size of (n×n) as input to be calculated by the DCT formula and then stored. The 64 basic functions are shown by an image where the horizontal frequencies decrease from right to left, and vertical frequencies decrease from bottom to top. It produces an image after calculating the summation of sinusoids of differing magnitudes and frequencies. Furthermore, two approaches are used to process the DCT using Image Processing Toolbox software. The first approach is to use the FFT-based algorithm, 2-D discrete cosine transform, to process large inputs in a fast manner. The second approach is to use the DCT transform matrix as an output of the function. It is similar to the Discrete Fourier Transform as both transform a signal from the spatial domain to the frequency domain. The methodology has the ability to approximate lines with fewer coefficients. It is suitable for image compression applications because the visually important data of the image is concentrated in a few coefficients, and is also suitable for small square inputs, such as (8x8) or (16x16). Another advantage of the method is that the output for constant matrices is composed of a large number of zero values. However, it requires a quantization step to process the values in each DCT block, which produces an integer-valued output. Without the quantization step, the output values are real-valued. Equation 1 describes the 1D DCT [19].

$$F(u) = \left(\frac{2}{N}\right)^{\frac{1}{2}} \sum_{i=0}^{N-1} A(i).\cos[\frac{\pi.u}{2.N}(2i+1)]f(i) \qquad (1)$$

Because it is an invertible transform, the inverse is shown in Equation 2.

$$A(i) = \begin{cases} \frac{1}{\sqrt{2}} & for\ \varepsilon = 0 \\ 1 & otherwise \end{cases} \qquad (2)$$

c) Discrete Fourier Transform Features: DFT is used for digital signal processing. The method transforms an aperiodic signal from the time domain to the frequency domain. The major applications of DFT are calculating a signal's frequency spectrum and finding a system's frequency response from the system's impulse response. We can extract features either from the time or frequency domain. 90% of computations in CNNs are convolutions, and Fast Fourier Transform diminishes the intensity of such computations by replacing them with input and filter matrices, which are then converted into the frequency domain to perform multiplications. The output is then processed by inverse FFT for it to be transformed back into the time domain. The method reduces the complexity of computations from exponential to logarithmic. It is easy to program and implement in any coding language. One disadvantage includes having a frequency domain matrix that is more complex than its input. It also requires additional memory space, bandwidth, and long computational time. A digital computer cannot directly use this equation to analyze the signal. It must use samples of the output of the equation to compute an approximation of the spectrum of the input signal as shown in Equation 3 and its inverse in Equation 4.

$$\mathrm{A}(\mathrm{e}^{\mathrm{j}\,\omega}) = \sum_{\mathrm{i}=-\infty}^{\infty} \mathrm{a}\,(\,\mathrm{i})\mathrm{e}^{-\,\mathrm{j}\,\mathrm{i}\,\omega} \qquad (3)$$

$$\mathrm{A}(\mathrm{i}) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \mathrm{A}\,(\,\mathrm{e}^{\mathrm{j}\,\omega})\mathrm{e}^{\,\mathrm{j}\,\mathrm{i}\,\omega}\,\mathrm{d}\omega \qquad (4)$$

d) Gabor Features: The Gabor feature extraction method extracts the Gabor features of grey-scale images from different scales and orientations [20]. These features are extracted from gray-scale character images by Gabor filters. These filters are a sinusoidal signal with a given frequency and orientation and are constructed from statistical information of character structures. They are linear filters used for texture analysis to interpret frequency content in a specific region. The output of the Gabor filters is used to construct histogram features to better differentiate between the extracted features. To use this method, first, calculate Gabor features at a specific scale and orientation to obtain a set of filters. Then, each filter is convolved with the image to acquire 40 representations of each image. Each image offers a feature vector. After convolution, the Response Matrices are converted to feature vectors. Feature vectors are composed of mean amplitude, phase-amplitude, local energy, or orientation having max energy local. Eventually, the matrices are appended to create a feature matrix. For low-quality images, an adaptive sigmoid function is implemented on the output of Gabor filters to enhance the performance. The method is suitable for low-quality recognition and exhibits low invariant properties of the extracted features. However, it must be used several times from different angles. If the filter has a DC component, it will produce an unwanted effect of enhancing bright smooth regions. The filter's bandwidth must be narrow to constrict the DC component.

C. *Classification*

Image classification refers to the process of defining a set of target classes, which are objects to identify in images, and training a model to distinguish between them using labeled images. In other words, it is the process of analyzing an image and then categorizing the class or group the image falls under. Classification involves the comparison of the image's patterns to the target pattern. For the model to classify images, data is preprocessed, and the desired objects are detected and then labeled within the image set. The proposed model uses the following 4 classification algorithms: Support Vector Machine (SVM), Random Forest (RF), Fast Decision Trees (FDT), and Naïve Bayes (NB). These are briefly described next.

a) Support Vector Machine: SVM is a linear model for classification and regression problems. The algorithm simply creates a line or a hyperplane to divide the data into classes [21-22]. We specify the support vectors, which are the points closest to the line from both classes, and then calculate the margin, which is the distance between the line and the support vectors. The method works to raise the value of this margin to reach the optimal hyperplane. A hyperplane is a flat 1-dimensional subset of the space that divides the data into two individual sections in an n-dimensional Euclidean space. If the data is not linearly separable, an additional dimension must be added and then the decision boundary is set to its original dimensions using mathematical transformation. Afterward, place the points in one array, the classes they are a part of in another array, and train the SVM model the dataset using a linear kernel. Furthermore, SVM could be implemented using the scikit-learn library by importing it, creating an object, constructing the mode, and predicting results. The method is suitable for many practical problems with a clear margin of separation. It can solve both linear and non-linear problems and works best in high dimensional spaces, and in situations where the number of samples is less than the number of dimensions. It is memory efficient as it uses a subset of the support vectors in the decision function. However, it has a high training time when the dataset is large and is not suitable for a dataset with noise also, it does not offer probability estimates. Equation 5 show the decision function which based on the decision rule as shown in Equation 6 and when $-x$ as y, we get Equation 7.

$$C = \begin{cases} 1\ if\ \vec{A}.\vec{b} + y \geq 0 \\ -1\ if\ \vec{A}.\vec{b} + y < 0 \end{cases} \qquad (5)$$

$$\vec{A}.\vec{b} - x \geq 0 \qquad (6)$$

$$\vec{A}.\vec{b} + y \geq 0 \qquad (6)$$

b) Random Forest: The Random Forest model is based on decision trees that operate as an ensemble and are trained with the bagging method, a collection of learning models [23]. The method builds several decision trees and combines them to obtain a stable prediction. Firstly, it takes random records from a dataset, and individual decision trees are built for each sample. The feature space is restricted when constructing a tree to produce trees that could be distinguished. Each decision tree produces an output based on Majority Voting or Averaging for classification. The method can easily size the significance of each input feature on the prediction. It has high versatility and can produce accurate prediction results due to the default hyperparameters used. It is fast to train and has the ability to handle datasets with continuous variables in regression, and categorical variables in classification. It exhibits high stability as results are based on Majority Voting or Averaging. Furthermore, it makes full use of the CPU to construct random forests as each tree is shaped based on distinct data and attributes. It does not require the separation of data for training and testing as 30% of the dataset will not be used by the decision tree.

However, overfitting could occur if there aren't enough trees in the forest and the method could be inefficient when a huge number of trees are inputted. Additionally, higher accuracy needs more trees, leading to a slower operating model.

c) Fast Decision Trees: It is based on a conditional independence assumption. The algorithm has a pseudo-polynomial time complexity, which depends on the size of the training data and the number of attributes [24]. The decision tree is obtained from a set of labeled training instances represented by a tuple of attribute values and a class label. The tree is incrementally built using binary recursive partitioning, which separates training data into sections on the feature that results in the largest information gain (IG) and creates a hierarchical structure. This step reduces the uncertainty towards the end outcome. It clusters data into dense and empty regions and classifies them based on their classes. The algorithm first specifies a test for the root node and creates a branch for all possible outcomes. It then splits each branch recursively into subsets and terminates the recursion if all its instances have the same class. Setting a limitation on the depth of the tree avoids overfitting. The method has the same speed as naïve Bayes but is more accurate. It requires no parameters, excludes insignificant features, and has the ability to manage mixed-type data. It is easy to interpret for small-sized trees and is speedy for unknown records classification. However, it does not properly handle missing values and it could be biased toward branches with a greater number of levels. Furthermore, interpreting big sized can be challenging and maybe counter intuitive.

d) Naïve Bayes: The method is based on the Bayes theorem [25]. For each class variable, the methodology assumes that the value of a certain characteristic is independent of the value of any other characteristic. It uses maximum likelihood for parameter estimation and therefore, it is unnecessary to use Bayesian methods or Bayesian probability. The method organizes the dataset into a frequency table. Second, it creates a likelihood table using probabilities like the probability of playing and Overcast probability. Next, the posterior probability of each class is computed using the Naïve Bayesian equation. The highest posterior probability class is the result of the prediction. The method is suitable for recommendation systems, document classification, sentiment analysis, and spam filtering. It is better suited for categorical input variables than numerical variables. It also requires a minimum amount of training data to predict the necessary parameters. It is faster and easier to implement than many methods and is useful for large-sized data sets and multi-class prediction. However, predictors and features are independent and if the categorical variable has a category that has not been examined in the training set, a Zero Frequency state occurs, which means the classifier will not complete the classification and the probability will be 0. It is also a bad estimator. The Bayes theorem is based on the Equation 8 and the classification rule for Naïve Bayes is based on the Equation 9 and Equation 10.

$$P(X \mid Y) = \frac{P(Y \mid X)\, P(X)}{P(B)} \tag{8}$$

$$P(n \mid m_1, m_2, \ldots\ldots, m_s) \propto P(\,n) \prod_{i=1}^{s} P(\,m_i \mid n) \tag{9}$$

$$\hat{n} = \arg\lim_{n} P(\,n) \prod_{i=1}^{s} P(\,m_i \mid n) \tag{10}$$

IV. Results and Discussions

The proposed methodology was tested on the LANDMASS dataset previously described. The dataset acted as an input to extract the features present in the seismic images using four different extraction methods, Discrete Wavelet Transform for Features, Discrete Cosine Transform Features, Discrete Fourier Transform Features, and Gabor Features. Afterward, each feature extraction method used a classifier to classify the obtained features based on the four different seismic classes, Chaotic, Fault, Horizon, and Salt Dome. The accuracy, precision, recall, and F1 measure results of the classification methods, Support Vector Machine, Random Forest, Fast Decision Trees, and Naïve Bayes. As shown in Table 1, Table 2, and Table 4, the RF method yielded the highest accuracy, recall, and F1 measure among all different classification techniques used in the study. The highest precision was attained using the SVM method. However, RF managed to obtain high values in this success-measuring parameter for DWT features, DCT features, and Gabor Features. Therefore, the highest achieving classifier was the Random Forest method to classify the extraction methods previously stated.

TABLE I. Comparison of different classifier results based on the DWT Features

	Accuracy	Precision	Recall	F1 Measure
RF	87.46	0.86	0.81	0.83
FDT	78.78	0.74	0.72	0.73
Naïve Byes	53.82	0.52	0.56	0.54
SVM	65.76	1.00	0.22	0.36

TABLE II. Comparison of different classifier results based on the DCT Features

	Accuracy	Precision	Recall	F1 Measure
RF	94.17	0.92	0.92	0.92
FDT	91.57	0.89	0.89	0.89
Naïve Byes	72.13	0.68	0.77	0.73
SVM	65.76	1.00	0.22	0.36

In Table 3, the highest accuracy was achieved by RF with a value of 86.24% with precision 0.79, recall 0.85. The highest recall was realized by Naïve Byes with a value of 0.96, and the highest accuracy and F1 measure were achieved by RF with a value of 86.24%, and 0.82, respectively. The highest performing classifier based on the DFT feature extracting method is RF, which received high scores in all success-measuring parameters. In Table 4, the highest accuracy was achieved by RF with a value of 76.91% with precision 0.73, recall 0.57.

TABLE III. COMPARISON OF DIFFERENT CLASSIFIER RESULTS BASED ON THE DFT FEATURES

	Accuracy	Precision	Recall	F1 Measure
RF	86.24	0.79	0.85	0.82
FDT	78.62	0.74	0.77	0.75
Naïve Byes	30.33	0.29	0.96	0.45
SVM	69.44	0.94	0.32	0.48

TABLE IV. COMPARISON OF DIFFERENT CLASSIFIER RESULTS BASED ON THE GABOR FEATURES

	Accuracy	Precision	Recall	F1 Measure
RF	76.91	0.73	0.57	0.64
FDT	59.45	0.45	0.38	0.41
Naïve Byes	57.81	0.47	0.52	0.49
SVM	65.76	1.00	0.22	0.36

The highest achieved accuracy was obtained by the Random Forest classification method based on DCT features with a value of 94.17% as shown in Table 2. The highest obtained precision was again achieved by DCT features after using the RF method with a value of 0.92. DCT successfully attained the highest F1 measure value of 0.92 and has a high 0.92 precision score. Therefore, the combination of Random Forest classification method and the Discrete Cosine Transform Features extraction method leads to the most desirable results to classify seismic structures in seismic images.

V. CONCLUSION

Novel features in seismic images are used to categorize different classes of seismic structures. The classification of seismic facies supports in the approximation of grain size, mineralogy, the permeability of deposition units, and in the identification of areas of interest. The proposed methodology uses the LANDMASS database, composed of two datasets, LANDMASS-1, with 17667 images, and LANDMASS-2, with 4000 images. The datasets contain seismic images of four different classes of seismic structures, Chaotic, Fault, Horizon, and Salt Dome. The feature extraction methods used were Discrete Wavelet Transform for Features, Discrete Cosine Transform Features, Discrete Fourier Transform Features, and Gabor Features. Next, classification methods were implemented to classify the novel features. These classification methods are Support Vector Machine, Random Forest, Fast Decision Trees, and Naïve Bayes. The results of this paper suggest that the combination of Random Forest classification method and the Discrete Cosine Transform Features extraction method acquires the highest achieving results to classify novel features of seismic structures in seismic images. To interpret the results and measure the success of feature-extracting and classification methods, accuracy, precision, recall, and F1 measure were the success-measuring parameters. The achieved accuracy of this combination is 94.17% with precision, recall, and F1 measure of 0.92.

REFERENCES

[1] De Santis, V., Caldara, M. and Pennetta, L., 2020. "Continuous" Backstepping of Holocene Coastal Barrier Systems into Incised Valleys: Insights from the Ofanto and Carapelle-Cervaro Valleys. *Water*, *12*(6), p.1799.

[2] Cox, D.R., Newton, A.M. and Huuse, M., 2020. An introduction to seismic reflection data: Acquisition, processing and interpretation. In *Regional Geology and Tectonics* (pp. 571-603). Elsevier.

[3] Crutchley, G.J. and Kopp, H., 2018. Reflection and refraction seismic methods. In *Submarine geomorphology* (pp. 43-62). Springer, Cham.

[4] Amin, A., Deriche, M., Qureshi, M.A. and Memon, K.H., 2019. New attributes for salt dome detection in 3D seismic data using higher order SVD. *Journal of Applied Geophysics*, *163*, pp.108-116.

[5] "LANDMASS | Datasets, Simulators & Codes | Center for Energy & Geo Processing – CeGP," *Gatech.edu*, 2022. https://cegp.ece.gatech.edu/codedata/landmass/ (accessed July. 23, 2022).

[6] Milosavljević, A., 2020. Identification of salt deposits on seismic images using deep learning method for semantic segmentation. *ISPRS International Journal of Geo-Information*, *9*(1), p.24.

[7] Hosseini-Fard, E., Roshandel-Kahoo, A., Soleimani-Monfared, M., Khayer, K. and Ahmadi-Fard, A.R., 2022. Automatic seismic image segmentation by introducing a novel strategy in histogram of oriented gradients. *Journal of Petroleum Science and Engineering*, *209*, p.109971.

[8] Wang, Z., Di, H., Shafiq, M.A., Alaudah, Y. and AlRegib, G., 2018. Successful leveraging of image processing and machine learning in seismic structural interpretation: A review. *The Leading Edge*, *37*(6), pp.451-461.

[9] Gao, K., Huang, L., Zheng, Y., Lin, R., Hu, H. and Cladohous, T., 2022. Automatic fault detection on seismic images using a multiscale attention convolutional neural network. *Geophysics*, *87*(1), pp.N13-N29.

[10] Xiong, W., Ji, X., Ma, Y., Wang, Y., AlBinHassan, N.M., Ali, M.N. and Luo, Y., 2018. Seismic fault detection with convolutional neural network. *Geophysics*, *83*(5), pp.O97-O103.

[11] ul Islam, M.S., 2020. Using deep learning based methods to classify salt bodies in seismic images. *Journal of Applied Geophysics*, *178*, p.104054.

[12] Zhang, Y., Lin, H., Li, Y. and Ma, H., 2019. A patch based denoising method using deep convolutional neural network for seismic image. *IEEE Access*, *7*, pp.156883-156894.

[13] Li, S., Liu, B., Ren, Y., Chen, Y., Yang, S., Wang, Y. and Jiang, P., 2019. Deep-learning inversion of seismic data. *arXiv preprint arXiv:1901.07733*.

[14] Wu, X., Yan, S., Qi, J. and Zeng, H., 2020. Deep learning for characterizing paleokarst collapse features in 3-D seismic images. *Journal of Geophysical Research: Solid Earth*, *125*(9), p.e2020JB019685.

[15] An, Y., Guo, J., Ye, Q., Childs, C., Walsh, J. and Dong, R., 2021. Deep convolutional neural network for automatic fault recognition from 3D seismic datasets. *Computers & Geosciences*, *153*, p.104776.

[16] Latif, G., Butt, M.M., Khan, A.H., Butt, O. and Iskandar, D.A., 2017, April. Multiclass brain Glioma tumor classification using block-based 3D Wavelet features of MR images. In *2017 4th International Conference on Electrical and Electronic Engineering (ICEEE)* (pp. 333-337). IEEE.

[17] Latif, G., Iskandar, D.A. and Alghazo, J., 2018, September. Multiclass brain tumor classification using region growing based tumor segmentation and ensemble wavelet features. In *Proceedings of the 2018 International Conference on Computing and Big Data* (pp. 67-72).

[18] Abbas, H. and George, L.E., 2020, November. Image Classification Schemes Based on Statistical Moments of Wavelet and Sliced Radial

Energy Distribution of DCT. In *IOP Conference Series: Materials Science and Engineering* (Vol. 928, No. 7, p. 072116). IOP Publishing.

[19] Latif, G., Iskandar, D.A., Jaffar, A. and Butt, M.M., 2017. Multimodal brain tumor segmentation using neighboring image features. *Journal of Telecommunication, Electronic and Computer Engineering (JTEC)*, *9*(2-9), pp.37-42.

[20] Cruz-Ramos, C., Garcia-Salgado, B.P., Reyes-Reyes, R., Ponomaryov, V. and Sadovnychiy, S., 2021. Gabor features extraction and land-cover classification of urban hyperspectral images for remote sensing applications. *Remote Sensing*, *13*(15), p.2914.

[21] Haque, M.S.M., Latif, G., Hasan, M.R., Arifuzzaman, M., Shafin, S.S. and Rahman, Q.A., 2019. Scalable parallel SVM on cloud clusters for large datasets classification.

[22] Latif, G., Ben Brahim, G., Iskandar, D.N.F., Bashar, A. and Alghazo, J., 2022. Glioma Tumors' Classification Using Deep-Neural-Network-Based Features with SVM Classifier. *Diagnostics*, *12*(4), p.1018.

[23] Latif, G., Al Anezi, F.Y., Zikria, M. and Alghazo, J., 2020, January. EEG-ECG Signals Classification for Arrhythmia Detection using Decision Trees. In *2020 Fourth International Conference on Inventive Systems and Control (ICISC)* (pp. 192-196). IEEE.

[24] Liu, P., 2021. Extracting Rules from Trained Machine Learning Models with Applications in Bioinformatics.

[25] Kaur, G. and Oberoi, A., 2020. Novel approach for brain tumor detection based on Naïve Bayes classification. In *Data management, analytics and innovation* (pp. 451-462). Springer, Singapore.