National College of Ireland

Deep Neural Network for Seismic Image Segmentation and Detection of Salt Domes

Msc Research Project
Data Analytics

Nikhil Chambial
Student ID: x20232713

School of Computing
National College of Ireland

Supervisor: Mr. Taimur Hafeez

National College of Ireland
Project Submission Sheet
School of Computing

Student Name:	Nikhil Chambial
Student ID:	x20232713
Programme:	Data Analytics
Year:	2022
Module:	Msc Research Project
Supervisor:	Mr. Taimur Hafeez
Submission Due Date:	19/09/2022
Project Title:	**Deep Neural Network for Seismic Image Segmentation and Detection of Salt Domes**
Word Count:	7476
Page Count:	21

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.

ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature:	
Date:	18th September 2022

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST:

Attach a completed copy of this sheet to each project (including multiple copies).	☐
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	☐
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	☐

Assignments that are submitted to the Programme Coordinator office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Deep Neural Network for Seismic Image Segmentation and Detection of Salt Domes

Nikhil Chambial
x20232713

Abstract

Historically, seismic images have been essential in understanding sub-surface structures. Salt Domes is a sub-surface structure that traps hydrocarbons. Domain experts such as geologists use these seismic images to detect salt domes and explore oil and natural gas reservoirs. However, domain experts require substantial computational power to analyze these seismic images, and the classification is also error-prone. Moreover, the precise position of the salt deposit is critical since drilling in the wrong spot would result in significant expenses for corporations as well as environmental harm. As a result, computational systems can stimulate experts in categorizing subsurface structures such as salt domes to fasten the analysis process, which is critical to the industry's growth. Deep learning's growing popularity prompted academics to apply such technologies to seismic data. While these methods have produced promising results, the difficulty in determining an appropriate initialization for adjusting the model's parameters is a common issue in deep machine learning systems. Often the random initialization results in longer training sessions, failure to locate the solution, and disappearing gradients till the first layers owing to back-propagation. To solve this problem, modern deep neural network models with transfer learning are used to provide an initialization point for the model's parameters and improve the performance. Most of the deep learning models in the past showed sub-par performance due to the scarcity of data that has been resolved in this study by using data augmentation techniques such as rotation and flipping. This study has implemented a state-of-the-art encoder-decoder model called Unet to detect salt domes in seismic images. The model has been trained and tested on the publicly available seismic image dataset provided by TGS. The Unet model is evaluated and optimized using various evaluation metrics such as Intersection over Union(IoU) and Accuracy, with values of 0.80 and 97%, respectively.

Keywords- Salt Domes, Seismic Images, UNet, IoU(Intersection over Union), Data Augmentation, Convolution Neural Network

1 Introduction

1.1 Background and Motivation

Huge salt deposits are present in many places on Earth, and some locations also have significant oil and gas resources. Oil and gas companies may extract oil and gas more effectively and in less harmful ways using seismic imaging of salt bodies. Seismic pictures

aid in the visualization of the subsurface structures needed to find hydrocarbon fuel. Sound waves are emitted by receivers called geophones and are reflected by the subsurface structures, which create seismic images. As seen in Figure 1 [1], a three-dimensional image of the subterranean rocks is made using the recorded reflected sound signal. The captured pictures are known as seismic images, and they depict the borders of various rock types.

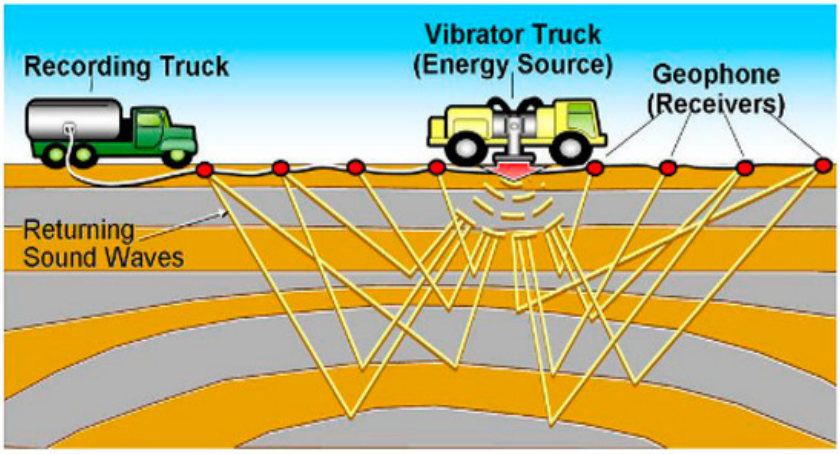

Figure 1: Seismic Image Collection

Even though it is highly challenging to identify the rock types from these seismic pictures, the reflected signal can reveal the boundaries between the rocks at the site of contact. This seismic picture interpretation is primarily used to locate salt domes that store hydrocarbon hydrocarbons like oil and natural gas, as seen in Figure 2 [2]. The properties of salt domes are numerous, complex, and only interpretable by specialists like geologists. Since salt domes have a lower density than the rocks surrounding and a higher seismic velocity than the rocks nearby, the seismic pictures of the salt dome's boundary exhibit a sharp reflection, making it easier for the experts to locate them.

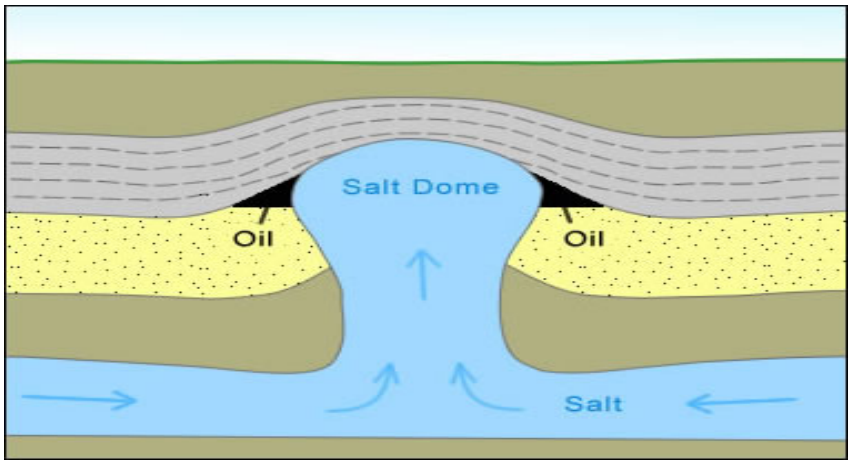

Figure 2: Oil Reservoir trapped by Salt Dome

[1] https://www.petroman.com
[2] https://www.geology.com

It takes a lot of processing power and domain knowledge to analyze seismic pictures to comprehend the subsurface structures. Seismic pictures are stored, rendered, and interpreted using high-end computer hardware and software. Experts interpret these seismic images to detect numerous structures such as salt domes, faults, and channels. Using powerful computers to render seismic images is better than manual interpretation but still requires much time to show conclusive results. The manual correction and modification of human interpretations introduce bias and forecasting inaccuracies. Much work has gone into automating, deciphering seismic pictures, and producing accurate forecasts. These challenges motivate this research study to create an automated deep learning framework that can accurately detect salt domes in the seismic images while improving the classification accuracy and reducing the time taken to interpret the seismic images.

1.2 Research Question

"To what extent can deep machine learning techniques in combination with data augmentation techniques help to detect salt domes under earth's surface by segmenting the seismic images?"

This research aims to address the above research question using Unet architecture (Ronneberger et al.; 2015) to create a deep learning model to detect salt domes in seismic images accurately. This research aims to develop an automated deep learning model that can categorize each seismic picture pixel as either salt or non-salt. This research has used TGS's seismic image dataset, which includes 22000 images with corresponding salt masks. The dataset also includes depth data in addition to the salt masks. As this is an image segmentation model, a deep learning approach would be best suited to detect salt domes. Although CNN algorithms have been used since the 1990s, deep CNN's development has completely changed how this field is approached since a Deep CNN network called Alexnet ultimately defeated the runners-up by a wide margin in the Imagenet Large Scale Visual Recognition Challenge (Guo et al.; 2020). On the TGS seismic image dataset, many deep learning models, including VGGNet(Alfarhan, Deriche and Maalej; 2020), Segnet (Badrinarayanan et al.; 2017), Resnet50, ResNext, and DenseNet, have been used in the past. However, these models achieved high accuracy but faced the issue of overfitting due to scarcity of data. This research aims to overcome this issue using data augmentation techniques to increase the dataset volume.

1.3 Research Objectives

This research paper pursues various techniques to fulfill the objective of research question and is organized into different sections which are elaborated below:

Section 2: Discussion of traditional and modern research in salt dome detection in seismic images.
Section 3: Discusses the adopted methodology in this research.
Section 4: Demonstrates the Design Specification .
Section 5: Presents the implementation of the Unet architecture.
Section 6: Discusses the model training parameters.

Section 7: Evaluation of results of the implemented model.
Section 8: Conclusion of the research.

2 Related Work

The research on detecting salt domes in seismic images can broadly be classified into the traditional Feature Extraction Approach and Modern Deep Learning Methods. Most methods for locating and understanding salt domes were developed by analyzing the characteristics of two-dimensional seismic images. It is a segmentation problem as the features are taken from the seismic images and used to find the salt borders.

2.1 Traditional Methods

To identify salt dome borders in seismic pictures, (Lomask et al.; 2006) applied the Normalized Cuts and Image Segmentation technique that (Shi and Malik; 2000) had previously developed. Due to the potential size of the weight matrix and the associated high processing demands, the computations were divided across a parallel network, which turned out to be extremely precise and efficient. NCIS handled the problem of local discontinuities by ignoring the local data and extracting the global characteristics from the images and proving to be an efficient approach to salt dome detection. As opposed to the reflectivity idea employed in NCIS algorithms, textural alterations are typically a superior method of classifying salt borders in seismic images, as just using the variations in amplitude can lead to incorrect detection. (Hegazy* and AlRegib; 2014) suggested that in order to obtain the smooth delineated region, the salt borders may be detected utilizing three different seismic image texture features, which are smoothness, edge content, and direction. The authors employed the eigenvalues of the moment of inertia of any pixel (x, y) to extract the direction attribute of any specific pixel since the salt content area in the seismic pictures lacked direction. In other instances, the directionality characteristic could not identify salt locations since several seismic images contained non-salt areas with no direction that were consequently labeled as salt regions. Since salt dome borders have a high intensity reflected in the pixels, edge detection is instrumental in identifying salt domes in seismic pictures. Sobel Edge Detector (Woods; 1992) used the derivative approach to detect the salt dome boundary. Sobel Edge Detector applies two masks on image slices where the first mask detects the boundaries of the salt domes. In contrast, the second mask is responsible for detecting if there are any discontinuities in any direction. For example, (Jing et al.; 2007) took inspiration from Sobel Edge Detector and used weighted masks to detect salt boundaries. This approach showed an accuracy of 81.1% in detecting salt Domes. The method produced decent results but was only accurate when the seismic pictures' amplitude fluctuations in the x and y directions were of the same order. Another variation of Sobel Edge Detector was adopted by (Aqrawi et al.; 2011), who overcame the amplitude variation issue by using 3D Sobel Edge Detector to detect salt domes in seismic images. The amplitude fluctuations in the 2D model were addressed by this method, but the abnormalities in identifying borders on the diagonals of seismic pictures persisted. This limitation was overcome by (Asjad and Mohamed; 2015) who used a 3D edge detector that combined the gradient maps of three directions of the salt dome edges. This approach showed better results than the Sobel Detector, as it could correctly detect salt domes in seismic images with high amplitude variations and low intensity, giving an accuracy of 89%.

2.2 Deep Learning Seismic Image Analysis

Deep learning algorithms have been used to solve the problems posed by traditional methods by adequately identifying the salt borders by readily extracting any number of variables from the seismic images. Deep learning technology has produced significantly superior results compared to the traditional techniques of salt identification based on edges and textures.

2.2.1 Standard Deep Learning Models

(Di et al.; 2018) employed a multi-layer perceptron and CNN combination to locate salt bodies in seismic pictures and assess how well it performed compared to other approaches. The writers looked at the key elements that aided in comprehending and recognizing the various characteristics of seismic pictures. It was discovered that CNN could effectively produce the features in the seismic images. Previously, interpretations of seismic images required experienced interpreters, and the characteristics needed to be fine-tuned to acquire good performance. With this model, the random noise in seismic pictures could be effectively removed, and the salt bodies and distinctive patterns could be recognized accurately. A four-step process was suggested by (Pochet et al.; 2018) to locate faults in the seismic pictures. Due to the lack of high-quality photos, the initial step was to build synthetic seismic images made up of faults. Using the Ricker wavelet on reconstructed photos and adding random noise to the images, a collection of around 500 images, each measuring 572 by 572 pixels, was produced. If a pixel in the corresponding binary picture was covered by a white pixel, it was regarded as a fault, and vice versa. Finally, the architecture was developed, incorporating the Rectified Linear Unit activation function for quicker training than other activation functions and max-pooling for nonlinear downsampling. On the artificial image dataset, this model displayed accuracy of 0.98. (Waldeland et al.; 2018) conducted the study on seismic imagery for salt body detection using CNN. Five convolutional layers and one linked layer make up the CNN architecture suggested by the authors. Finally, the authors classified pixels as salt or no salt pixels using a two-node SoftMax loss function. The suggested model was trained using an inline slice that had been manually labeled and several little cubes placed around the relevant pixels. The study was conducted using seismic data from the Netherlands F3 block. (Pochet et al.; 2018) suggested a unique approach to locate salt bodies using electromagnetic algorithms using data from towed streamers. A batch normalization layer, two max-pooling layers, a SoftMax loss function, and two fully connected layers were all present in each of the four convolutional layers of the CNN model. Rectified Linear Unit and Dropout Layers were also employed in this model to facilitate quicker training and prevent overfitting. The dropout rate for Convolutional layers was adjusted to 0.1 and set as 0.2 for fully connected layers. The accuracy for the noise-free data was 92.62 percent, indicating that the model performed well. The model provided an accuracy of over 80% even when there was data noise. (Xiong et al.; 2018) constructed synthetic seismic images by making a few alterations such as inserting folds, shears, faults, and noise. The output of the CNN model was a class label for the patch's center point. Batch normalization, four convolutional layers, four maximum pooling layers, a nonlinear RELU function, and an activation function made up the model's architecture. The outputs were finally subjected to the SoftMax function before being converted to probabilities. The authors observed the model accuracy using several loss functions, including CE Loss, BCE Loss, and Focal Loss. BCE loss function fared better than CE and Focal loss, obtaining 91.21% accur-

acy.(El Zini et al.; 2019) used transfer learning to adopt a unique exploiting bright spot detection method. SeisNet architecture, a deep convolutional neural network, performed better than previous approaches, with an accuracy of 95.6%. Noise addition, rotation, and cropping data augmentation techniques were utilized to produce additional synthetic data, which improved the F1 score by up to 50%. The transfer learning approach was used to increase accuracy by 10% by transferring information obtained from vast amounts of unlabeled data to the classification stage.

2.2.2 UNET Deep Learning Models

(Shi et al.; 2018) detected the salt domes in seismic images by creating image segmentation architecture. Drawing inspiration from previous works such as Segnet Architecture by (Badrinarayanan et al.; 2017) and UNet Architecture by (Ronneberger et al.; 2015) for automatic detection of salt bodies in the seismic pictures, they developed an Unet encoder and decoder architecture. In order to correctly recognize salt bodies, (Zeng et al.; 2019) suggested the Resnet framework in conjunction with U-Net architecture for residual learning. They used the activation function of the exponential linear unit and the Lovasz-Softmax loss function in the model. Additionally, K-fold cross-validation was carried out to increase accuracy. Compared to the manual interpretation of seismic images, the model produced good results when deployed on the SEG-SEAM dataset. (Liu et al.; 2019) developed a unique Se-FPN architecture to identify salt bodies in seismic imagery. The encoder portion of the model used SENet-154 and has five convolutional layers for downsampling. The authors employed batch normalization and ReLU activation function to combat overfitting, taking inspiration from earlier publications. Using stochastic gradient descent, the authors trained the model with a parameter size of 24, momentum of 0.5, and a learning rate of 0.0001. In the first five epochs, Binary Cross Entropy loss was applied, and for the remaining epochs, Lovasz Softmax was applied. The model's IoU score of 87.48 percent demonstrated remarkable performance compared to comparable models, including Deeplab and FCN.

The U-Net backbone was used for most of the prior research; however, (Ma and Li; 2021) developed a Rotated U-Net model for fault identification in the seismic images. When features were extracted from pictures using up-sampling, the researchers noticed an information loss at the U-Net model's foundation. They were inspired by Mobilenetv2, where the residuals extracted the picture characteristics by ascending and descending. Then an inverse sampling encoder-decoder block was put in front of the U-Net model. To combat the overfitting issue, the input was passed through a three-by-three convolutional layer, a Dropout Layer, and finally, a ReLU activation layer. The input image of dimension 128*128*128 was used along with the parameters- learning rate as 0.0001, epochs as 100, and batch size as 1. This method demonstrated greater than 95% accuracy in accurately locating faults in seismic pictures. (Zhang et al.; 2019) suggested an ensemble-based approach for identifying seismic features. In order to build an ensemble technique, the authors pooled the output from the two models they had created using Segnet architecture and U-Net architecture backbones. Compared to the Segnet architecture model's accuracy of 92.40%, the U-Net architecture model demonstrated a 91.89% accuracy. However, the ensemble model developed employing both of these models had an enhanced accuracy of 96.57%.

2.2.3 UNET with Transfer Learning Models

(Alfarhan, Maalej and Deriche; 2020) suggested combining the U-Net and ResNet to distinguish salt bodies in seismic pictures. The suggested model demonstrated high accuracy, but no tangible outcomes could be demonstrated due to the lack of a sizable dataset of annotated seismic pictures. In (Alfarhan, Deriche and Maalej; 2020), the authors of this study once more suggested a brand-new, enhanced architecture. They used the labeled data from the F3 block for their study. The authors offered two distinct models. In the first model, the pooling and convolutional layers of the VGG19 architecture served as the encoder, and ResNet34, a network noted for its ability to handle the vanishing gradient issues, served as the encoder in the second network. Each convolutional layer in both models was followed by a reLU activation layer and a batch normalization layer. The U-Net encoder was swapped out for ResNet34 and VGG19 to generate the two variations of Unet architecture. Two techniques—random weight initialization and pre-trained weights—were used to train both systems. The photos were initially reduced to a size of 128 by 128, then normalized. With a learning rate of 0.0001 and a fading factor of 0.5, the four models were generated using these pictures over 100 epochs. Pre-trained networks outperformed networks trained using random initialization. The U-Net was used as the reference model while comparing the effectiveness of the four networks. The pre-trained models appeared to demonstrate significant advancements in defect identification. None of the four models outperformed the baseline U-Net model, which displayed an IoU metric value of 0.9778, even though all four models significantly increased defect detection. (Guo et al.; 2020) also suggested the U-Net design, which included edge prediction, an encoder, and a decoder to enhance the performance of salt dome identification. To link feature maps and the appropriate masks, they employed jump connections. At each phase, the encoder component took care of downsampling the seismic images. ResNet-34 served as the foundation of the whole model architecture. The decoder component handled the incremental upsampling and delivered the final results of the same size as the original seismic image. Due to a lack of data, the authors employed pre-trained ImageNet weights for the encoder to enhance the outcomes. The model's accuracy was increased using various parameters, including scSE, which increased the accuracy to around 85%, and deep supervision, which increased the accuracy further to around 87%. After testing the edge branch with various weights, the optimum weight of 0.1 increased accuracy by the maximum. (Alfarhan, Deriche, Maalej, AlRegib and Al-Marzouqi; 2020) employed a similar strategy, employing an Unet encoder-decoder architecture in which high-level features were transferred from the encoder part of the architecture to the decoder part for the extraction of low-level information via skip-level links. A sigmoid activation function and three convolutional layers made up the model's final section, which displayed each output as a probability. According to this study, the pre-trained ResNet model had a more excellent accuracy score of 98% compared to ResNet, which was created from scratch. A different approach of using a Mask Region-based Convolutional Neural Network was proposed by (Arsha and Thulasidharan; 2021) for delineating salt bodies in the seismic images. The segmentation model Mask R-CNN was used to classify each pixel of the seismic image as either a salt-containing or non-salt-containing pixel. The Mask R-CNN model used CNN backbone for feature extraction from seismic images. The researchers proposed to use RESNET101 as the backbone of the model. The performance of the Mask R-CNN was much better than the CNN model as it showed an accuracy of 0.93 compared to 0.58 of the CNN model. Another research was conducted by (ul Islam;

2020), who tackled this issue as a segmentation challenge by categorizing each pixel as a salt-containing pixel or a non-salt-containing pixel. These pixels were then used to form a mask for each seismic picture, where a white pixel denoted the presence of salt, whereas the black pixel denoted its absence. Before feeding the neural network, the seismic pictures were initially downsized to 101 by 101 by 1. The pictures were standardized within 0-1 after array conversion. The network included Resnet blocks in conjunction with the U-Net base design. With the increasing depth of the encoder in the network, the picture size reduced, and with the decrease in the depth of the decoder in the network, the picture size increased. With a modest adjustment to enhance the dependency of channels, the residual blocks were used to solve the overfitting problem. Additionally, outcomes were 25% better because of the introduction of Imagenet weights.

Research Paper	Model Name	Year	Dataset	Acc/IoU
Arsha and Thulasidharan (2021)	Mask R-CNN	2021	TGS Salt	0.93/NA
Alfarhan, Maalej and Deriche (2020)	Unet-Resnet	2020	Netherland F3	0.99/NA
Liu et al. (2019)	Se-FPN	2019	TGS Salt	NA/0.87
ul Islam (2020)	U-Net with Se-ResNet	2020	TGS Salt	NA/0.86
Zhang et al. (2019)	Ensemble Learning	2020	Netherlands F3	0.96/NA
El Zini et al. (2019)	SeisNet	2019	TGS Salt	NA/0.95
Ma and Li (2021)	Mobilenetv2	2021	TGS Salt	NA/0.96
Tschannen et al. (2020)	Ensemble Learning	2021	Netherlands F3	0.96/NA

Table 1: Summary of related work

3 Methodology

This research study's technique is similar to the Cross-Industry Process for Data-Mining (CRISP-DM). The seismic salt dome identification technique is depicted systematically in the Figure 3, which includes the phases of business understanding, data understanding, data preparation, modeling, evaluation, and results. The model output will effectively identify salt domes in seismic images.

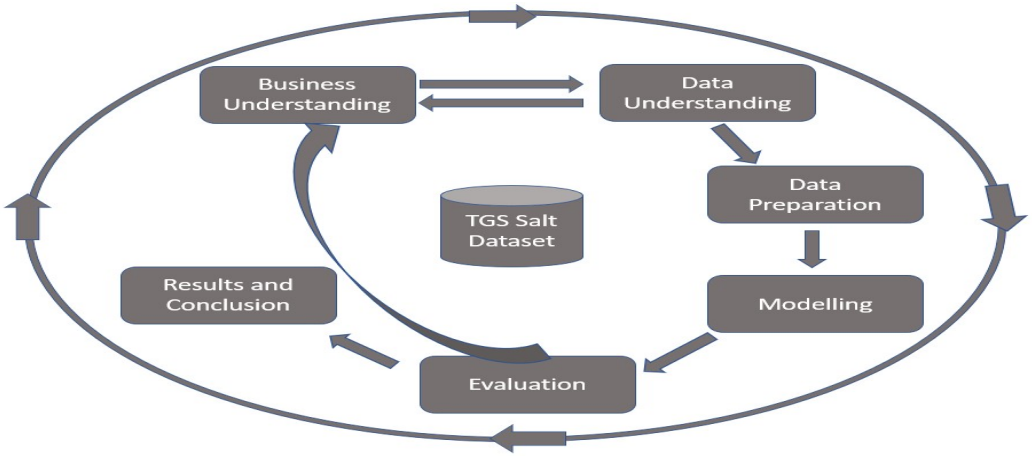

Figure 3: Salt Dome Detection Design Methodology

3.1 Business Understanding

The oil and gas industries are constantly dealing with industry-specific issues, such as a lack of coverage in intricate processes, performance improvement problems, equipment issues, and environmental regulations. Machine learning can help to transform massive datasets and efficiently make decisions to explore oil and gas in a particular area Alfarhan, Maalej and Deriche (2020). This would further help the oil companies to reduce costs related to hiring experts and expensive exploration equipment and would also pose less threat to the environment. In addition, the companies are looking for a cheaper alternative to expensive sensors and handling the vast amount of exploration data. Manually identifying salt domes is a laborious process that is expensive due to the need for qualified interpreters. The time commitment and error-proneness of conventional procedures for identifying salt domes are its significant drawbacks. Furthermore, these techniques call for specialists to extract precise seismic characteristics, which they cannot achieve in noisy seismic pictures. Deep learning systems have outperformed more conventional approaches when identifying salt domes in seismic picturesul Islam (2020). Adopting deep learning algorithms to automate salt dome recognition would thus be highly advantageous to the oil and gas sector.

3.2 Data Understanding

This dataset [3] used in this research is made accessible to the public on the Kaggle platform by the Geophysical Company TGS. The data includes seismic pictures taken at numerous arbitrary subsurface sites. The seismic images in the dataset were obtained using the reflection seismology method, which creates seismic images utilizing reflected light from rocky layers under the earth's surface. Each pixel in the seismic images may be classified as either salt or not; these are all 101 by 101 pixels in size. A CSV file with each seismic image's depth is also available. The training dataset also contains a pixel-by-pixel classification of the seismic image, also known as a seismic image mask. There are 4000 seismic images and 4000 mask files in the training dataset. The testing dataset also consists of 12000 seismic test photos that will be used to assess the model's accuracy and other evaluation measures. However, in the absence of masks for the seismic images in the test dataset, we cannot use these images to test the model's accuracy. Figure 4 shows the seismic image and the corresponding salt mask.

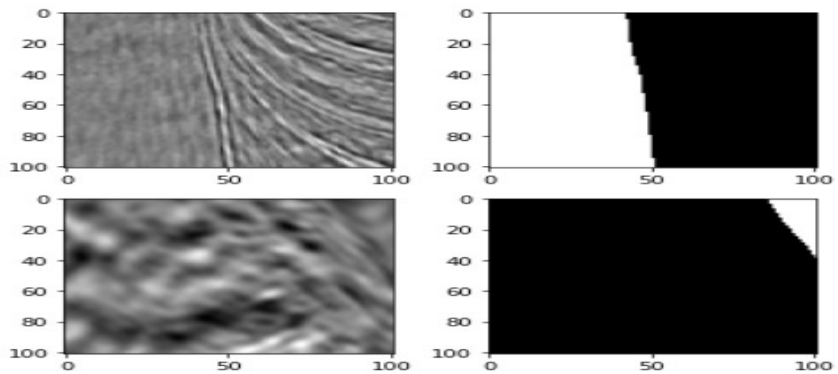

Figure 4: Seismic Image and Salt Mask

[3] https://www.kaggle.com/competitions/tgs-salt-identification-challenge/data

3.3 Data Preparation

3.3.1 Data conversion and loading

The data for this investigation consists of photos with a 101*101 pixel dimension and their respective masks in a separate folder called Masks, sharing the same name. The seismic images in the Train folder are used as training data for the model, and the respective masks are used as training labels for the model. Images are initially translated into grayscale after being read from the dataset directory. The default seismic image proportions are 101*101*3. As the original seismic image is of the size 101*101*3, after this seismic image passes through the max pooling layer, the final output of the model cannot be maintained as the original image. Therefore, the seismic images were converted to the dimension 128*128*3. To feed the initial layer of the neural network, the picture is further condensed to 128*128*1. The images are loaded and then transformed to Numpy Arrays to increase processing flexibility. The values in the image array are normalized to 0.0 and 1.0 taking inspiration from ul Islam (2020). Seismic image masks also go through a similar process.

3.3.2 Data Splitting

Since the TGS Seismic Dataset only contains Train and Test Data, the dataset was divide into Train and Validation datasets in the ratio 80/20.

3.3.3 Data augmentation

There are few seismic photos in the TGS dataset; thus, more seismic images are needed to build a more reliable model. The seismic image dataset was increased using the data augmentation technique of flipping. Left-right and up-down flipping techniques were employed to increase the training dataset from 4000 images to 12000 images, as seen in the Figure 5.

DATA AUGMENTATION

```
[ ]  # Data Augmentation by flipping the Images and Masks up-down and left-right
     X_train_data = np.append(X_train_data, [np.fliplr(x) for x in X], axis=0)
     Y_train_data = np.append(Y_train_data, [np.fliplr(x) for x in Y], axis=0)
     X_train_data = np.append(X_train_data, [np.flipud(x) for x in X], axis=0)
     Y_train_data = np.append(Y_train_data, [np.flipud(x) for x in Y], axis=0)

     del X, Y

     print('Seismic Image Training Dataset shape:', X_train_data.shape)
     print('Seismic Image Validation Dataset shape:', X_eval_data.shape)
     print('Seismic Image Training Dataset shape:', Y_train_data.shape)
     print('Seismic Image Validation Dataset shape:', Y_eval_data.shape)

     Seismic Image Training Dataset shape: (11200, 128, 128, 1)
     Seismic Image Validation Dataset shape: (800, 128, 128, 1)
     Seismic Image Training Dataset shape: (11200, 128, 128, 1)
     Seismic Image Validation Dataset shape: (800, 128, 128, 1)
```

Figure 5: Data Augmentation

4 Design Specification

As discussed in the related work, traditional salt dome detection methods in seismic images focused more on extracting different seismic features to make accurate predictions. However, these methods were cost-intensive, required professional help, and led to many prediction errors. However, in the recent past, deep learning methods have been very successful in image segmentation problems and require no manual intervention. For the specific case of segmenting seismic images to detect salt domes, the deep learning methods have been shown to automatically extract seismic features showing better performance than the traditional methods (Alfarhan, Deriche, Maalej, AlRegib and Al-Marzouqi; 2020). Therefore, the image segmentation deep learning approach has been adopted in this research, where each pixel in the seismic image is classified as a salt-containing pixel or a non-salt-containing pixel. The state-of-the-art model architecture UNET proposed by (Ronneberger et al.; 2015) has been adopted to create the model. The entire design of the research can be seen in Figure 6.

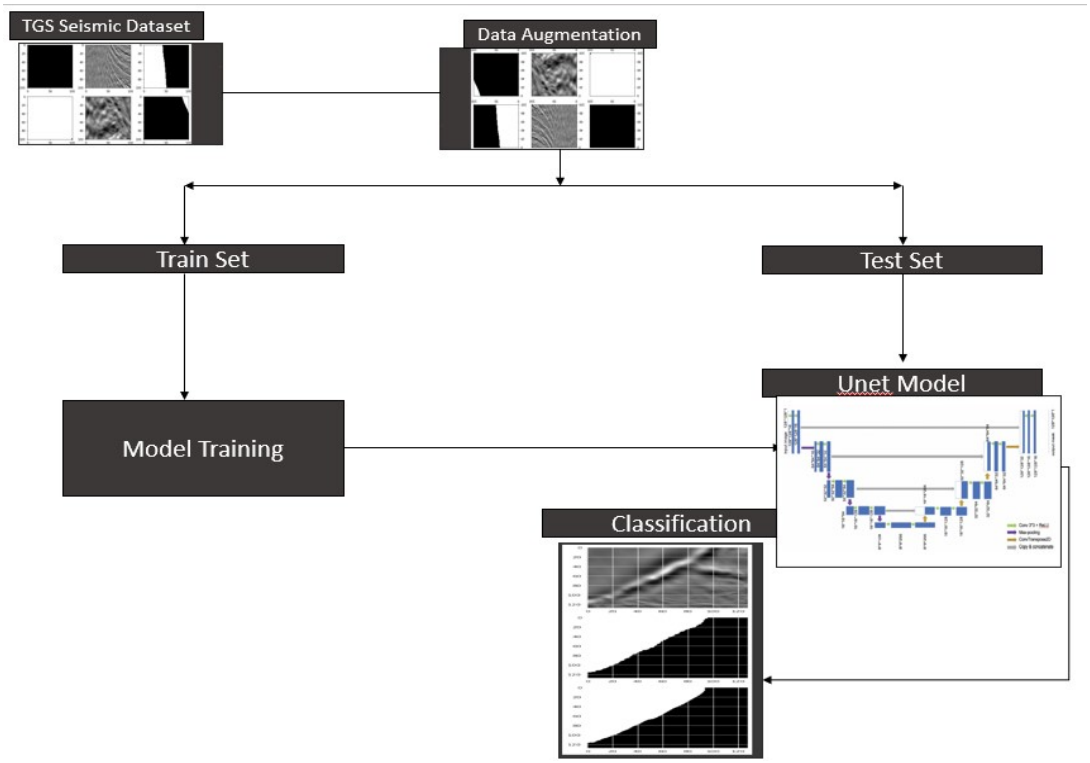

Figure 6: Data Processing

5 Model Implementation

U-Net is an enhanced, fully convolutional network consisting of three main building blocks. The encoder is the contraction part of the Unet architecture and is used to downsample the seismic image. The second component is called the compression layer, a seismic feature vector. The third component, called the Decoder, follows an expanding path and extracts the location of the salt dome in the whole seismic image. The U-Net architecture of down-sampling and then up-sampling the seismic images has been used as

the baseline model in this research. The concept of deep residual learning has been used to overcome the problem of degradation using skip connections (Alfarhan, Deriche, Maalej, AlRegib and Al-Marzouqi; 2020). The detailed implementation of the Unet architecture can be seen in the Figure 7.

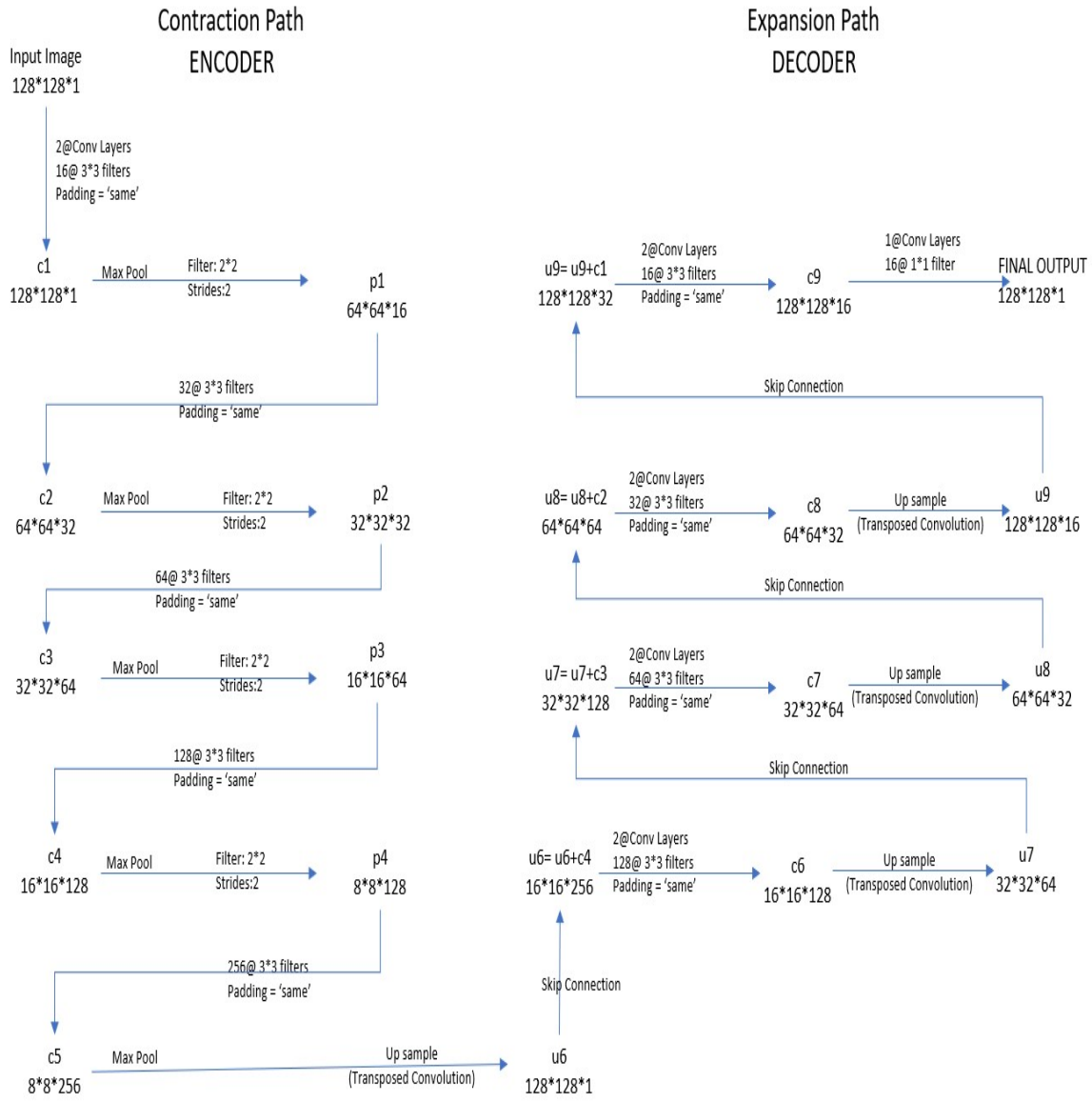

Figure 7: Detailed U-Net Architecture

Figure 8 shows that various operations in the architecture have been performed, such as Convolution, Maximum Pooling, Activation, Dropout, and Transpose Convolution. All of these components and operations have been discussed in detail below:

Convolution Operation: The convolution operation in this model has two inputs: 1. input image of size 128*128*3 where 3 denotes the number of channels. 2. 16 filters, also known as feature extractors of size 3*3. The below formula gives the relationship between the input and output of a convolution block:

$$x(output) = \left| \frac{x(input) + 2p - k}{s} \right| + 1 \tag{1}$$

where k is the number of filters, p is the padding size, and s is the stride size. As seen in the first layer of our Unet architecture, the x(input) of 128*128*3 gets converted to 128*128*16 after going through 2 convolution layers. 2@Conv in the figure means that two consecutive convolution layers are applied in the network. c1, c2, and others are the output tensors after the convolution operations.

Activation Layer: We have used Rectified Linear Unit activation function in the activation layer of the U-Net model as the previous related papers (Liu et al.; 2019) and (Ma and Li; 2021) show that ReLU works best with Unet architecture. Below is the formula for ReLU activation function:

ReLU Formula:

$$f(x) = \begin{cases} 0, if x \leq 0, \\ x, if x > 0 \end{cases} \tag{2}$$

Max Pooling Operation: Max pooling operation is used to decrease the size of feature maps. This is done to keep only the vital features of different regions of the image such that the context of the image is described correctly. This operation takes two inputs as parameters- the filter size and the stride length. The U-net model in this research takes filter size 2*2 and the stride as 2, as already seen in the architecture. p1, p2, and others are the outputs of the max pooling operations in the U-net architecture used in this research. The visual representation of the max pooling operation can be seen in the Figure 8.

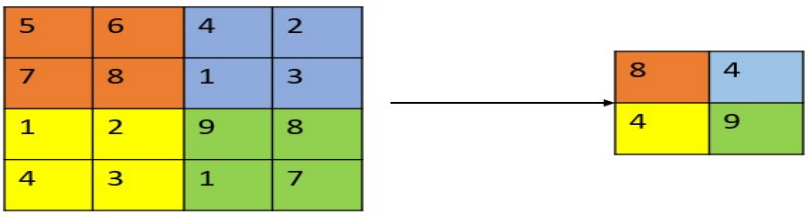

Figure 8: Max Pooling Operation with filter 2x2 and stride 2

Convolution and Max Pooling Operations are used to down-sample the seismic image to get a better context or receptive field. As seen in the encoder part of the U-net network, the input size decreased from 128*128*3 to 8*8*256.

Dropout Layer: Dropout layers are very vital to avoid the issue of overfitting and to decrease the training time of the deep neural network (Pochet et al.; 2018). The dropout layer is added to our model after each max pooling operation.

Transposed Convolution: After the down-sampling of the seismic image is done in the encoder, the transposed convolution layers are used for upsampling the images with the learnable parameters. The layer converts a low-resolution image to a high-resolution image by performing matrix multiplication of the input image matrix with the transpose of the filter matrix. As seen in the U-net architecture, the layers u6, u7, and so on are the output vectors of transposed convolution layers used for upsampling. A transpose convolution example can be seen in Figure 9.

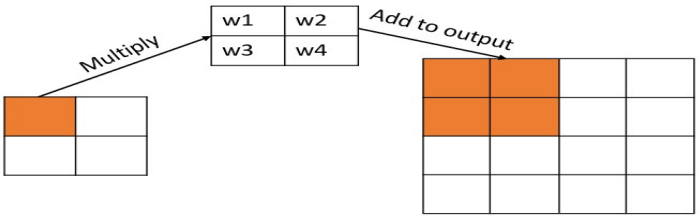

Figure 9: Transposed Convolution

Skip connections: We have used connections (Ronneberger et al.; 2015) in the model architecture to skip specific layers in the network and directly transfer the output of a layer to the last layers in the network. The benefit of using skip connections in the U-net architecture is that the trained features in the encoder part of the network can be directly used in the decoder part to up-sample the seismic image. Therefore, we have concatenated the outputs from transposed convolution layers with the learned feature maps gathered from the encoder. As seen in the U-net architecture, the layer u6(transposed convolution layer) = u6 + c4(feature map from the encoder layer).

6 Model Training

The U-net model was compiled using the Adam Optimizer as it showed the best performance compared to other optimizers. As each pixel of the seismic image could be either a salt-containing pixel or a non-salt-containing pixel, we have the binary cross entropy loss function in the model. Apart from that, Keras callbacks were implemented to train the model. Early Stopping patience was set to 30, so the execution could stop if the validation loss did not improve for 30 successive epochs. The learning rate was set to decay by a factor of 0.1 if validation loss did not improve for five successive epochs. Minimum value of learning rate was set to 1e12. We have used a batch size of 32 to fit the model over 200 epochs.

7 Model Evaluation

The Unet model was trained over 200 epochs and a batch size of 32. Each epoch consisted of 350 steps. The Keras callback functionality of early stopping could stop the model training at the 73rd epoch because the validation loss did not improve for 30 successive epochs.

```
Epoch 68: ReduceLROnPlateau reducing learning rate to 1.0000000082740371e-09.
350/350 [==============================] - 75s 215ms/step - loss: 0.0861 - acc: 0.9670 - iou_metric: 0.7749 - val_loss: 0.0953 - val_acc: 0.9635 - val_iou_metric: 0.7940 - lr: 1.0000e-08
Epoch 69/200
350/350 [==============================] - 75s 215ms/step - loss: 0.0867 - acc: 0.9669 - iou_metric: 0.7729 - val_loss: 0.0954 - val_acc: 0.9636 - val_iou_metric: 0.7956 - lr: 1.0000e-09
Epoch 70/200
350/350 [==============================] - 76s 218ms/step - loss: 0.0855 - acc: 0.9671 - iou_metric: 0.7739 - val_loss: 0.0953 - val_acc: 0.9636 - val_iou_metric: 0.7935 - lr: 1.0000e-09
Epoch 71/200
350/350 [==============================] - 76s 218ms/step - loss: 0.0852 - acc: 0.9675 - iou_metric: 0.7732 - val_loss: 0.0953 - val_acc: 0.9636 - val_iou_metric: 0.7959 - lr: 1.0000e-09
Epoch 72/200
350/350 [==============================] - 75s 215ms/step - loss: 0.0853 - acc: 0.9673 - iou_metric: 0.7739 - val_loss: 0.0952 - val_acc: 0.9637 - val_iou_metric: 0.7947 - lr: 1.0000e-09
Epoch 73/200
350/350 [==============================] - ETA: 0s - loss: 0.0851 - acc: 0.9673 - iou_metric: 0.7736Restoring model weights from the end of the best epoch: 43.

Epoch 73: ReduceLROnPlateau reducing learning rate to 1.0000000082740371e-10.
350/350 [==============================] - 76s 216ms/step - loss: 0.0851 - acc: 0.9673 - iou_metric: 0.7736 - val_loss: 0.0953 - val_acc: 0.9636 - val_iou_metric: 0.7959 - lr: 1.0000e-09
Epoch 73: early stopping
```

Figure 10: Model Training

7.1 Evaluation Metrics:

The performance of the Unet model to detect salt domes was evaluated using different model evaluation metrics. These metrics help determine the effectiveness of the model in accurately detecting salt domes in the seismic images.

7.1.1 Intersection over Union(IoU):

The evaluation metric Intersection over Union was used to evaluate the Unet model. IoU evaluation metric was used by (ul Islam; 2020) to find the overlapping percentage of masked training seismic image and the predicted seismic image mask. The formula for IoU is given below:

$$iou = \frac{IntersectionArea}{UnionArea} \tag{3}$$

Figure 11 graph shows the increase in the value of IoU with the increasing epochs. The model achieved the IoU metric value of around 80 percent for both training and validation datasets.

Figure 11: IoU vs Number of Epochs

7.1.2 Accuracy:

The second evaluation metric, Accuracy, was also used in the model to determine the performance. Accuracy is defined as the ratio between the number of correct predictions made by the model and the total number of predictions made by the model. Accuracy evaluation metric was used in earlier research such as (Arsha and Thulasidharan; 2021) to evaluate the performance of the U-Net model to detect salt in seismic images. Figure 12 shows the Accuracy vs. the number of epochs graph. The model achieved an excellent accuracy of around 97% for the training dataset and around 96% for the validation dataset.

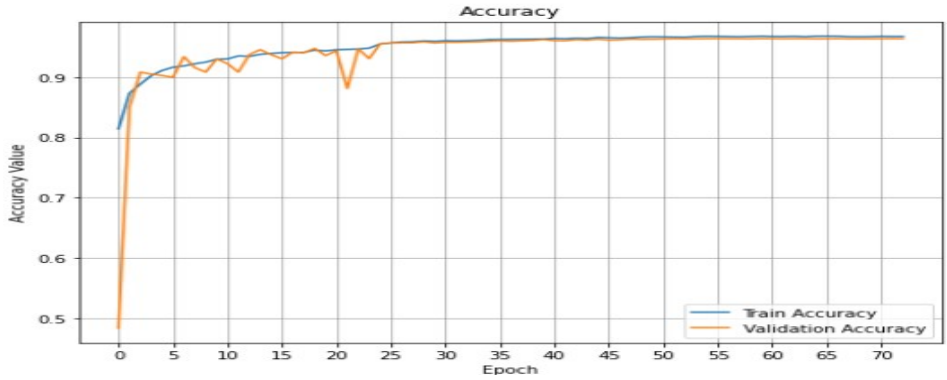

Figure 12: Accuracy vs Number of Epochs

7.1.3 Model Loss:

As seen in the Figure 13, the model loss curves show good performance implying that the model was effectively trained. However, the graph remained stable with the increasing number of epochs, suggesting that the model was not overfitted. Moreover, good convergence was also demonstrated by the gradual decrease in the curve.

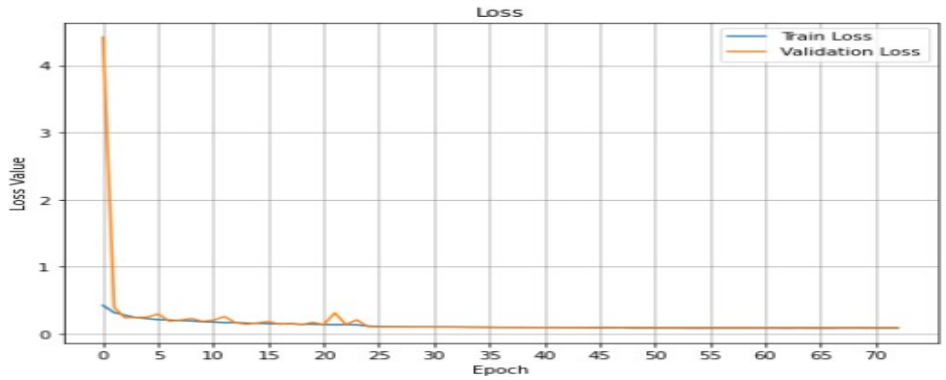

Figure 13: Loss vs Number of Epochs

7.2 Model Parameters:

7.2.1 Loss Function

The model was run using two loss functions: 'Binary Cross Entropy' and 'Weighted Cross Entropy.' Binary Cross Entropy is best suited for datasets where the data is equally distributed among different classes. In contrast, Weighted Cross Entropy is best suited for scenarios where the data in the dataset is skewed. As seen in the Table 2, the Binary Cross Entropy shows better results per the accuracy evaluation metric and is used to create the final model.

Epoch	Binary CE-Train	Binary CE-Valid	Weighted CE-Train	Weighted CE-Valid
1	0.8142	0.4836	0.8386	0.3480
5	0.9111	0.9027	0.8917	0.9078
10	0.9295	0.9289	0.9083	0.9068
15	0.9393	0.9281	0.9372	0.9292
20	0.9434	0.9356	0.9193	0.9319

Table 2: Accuracy Comparison for BCE and WCE

7.2.2 Optimizer

Various optimizers have been used in this research to find the best optimizer to eliminate the issue of overfitting. The optimizers are required in the model to fasten the loss minimization process. The process involves the initialization of weights and also bias, setting optimum values of learning rate and other parameters, and finally correcting the bias. We have used various optimizers like Adam, Adamax, and SGD optimizer while compiling the model and compared the accuracy of each optimizer with the increasing number of epochs. Table 3 compares the accuracy of different optimizers, and it can be seen that the Adam optimizer performs the best.

Epoch	Adam Optimizer	Adamax Optimizer	SGD Optimizer	Nadam Optimizer
1	0.8142	0.8116	0.7092	0.8421
5	0.9111	0.9151	0.8408	0.9256
10	0.9295	0.9342	0.8529	0.9404
15	0.9393	0.9394	0.8533	0.9476
20	0.9434	0.9448	0.8539	0.9591

Table 3: Accuracy Comparison for different optimizers

7.3 Discussion:

The above evaluation metrics prove that the model shows excellent performance in accurately detecting salt in seismic images. The model was used to create seismic masks for the training data images to compare them. The image segmentation results from our Unet model are shown in Figure 14. The first column in the figure shows the actual seismic image from the validation dataset in grayscale format. In contrast, the second column in the figure shows the seismic image salt masks for the seismic images in the first column. Finally, the third column shows the predicted salt masks by the trained

unet model. As seen in the figure, our unet model successfully correctly classifies the salt regions in the seismic images.

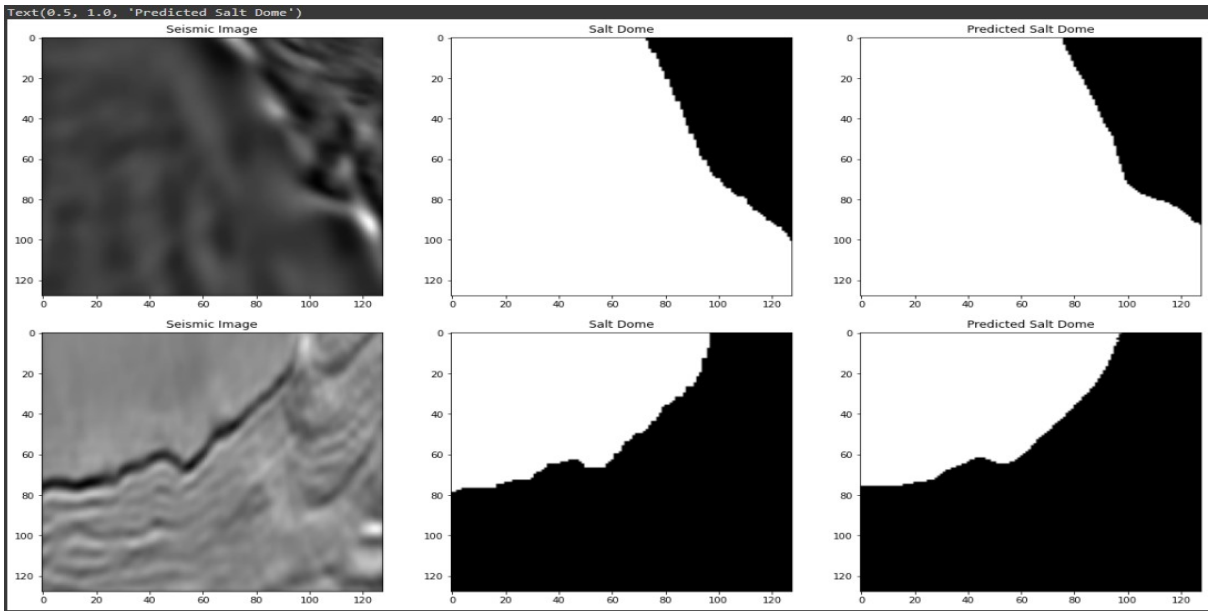

Figure 14: Loss vs Number of Epochs

The results of our model were compared with previous research works in detecting salt in seismic images. Our model was able to achieve the good accuracy when compared to the accuracy claimed by the authors of the previous related papers. Alfarhan, Maalej and Deriche (2020) were able to achieve higher accuracy of 99% as listed in their research paper, but the Netherland F3 dataset they used contained limited seismic images, and thus the performance can not be generalized.

8 Conclusion

In this research paper, we created a deep learning model-Unet to classify salt regions in seismic images correctly. The conventional methods to classify salt regions required seismic features that were hand-crafted by the experts, which often induced errors in the method and was also a time taking process. This Unet model was used with residual blocks to create the model architecture. The dataset was increased by using data augmentation techniques to create a more robust model. Various optimizers and loss functions were used in the training phase to compare the performance. The various model hyper-parameters such as learning rate, epochs, and filter weights were optimized to avoid over-fitting and get optimal results. Various performance evaluation metrics such as IoU were implemented to assess and compare the model's performance. The Unet model used in this research can be beneficial for researchers in the future who are facing issues finding a large dataset to segment seismic images. Apart from that, this model is highly cost-effective for the oil and gas industry as it can automatically classify salt regions without the requirement of an intervention by experts. In the future, the performance of this model can be assessed on a dataset consisting of complex seismic images. Moreover, the impact of using transfer learning where state-of-the-art pre-trained

networks such as Restnet50 and VGG19 can be used in the encoder part of the U-net network can be studied, and the results can be compared.

9 Acknowledgement

We would like to acknowledge and express gratitude toward the project supervisor Mr. Taimur Hafeez for continuously guiding and pushing us to achieve the targets throughout the period of this research.

References

Alfarhan, M., Deriche, M. and Maalej, A. (2020). Robust concurrent detection of salt domes and faults in seismic surveys using an improved unet architecture, *IEEE Access* .

Alfarhan, M., Deriche, M., Maalej, A., AlRegib, G. and Al-Marzouqi, H. (2020). Multiple events detection in seismic structures using a novel u-net variant, *2020 IEEE International Conference on Image Processing (ICIP)*, IEEE, pp. 2900–2904.

Alfarhan, M., Maalej, A. and Deriche, M. (2020). Concurrent detection of salt domes and faults using resnet with u-net, *2020 6th Conference on Data Science and Machine Learning Applications (CDMA)*, IEEE, pp. 118–122.

Aqrawi, A. A., Boe, T. H. and Barros, S. (2011). Detecting salt domes using a dip guided 3d sobel seismic attribute, *SEG Technical Program Expanded Abstracts 2011*, Society of Exploration Geophysicists, pp. 1014–1018.

Arsha, P. and Thulasidharan, P. P. (2021). Salt body segmentation in seismic images using mask r-cnn, *2021 International Conference on Communication, Control and Information Sciences (ICCISc)*, Vol. 1, IEEE, pp. 1–6.

Asjad, A. and Mohamed, D. (2015). A new approach for salt dome detection using a 3d multidirectional edge detector, *Applied Geophysics* **12**(3): 334–342.

Badrinarayanan, V., Kendall, A. and Cipolla, R. (2017). Segnet: A deep convolutional encoder-decoder architecture for image segmentation, *IEEE transactions on pattern analysis and machine intelligence* **39**(12): 2481–2495.

Di, H., Wang, Z. and AlRegib, G. (2018). Why using cnn for seismic interpretation? an investigation, *2018 SEG International Exposition and Annual Meeting*, OnePetro.

El Zini, J., Rizk, Y. and Awad, M. (2019). A deep transfer learning framework for seismic data analysis: A case study on bright spot detection, *IEEE Transactions on Geoscience and Remote Sensing* **58**(5): 3202–3212.

Guo, J., Xu, L., Ding, J., He, B., Dai, S. and Liu, F. (2020). A deep supervised edge optimization algorithm for salt body segmentation, *IEEE Geoscience and Remote Sensing Letters* **18**(10): 1746–1750.

Hegazy*, T. and AlRegib, G. (2014). Texture attributes for detecting salt bodies in seismic data, *SEG Technical Program Expanded Abstracts 2014*, Society of Exploration Geophysicists, pp. 1455–1459.

Jing, Z., Yanqing, Z., Zhigang, C. and Jianhua, L. (2007). Detecting boundary of salt dome in seismic data with edge-detection technique, *SEG Technical Program Expanded Abstracts 2007*, Society of Exploration Geophysicists, pp. 1392–1396.

Liu, B., Jing, H., Li, J., Li, Y., Qu, G. and Gu, R. (2019). Image segmentation of salt deposits using deep convolutional neural network, *2019 IEEE International Conference on Systems, Man and Cybernetics (SMC)*, IEEE, pp. 3304–3309.

Lomask, J., Clapp, B. and Biondi, B. (2006). Parallel implementation of image segmentation for tracking 3d salt boundaries, *68th EAGE Conference and Exhibition incorporating SPE EUROPEC 2006*, European Association of Geoscientists & Engineers, pp. cp–2.

Ma, Z. and Li, Y. (2021). Rotated-unet: A seismic fault identification network based on inverse sampling block construction, *2021 3rd International Academic Exchange Conference on Science and Technology Innovation (IAECST)*, IEEE, pp. 740–744.

Pochet, A., Diniz, P. H., Lopes, H. and Gattass, M. (2018). Seismic fault detection using convolutional neural networks trained on synthetic poststacked amplitude maps, *IEEE Geoscience and Remote Sensing Letters* **16**(3): 352–356.

Ronneberger, O., Fischer, P. and Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation, *International Conference on Medical image computing and computer-assisted intervention*, Springer, pp. 234–241.

Shi, J. and Malik, J. (2000). Normalized cuts and image segmentation, *IEEE Transactions on pattern analysis and machine intelligence* **22**(8): 888–905.

Shi, Y., Wu, X. and Fomel, S. (2018). Automatic salt-body classification using a deep convolutional neural network, *SEG Technical Program Expanded Abstracts 2018*, Society of Exploration Geophysicists, pp. 1971–1975.

Tschannen, V., Delescluse, M., Ettrich, N. and Keuper, J. (2020). Extracting horizon surfaces from 3d seismic data using deep learning, *Geophysics* **85**(3): N17–N26.

ul Islam, M. S. (2020). Using deep learning based methods to classify salt bodies in seismic images, *Journal of Applied Geophysics* **178**: 104054.

Waldeland, A. U., Jensen, A. C., Gelius, L.-J. and Solberg, A. H. S. (2018). Convolutional neural networks for automated seismic interpretation, *The Leading Edge* **37**(7): 529–537.

Woods, R. (1992). Digital image processing, 2nd, *Addision-Wesley* .

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): O97–O103.

Zeng, Y., Jiang, K. and Chen, J. (2019). Automatic seismic salt interpretation with deep convolutional neural networks, *Proceedings of the 2019 3rd international conference on information system and data mining*, pp. 16–20.

Zhang, Y., Liu, Y., Zhang, H. and Xue, H. (2019). Seismic facies analysis based on deep learning, *IEEE Geoscience and Remote Sensing Letters* **17**(7): 1119–1123.