Salt Identification Challenge

Abstract

Seismic survey is critical for exploring oil and gas assets, both onshore and offshore. The survey is similar to an ultra-sound; however, seismic wavelengths are typically much larger, sometimes in the order of 100m. Large wavelengths have some physical implications, but for now, we don't have to deal with that. It's just something to keep in mind while thinking about resolution.

Imaging salt has been a huge topic in the seismic industry. The Society of Exploration geophysicist alone has over 10,000 publications on salt imaging. Salt bodies are important for the hydrocarbon industry, as they usually form nice oil traps. Therefore, there's a clear motivation to delineate salt bodies in the subsurface. Geophysicist are accustomed to interpreting on 2D or 3D images that have been heavily processed. The standard work of seismic data analysis is open access can be found in the reference [1], [2] and [3].

Unfortunately, knowing where large salt deposits are precisely is very difficult. Professional seismic imaging still requires expert human interpretation of salt bodies. This leads to very subjective, highly variable renderings. More alarmingly, it leads to potentially dangerous situations for oil and gas company drillers. To create the most accurate seismic images, machine learning and computer vision can be used to build an algorithm that automatically and accurately identifies if a subsurface target is salt or not by performing image segmentation. In this project, we have used U-net that have proven to produce excellent results in the medical sciences.

Introduction

A seismic program is expensive and time consuming. Collecting and processing the data can take 12-18 months. However, sophisticated computer imaging of subsurface structures can enhance the likelihood of a successful wildcat well. So, it is used after a play has shown some promise by the basin analysis and aerial surveys. There are three basic types of seismic technologies used to help explore for oil and gas – 2D, 3D and 4D.

Advanced imaging not only helps find and produce oil and gas more efficiently but has advanced the identification and recovery of important subsalt reservoirs in offshore locations such as the Gulf of Mexico and Brazil.

Advances in the computer vision and deep learning can be used to identify salt regions, with minimal human intervention and thereby reduce prospective human errors. This not only increases the monetary value of the asset in form of minimized capital expenditure for the oil and gas industry but, at large, it is possible to reduce the resources spent to drill dry wells thereby reducing the carbon footprint of the petroleum industry.

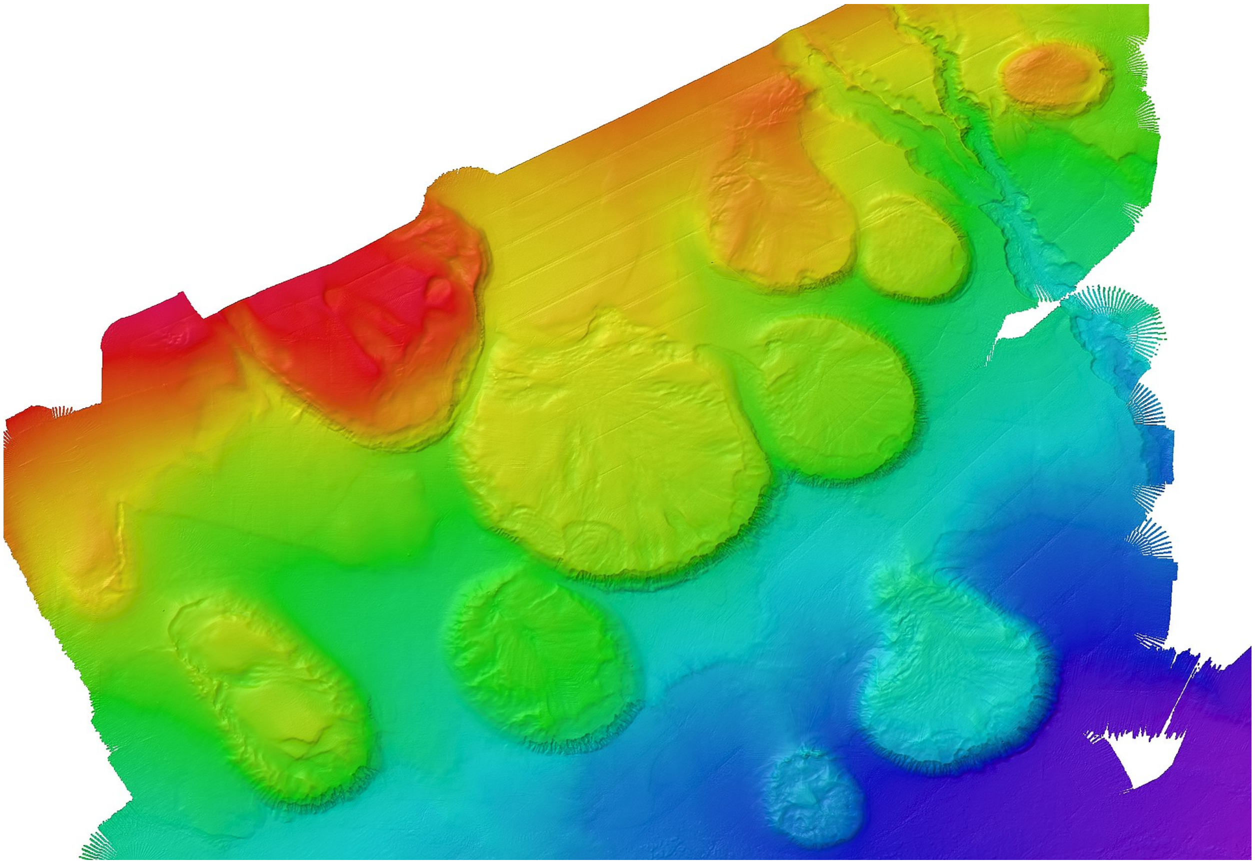

Figure 1. Sub surface salt features from offshore Gulf of Mexico. Courtesy: National Oceanic and Atmospheric Admin.

U-NET architecture, as shown in **Fig.2,** was first developed for the biomedical image segmentation. The underlying idea of this algorithm is to obviate the need for thousands of annotated training samples to train deep networks. Reference [4], presents the network and the training strategy that relies strongly on the use of data augmentation techniques to use the available annotated samples more efficiently. The architecture consists of a contracting path to capture context and a symmetric expanding path that enables precise localization. The paper shows how a U-NET can be trained end-to-end from very few images, with the idea for segmenting the neuronal structures in electron microscopic stacks.

The U-net architecture is synonymous with an encoder-decoder architecture. Essentially, it is a deep-learning framework comprising of two parts:

1. A contracting path similar to an encoder, to capture context via a compact feature map.
2. A symmetric expanding path similar to a decoder, which allows precise localization. This step is done to retain boundary information (spatial information) despite down sampling and max-pooling performed in the encoder stage.

Advantages of Using U-Net:

1. Computationally efficient
2. Trainable with a small data-set
3. Trained end-to-end
4. Works great for image segmentation problems

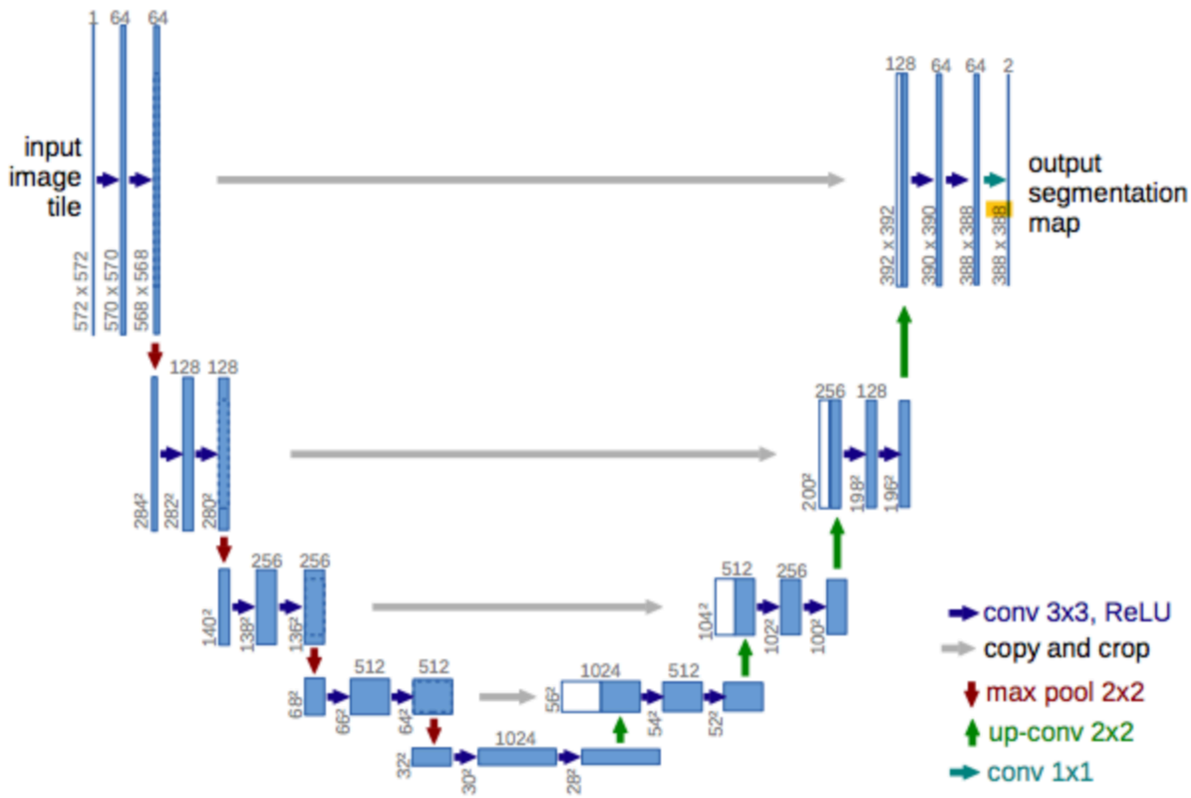

Figure 2. Architecture of U-NET. Ref: Ronneberger et al 2015

Methodology

First phase of the project is the collection of seismic image data that is curated by TGS. This can be done through Kaggle API or via a direct download onto local system. Considering that this project requires use of deep networks and processing of images, that work best with GPU, use of API is highly recommended. A schematic representation of the project workflow is shown in **Fig.3.**

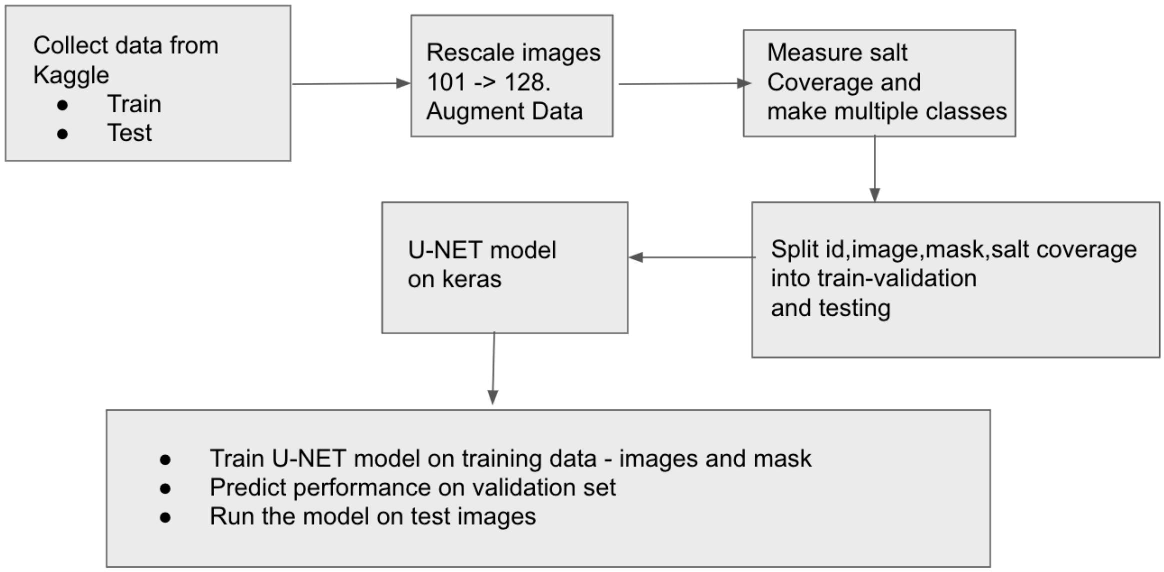

Figure 3. Flow chart describing project methodology

The training images and their respective run length encoded mask are illustrated in **Fig.4**

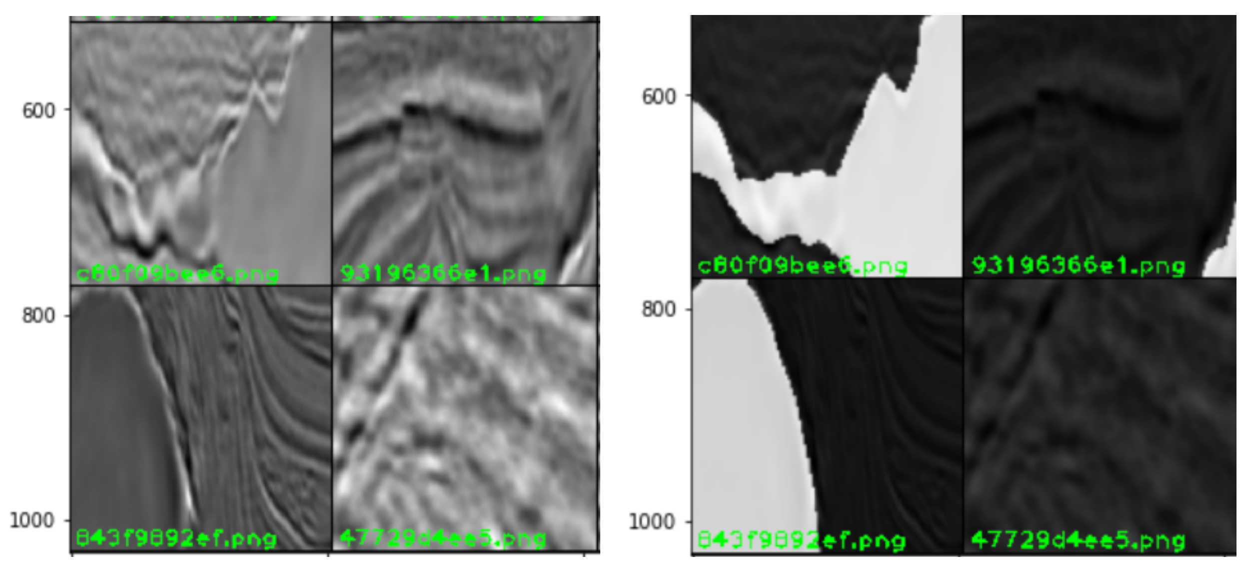

Figure 4. Training Images and Training masks

The distribution plot, **Fig.5**, illustrates the representation of training and testing images at various depths. It's clear that both training and testing have a similar distribution, which shows the training data has a really good mix.

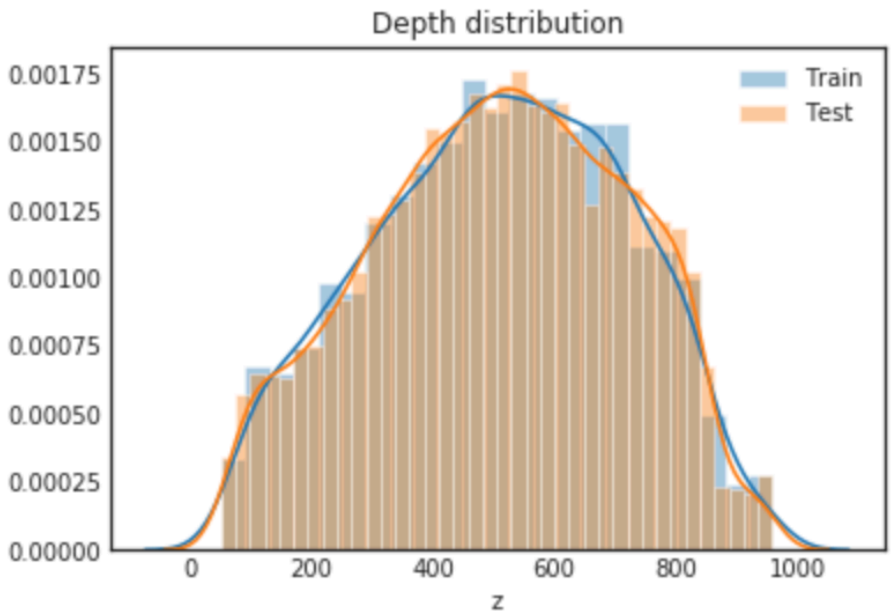

Figure 5. Depth Distribution of training and testing images

The next step involves upscaling the 101*101 sized images to 128*128 to adhere to architecture requirements of the U-Nets. **Fig.6** illustrates the upscaling process.

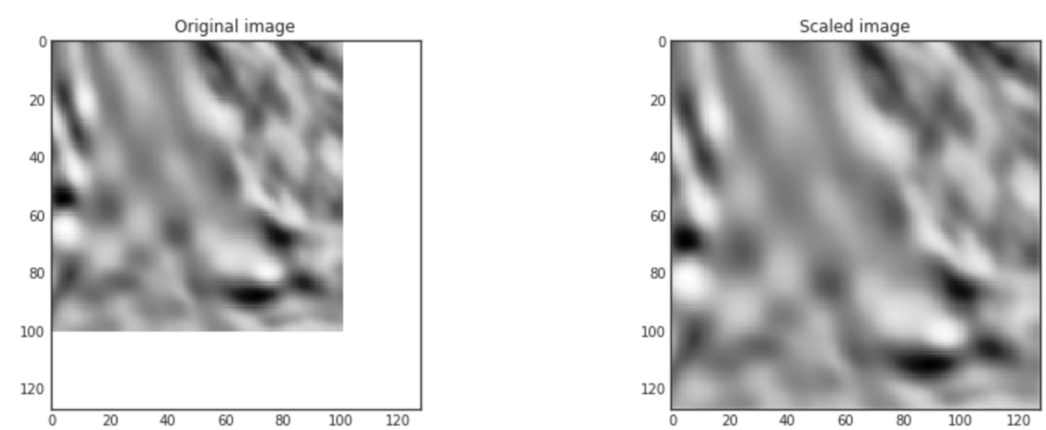

Figure 6. Upscaling images

It is also a good practice to augment the images by making minor alterations to our existing dataset. Minor changes such as flips, translations or rotations are treated as distinct images by the neural network. An example of flipped data is shown in **Fig.7**

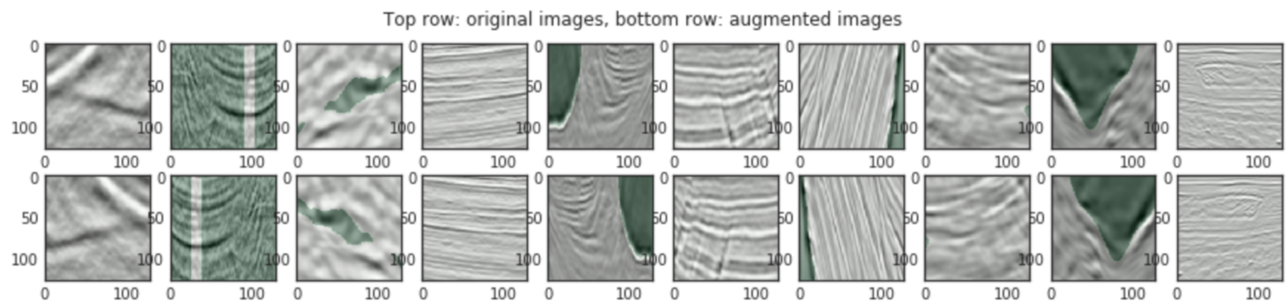

Figure 7. Augmentation of training images

The next step, as shown in **Fig.3,** is evaluating the amount of salt in each image mask. In the original images, as shown in **Fig.4**, the salt bodies are represented as white pixels and the sedimentary layer is represented as non-white regions. Therefore, calculation of salt coverage is a simple fraction of white pixels to all the pixels of the image.

After identification of salt coverage, it's possible to divide the images into multiple salt coverage class on the basis of its "saltiness". **Fig.8** illustrates the division of coverage classes on the basis of salt coverage

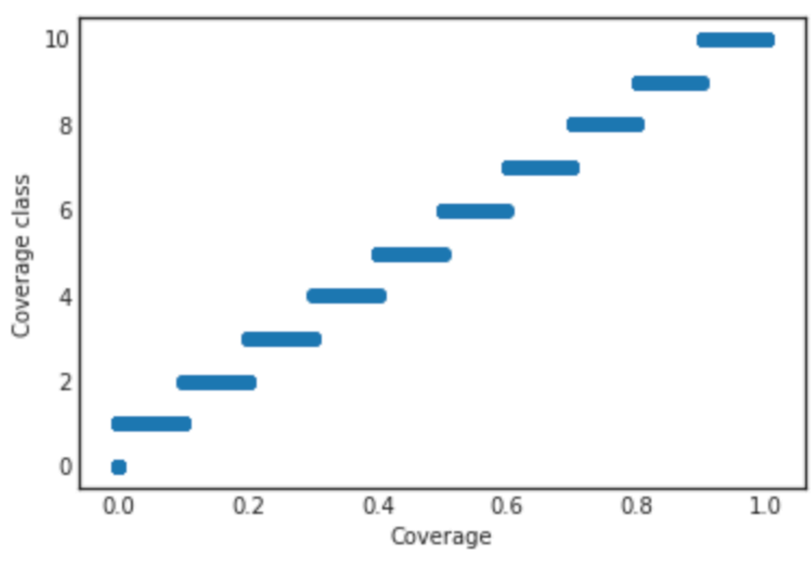

Figure 8. Salt coverage classes vs salt coverage %

Following classification of image "saltiness" into multiple coverage classes, the next step, as shown in **Fig.3**, requires division of the training data into two parts – training and validation. This is followed by developing a U-Net architecture using keras, similar to **Fig.2**.

Results and Discussion

The objective of this project is to identify the salt zone in an image. For such a task, the Intersection over Union (IOU) metrics is the recommended approach to score the performance of a model.

```
Epoch 00004: val_loss improved from 0.34075 to 0.32827, saving model to ./keras.model
Epoch 5/5
6400/6400 [==============================] - 42s 6ms/step - loss: 0.3012 - acc: 0.8788 - val_loss: 0.2834 - v
al_acc: 0.8843
```

Figure 9. Training model performance. (Limited model)

As illustrated in **Fig.10,** the model is able to predict the salt segments from the training images in some cases. This particular result was obtained with the model running with heavy constraints (just 10 epochs). The performance can be greatly enhanced by increasing number of epochs.

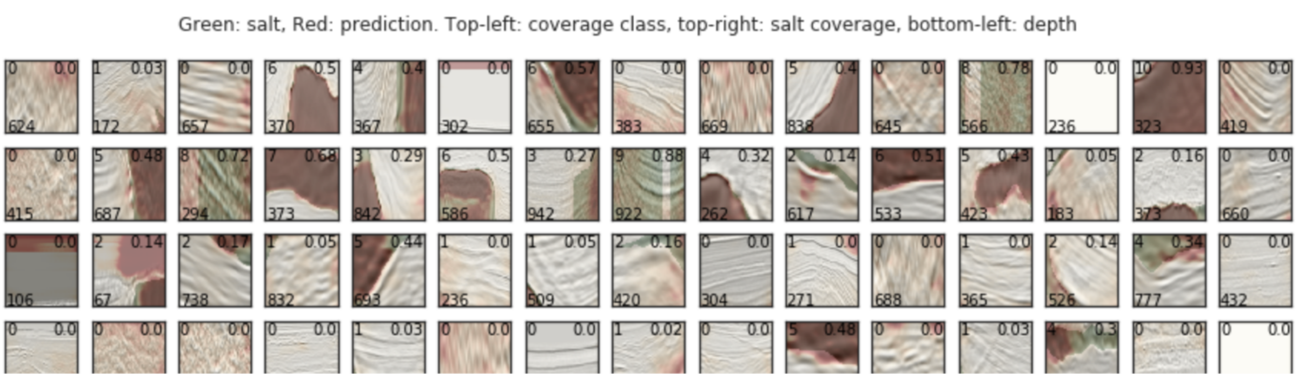

Figure 10. Model performance on the validation set

Even though **Fig.10** is a great way to visualize the results, the performance of the mode is not very clear from the collage of images above. **Fig.11,** that indicates IOU score of 63%, serves really well for that case.

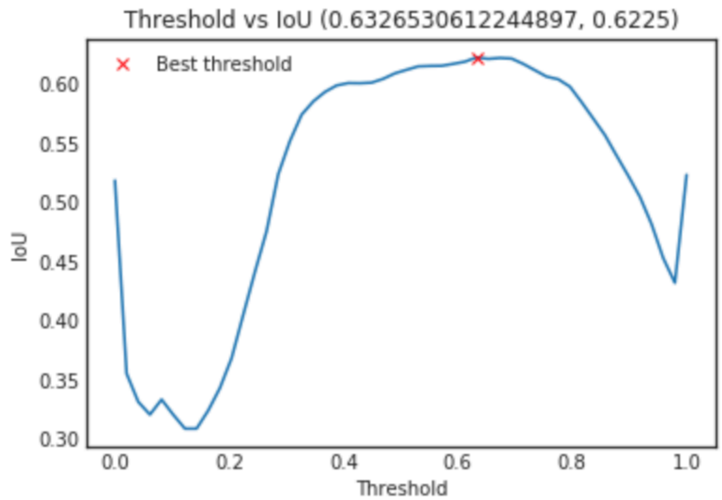

Figure 11. Threshold vs IOU for validation sample

Conclusion

1. A U-Net implementation for semantic segmentation that aims to delineate salts in the seismic images is developed.
2. This developed model will be iteratively improved by incorporating ResNet blocks in the future.
3. This project has been shared on the Kaggle and to the curators of the data, TGS, for feedback. Results pending.

Reference

1. Seismic Data Analysis, Oz Yilmaz
2. Seismic Data Analysis: Processing, Inversion and Interpretation of Seismic Data, Oz Yilmaz
3. Seismic Data Analysis Techniques in Hydrocarbon Exploration, Enwenode Onajite
4. U-Net: Convolutional Networks for Biomedical Image Segmentation, Olaf Ronneberger, Phillip Fischer, Thomas Brox