By:

APGCE ASIA PETROLEUM GEOSCIENCE CONFERENCE & EXHIBITION
'Asia's Premier Geoscience Event'

PETRONAS

GEOHACKATHON 2024

16 – 18 November 2024
Common Ground Bukit Bintang, Kuala Lumpur

Challenge 1

Predict the faults across a given 2D seismic dataset.

Team **DSGS101**

Chun Yan Lim (Reservoir Engineer)
Zulfadhli Mohd Zaki (Geoscientist)
Teo Sheng Pu (Data Engineer)
Max Wei Xiang Ooi (Data Scientist)
Izzudin Hussein (Data Scientist)

Sponsored by:

ConocoPhillips EnQuest ERCE ExxonMobil JX VIRID

01 – Meeting between Nigel and Rachael

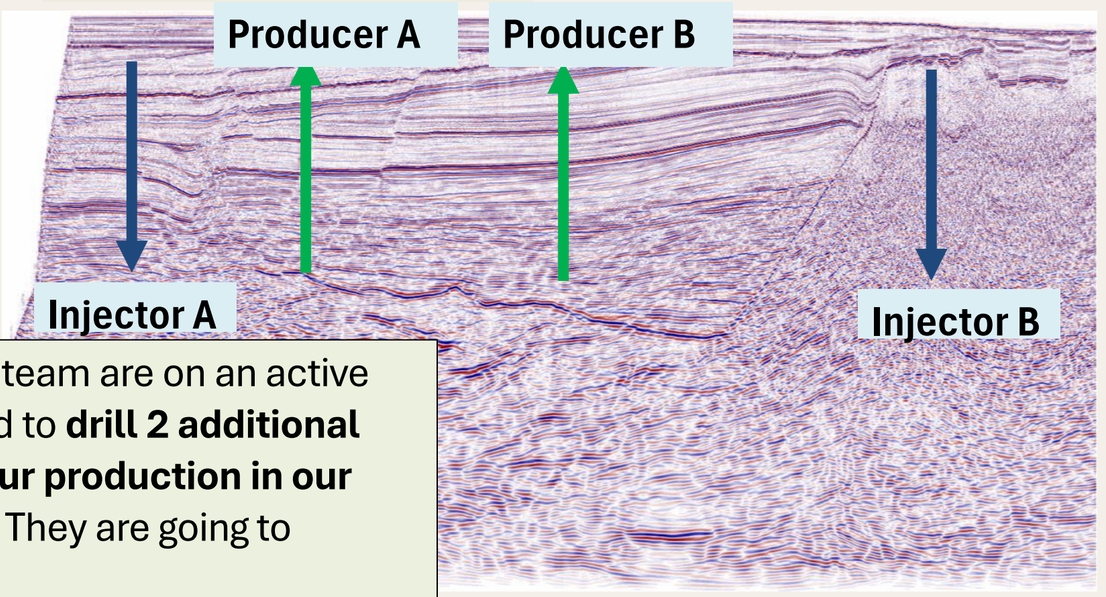
Mr. Nigel, Geoscience Mgr Data Rich Oil Company

Racheal, Ambitious Geoscientist

1. Hey Rachael, the reservoir team are on an active drilling campaign and planned to **drill 2 additional water injectors to support our production in our newly discovered field ZZZ.** They are going to contact the rig in 2 weeks!

2. Hi Mr Nigel. OMG! We have to stop them. Looking at some of the seismic images, Fault Block A and Fault Block B is clearly divided by a fault

But interpreting seismic for the whole field is so **time consuming**. It would take months, if not a year!



3. I have an IDEA. We can leverage our in-house DS team to built an AI tool

We have some interpreted seismic from our older field that we can use as dataset!

Seismic datasets

- 439 seismic images.
- 364 faults images, including 2 faults with missing seismic images.
- Only 362 seismic/faults images used.
- Train/Test split: 66% Train/33% Test

02 – The next day.....

Data Science Team

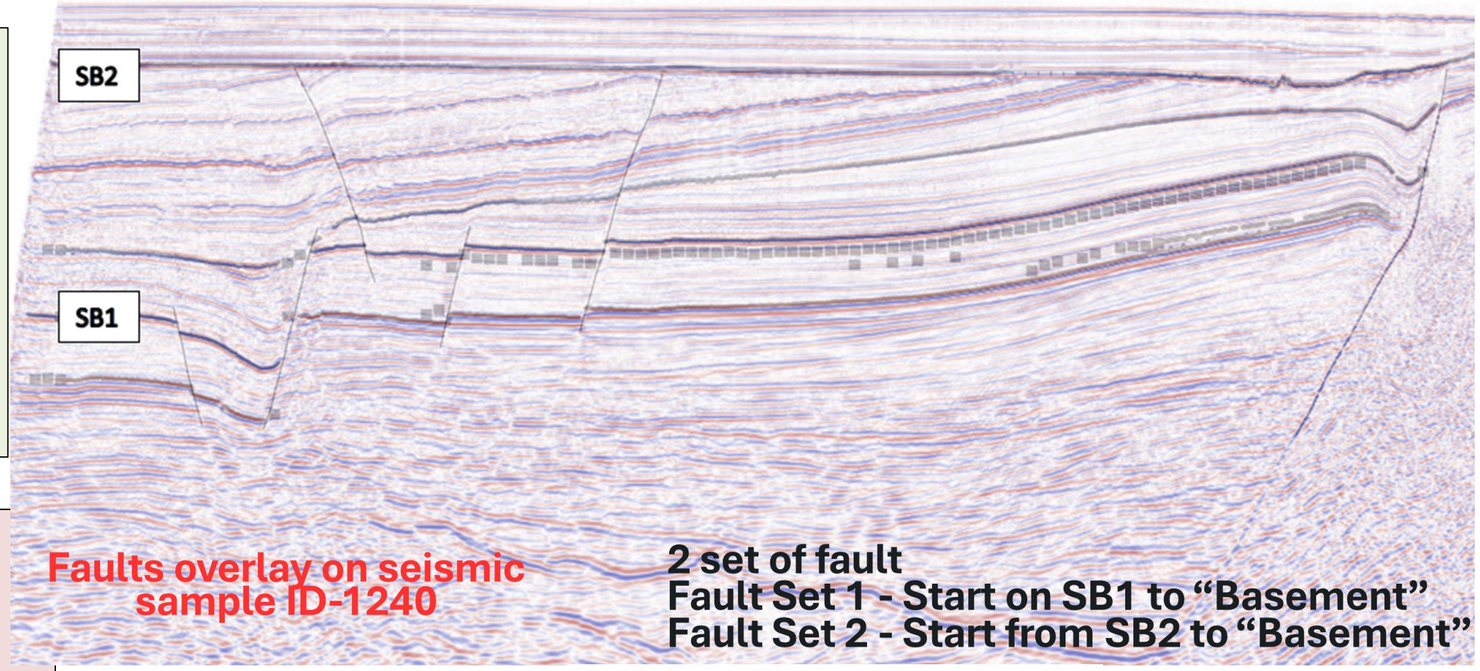
1. Hey Rachael, this is an impossible task! The **given dataset is so imbalance**. There is no incentive to force our model to accurately match the fault

Racheal, Very Ambitious Geoscientist

2. Hi Team!

I think we should pre-process our input data prior to training. Do you think we should **discretize** the dataset?

Also, we should use Focal Loss since the dataset is heavily imbalanced. Maybe hyperparameter tuning on the loss function will help ?



Faults overlay on seismic sample ID-1240

2 set of fault
Fault Set 1 - Start on SB1 to "Basement"
Fault Set 2 - Start from SB2 to "Basement"

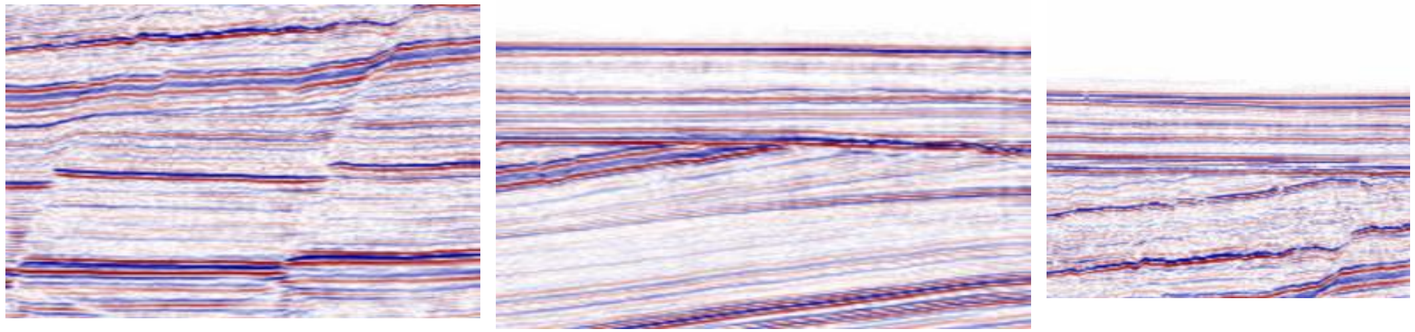


Image Discretization – To increase focus on imbalance dataset and increases the amount of dataset

03 – Methodology (Data Pre-Processing)

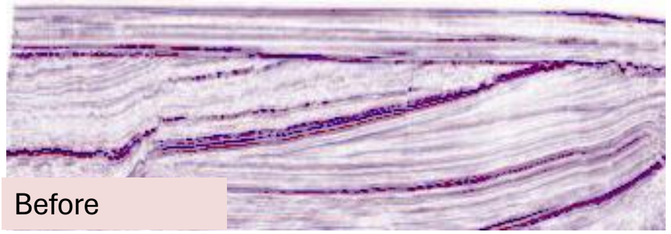
Data Science Team

1. That's a really good idea. We can increase

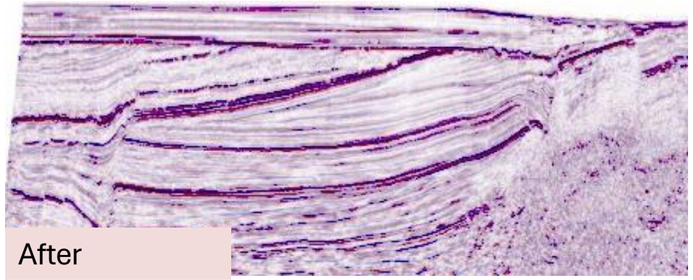
2. What do you think of the idea of increasing the width of the labels ?

It doesn't change the geology but will significantly help our model

Image Enhancement

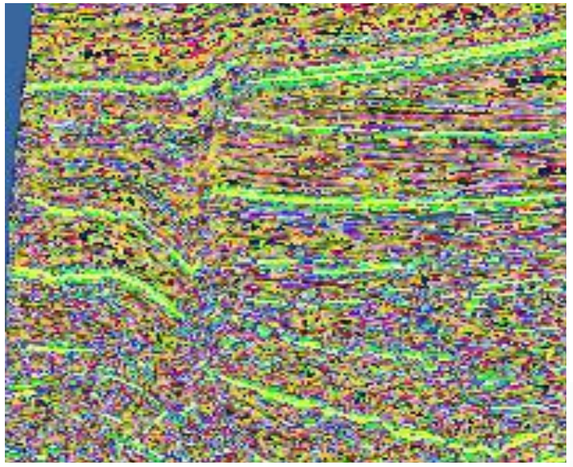


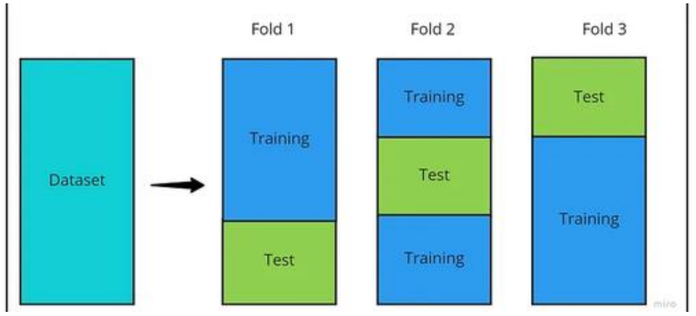
Before



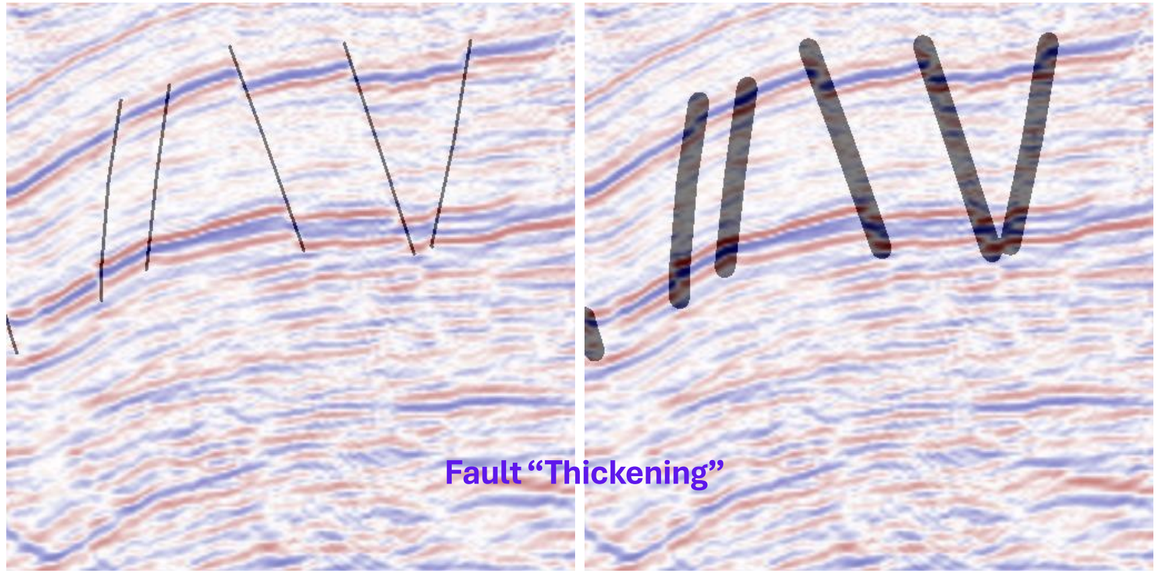
After

Normalization & Contrast





Cross Validation



Fault "Thickening"

Key Differentiators:

1. Tailored Pre-processing steps for seismic image
2. Enhancing Seismic Reflector
3. Wider fault to help interpretability
4. Addressing imbalanced data

04 - Results

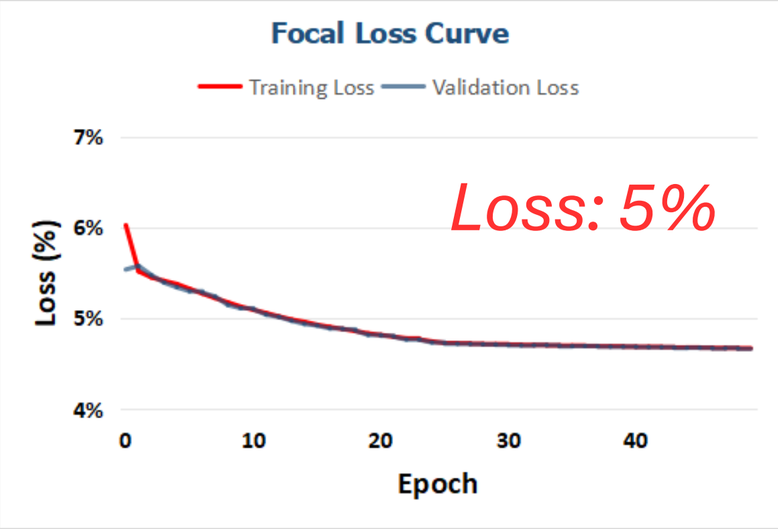
Our Model - **CrackNet**

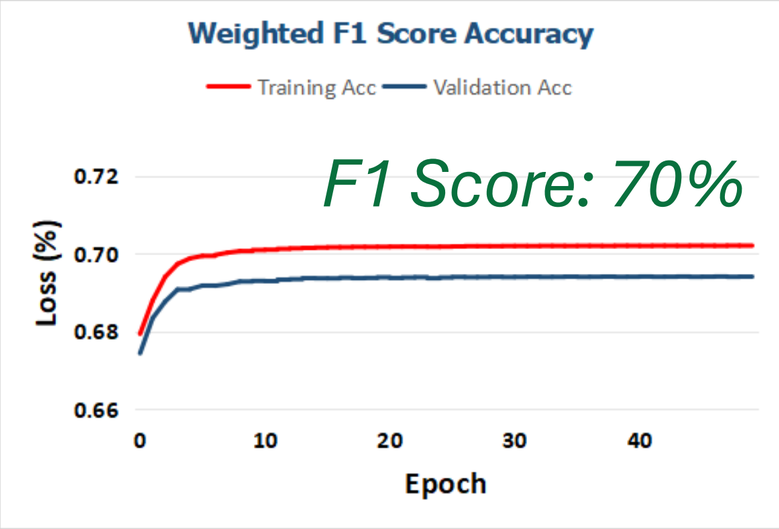
Epochs: 50
Image Size: 512 x 512
Batch Size: 6
Learning Rate: 0.0001
Evaluation Matrix: Weighted F1 Score (Addressing imbalanced data)
Loss: Enhanced Focal Loss with alpha

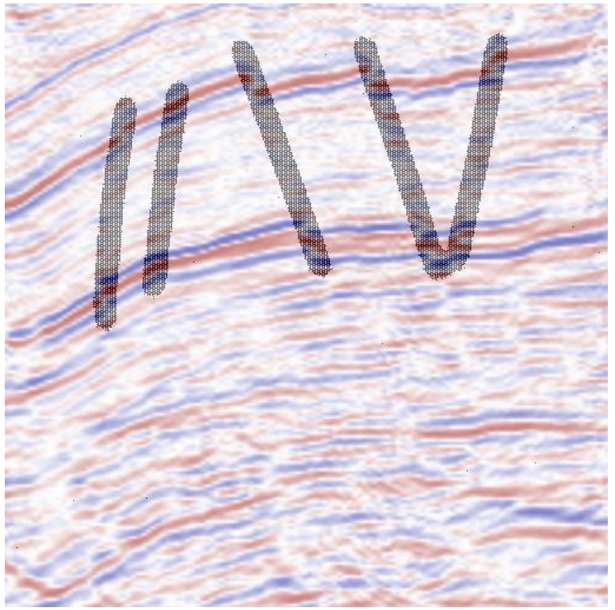
That's really impressive.

All our pre-processing work paid off

The CrackNet model is now able to capture the faults really well



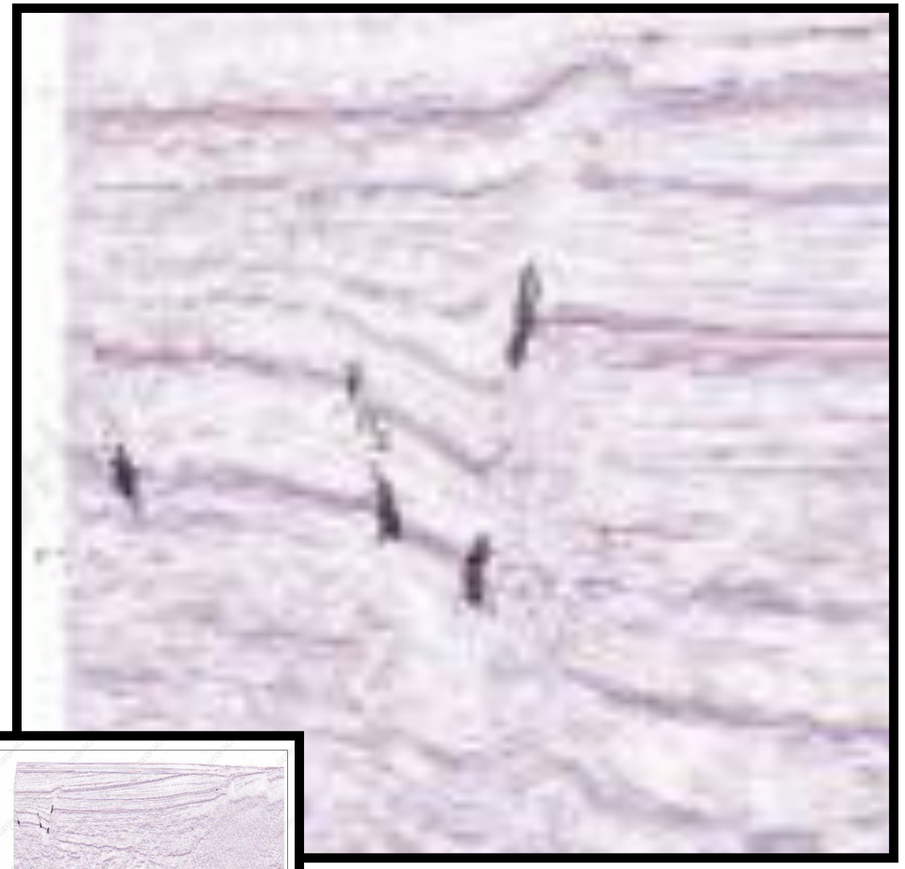


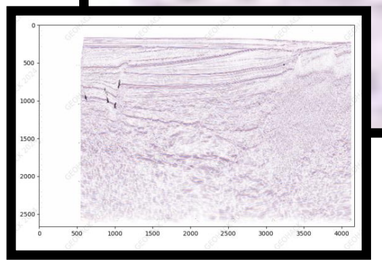


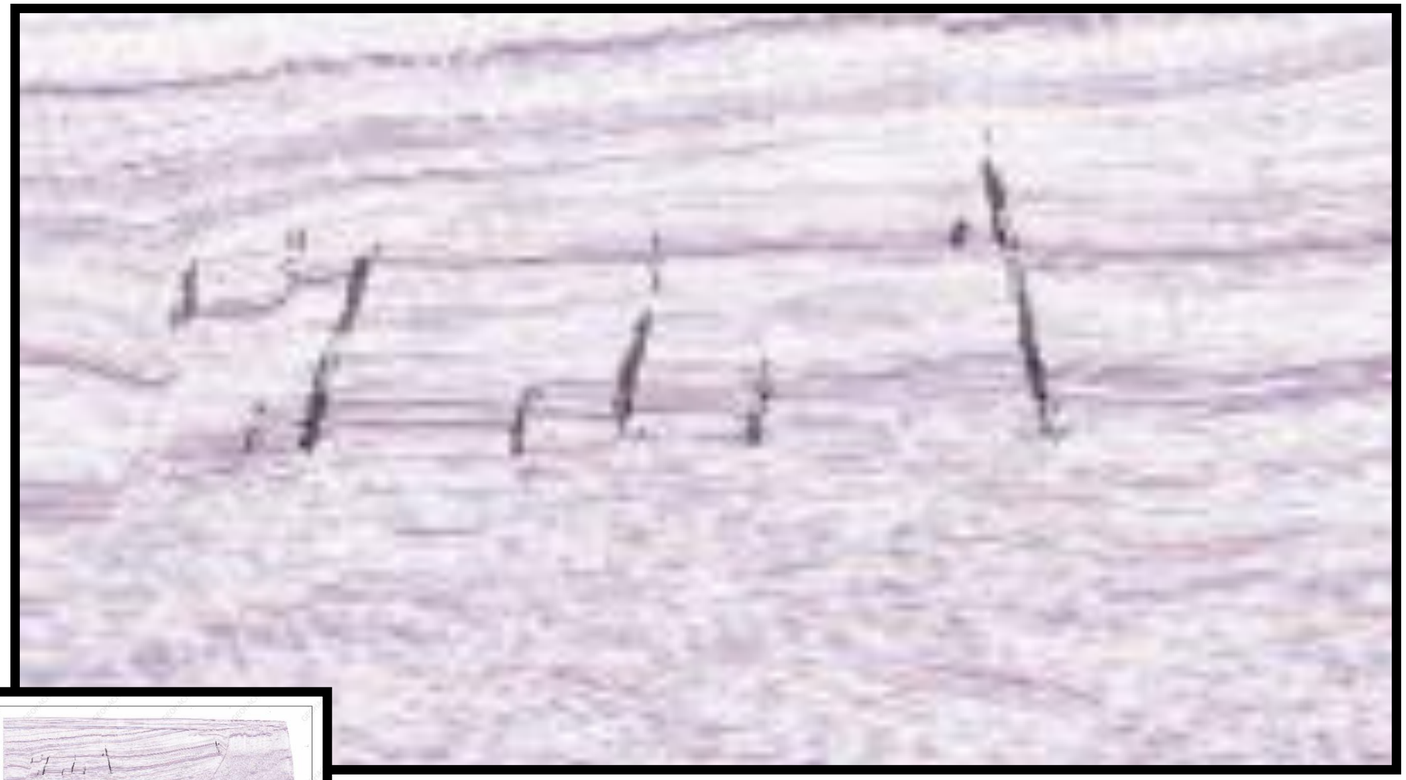
Our model CrackNet performance on the validation dataset with 3 folds cross validation

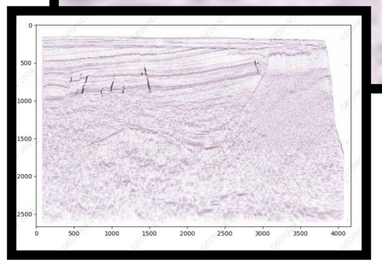
05 - Visualization of results

on holdout data

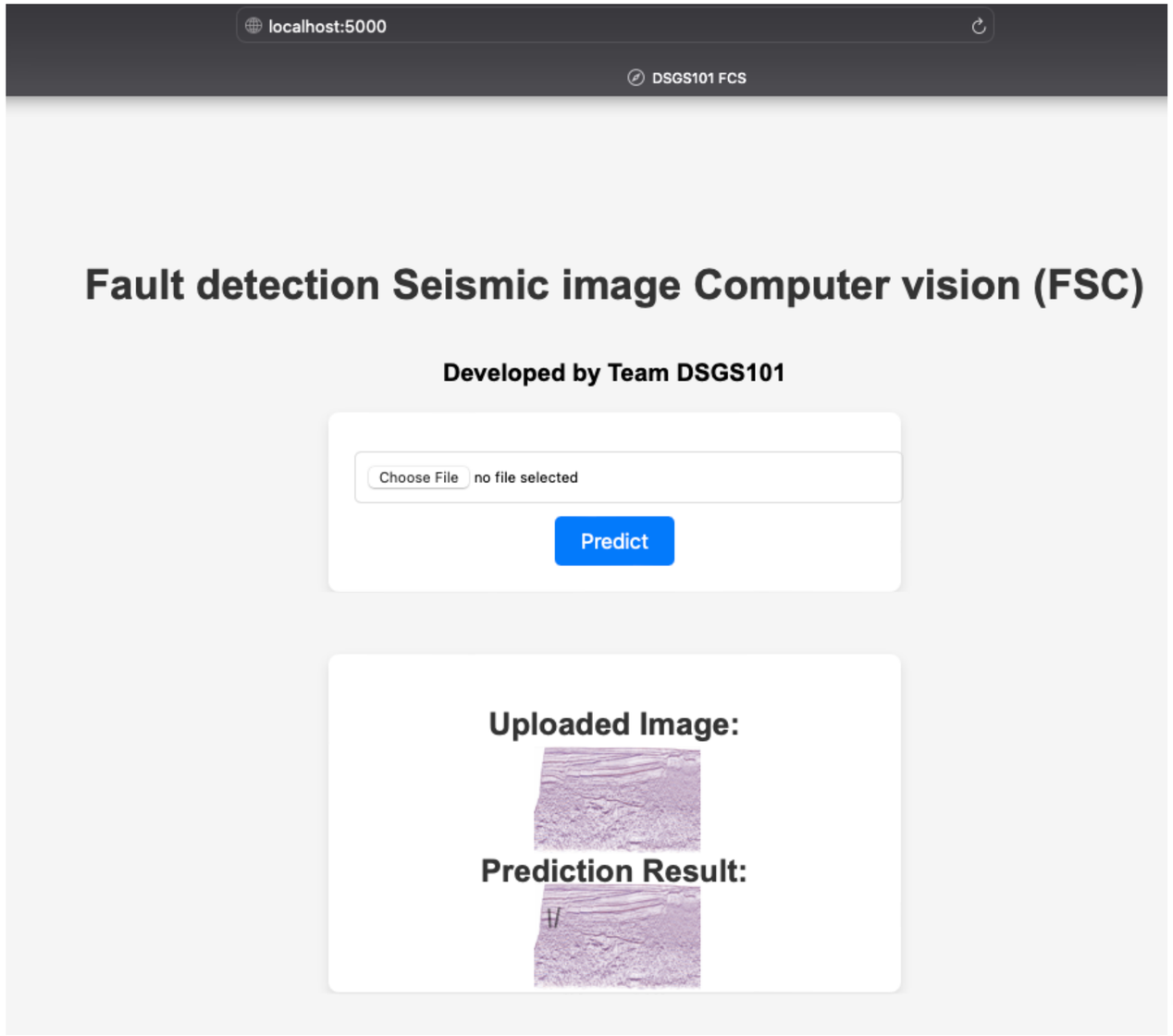








06 - How to interface with the model?



Web application developed with Python Flask framework.

The end user will upload their seismic image into the tool and the model will generate the fault that overlays on top of the seismic image.

1. Great job, team!

Now the reservoir team is able to use this to guide their drilling campaign

Thank you!

03 - Methodology

The HOW?

Data Pre-processing

Data Science Team

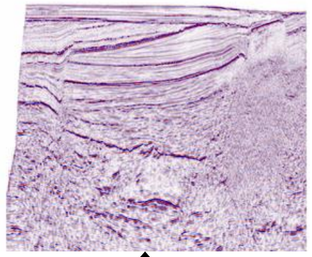
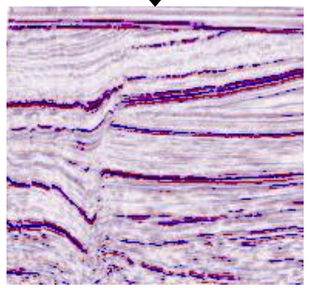


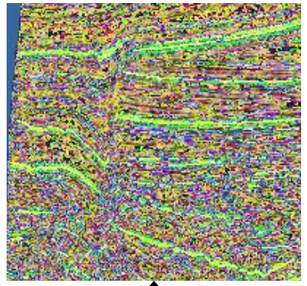
Image Size Reduction

Reduces memory usage

Ensure important feature are fed into the model

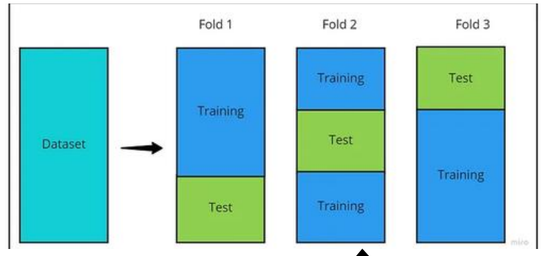
Image Enhancement





Normalization & Contrast

Ensure model compatibility



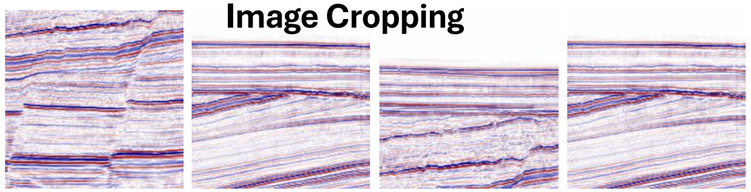
Improve training set label area

"Thickening" Fault

Cross Validation

increase the datasets size and focus on the contours

Image Cropping



Our Differentiator

1. Tailored Pre-processing steps for seismic image
2. Enhancing Seismic Reflector
3. Wider fault to help interpretability
4. Addressing imbalanced data

Our Model - **CrackNet**

Epochs: 50
Image Size: 512 x 512
Batch Size: 6
Learning Rate: 0.0001
Evaluation Matrix: Weighted F1 Score (Addressing imbalanced data)
Loss: Enhanced Focal Loss with alpha

BACK-UP