

[Open]

By:



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

## Challenge 1 Semantic Segmentation of Fault Lines on 2D-Seismic Data

Team 8 – The Fault in Our Stars

Adib Yusof, Kasyful Qaedi, Syaiful Afrizal, Aqil Izham, Fatin Irdina

Sponsored by:



VIRIDIEN



GEOHACKATHON

# Fault detection using CNN



Predicting the fault with more precise and faster processing.



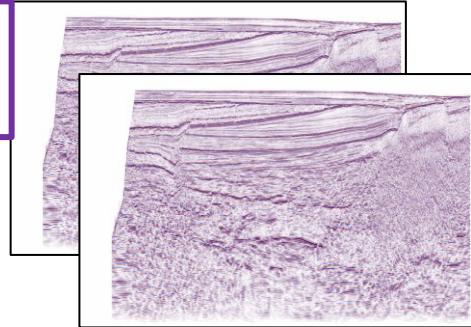
Mitigating risks associated with structural uncertainty.



Minimizing costs and reducing the need for repetitive manual interpretation tasks.

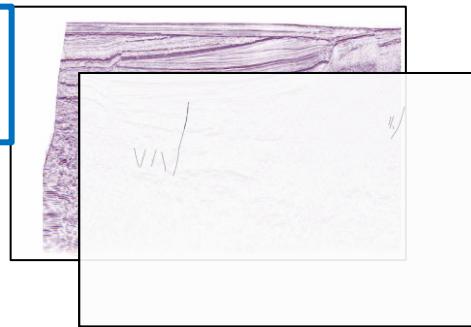
# Challenges

1



**Signal-to-noise ratio – Top to bottom portion of images show different quality**

2



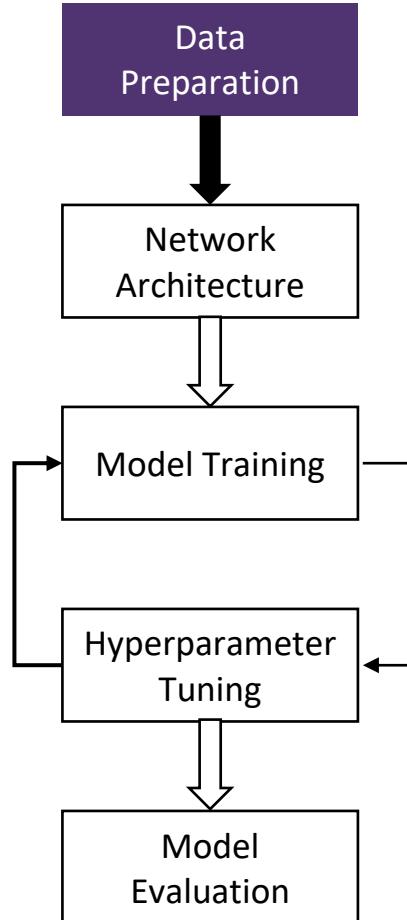
**Insufficient labelled data set.  
441 total of available data for seismic and fault label  
~50% are partially/non labelled**

3

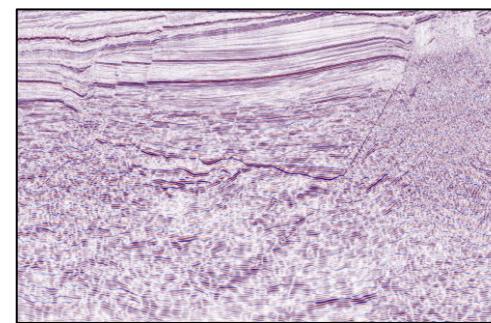
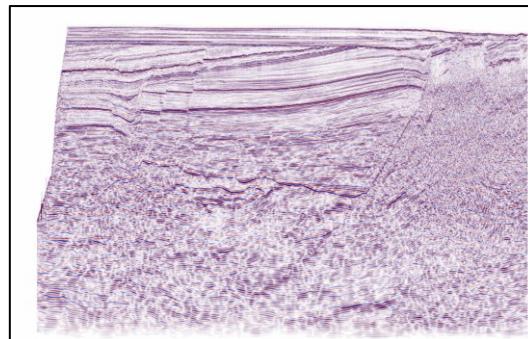


**Imbalanced labelled pixel data (biased).  
4166x2664 pixels and only 1% are labelled**

# Data Preparation



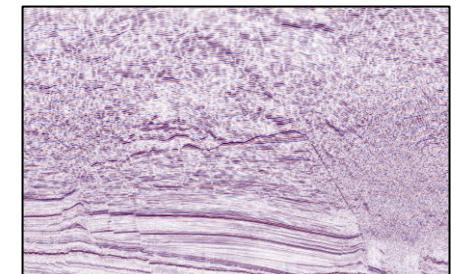
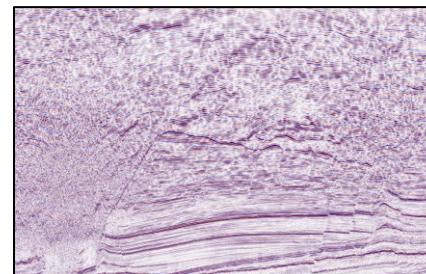
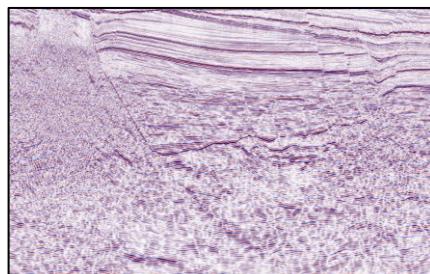
- **Data selection, cropping, augmenting, patching, resizing, normalizing,**



Flip Horizontal

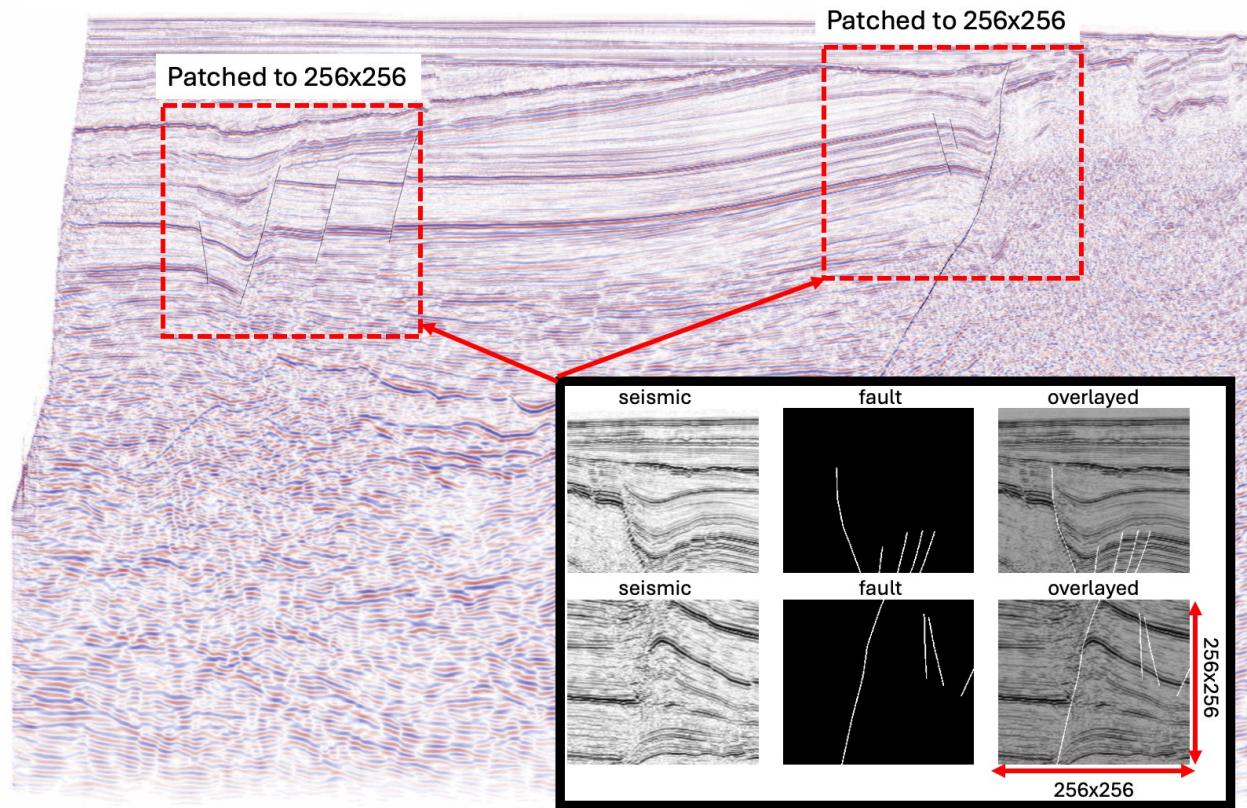
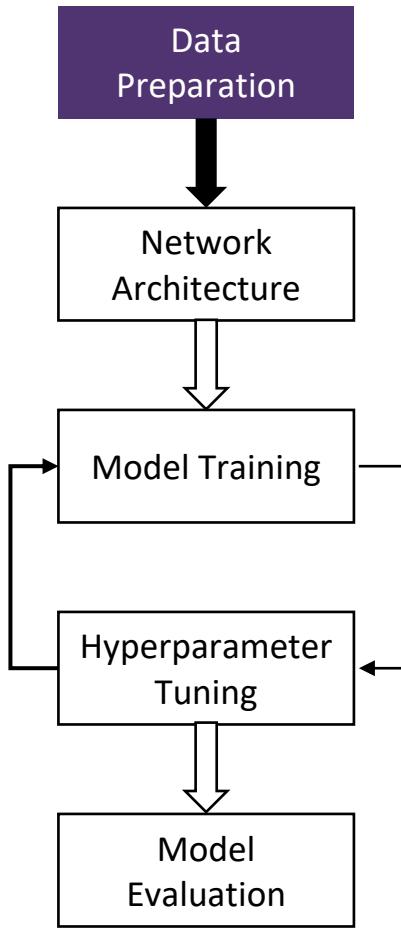
Flip Vertical

Flip Along Both Axes



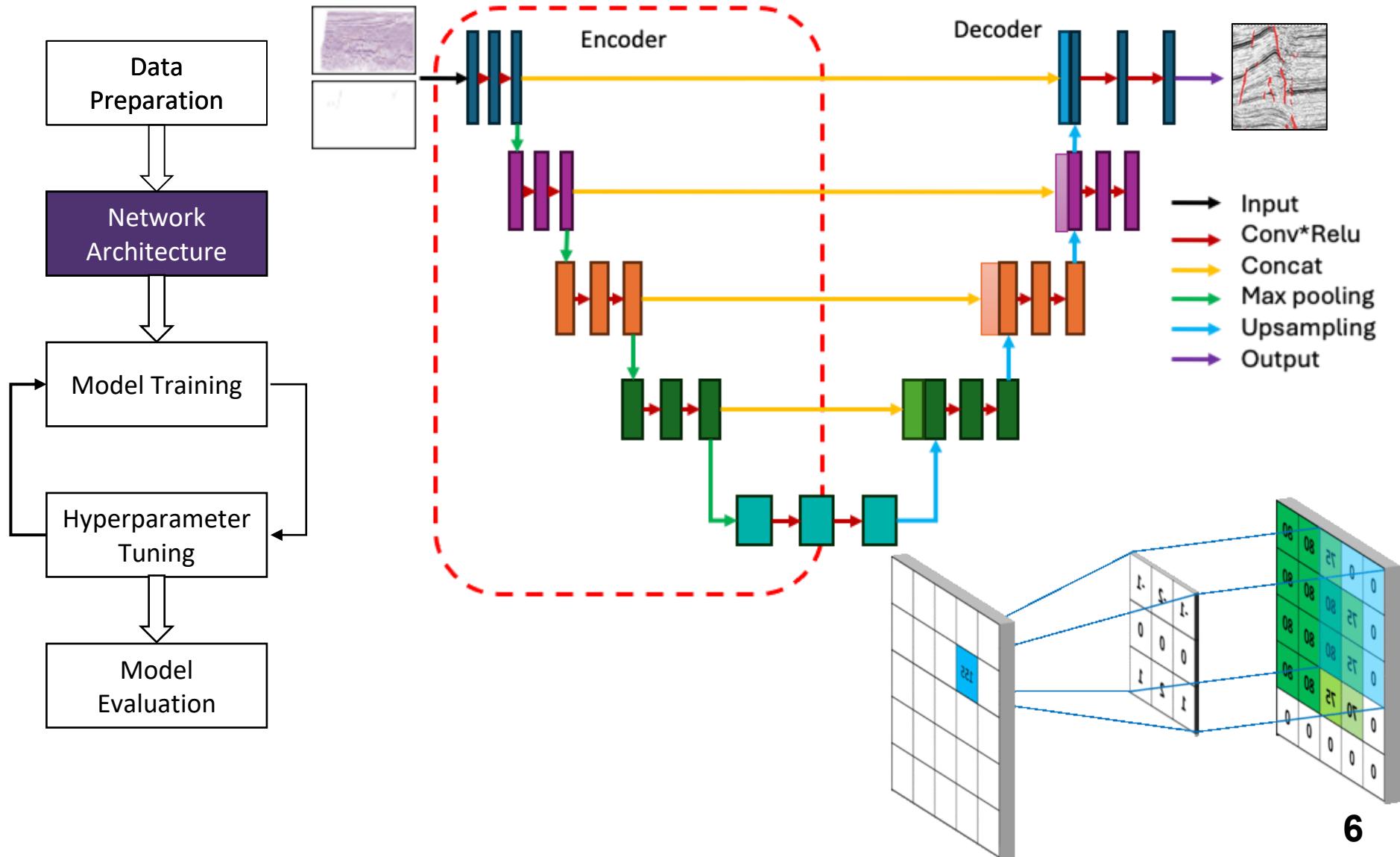
# Data Preparation

- Data selection, cropping, augmenting, **patching, resizing, normalizing**

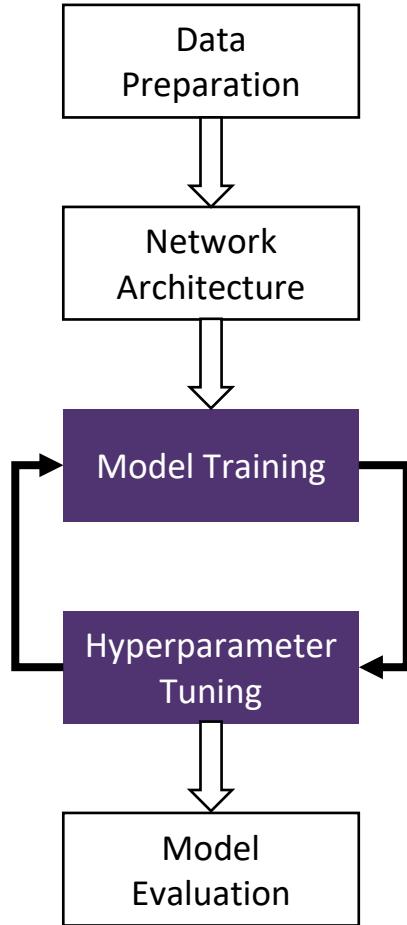




# Network Architecture – U-net

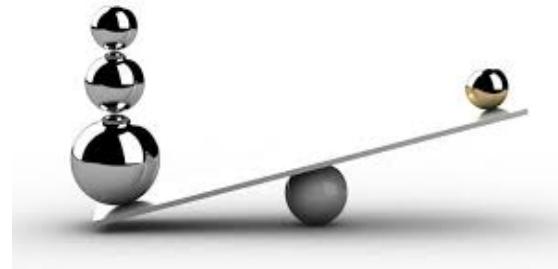


# Model Training & Hyperparameter Tuning



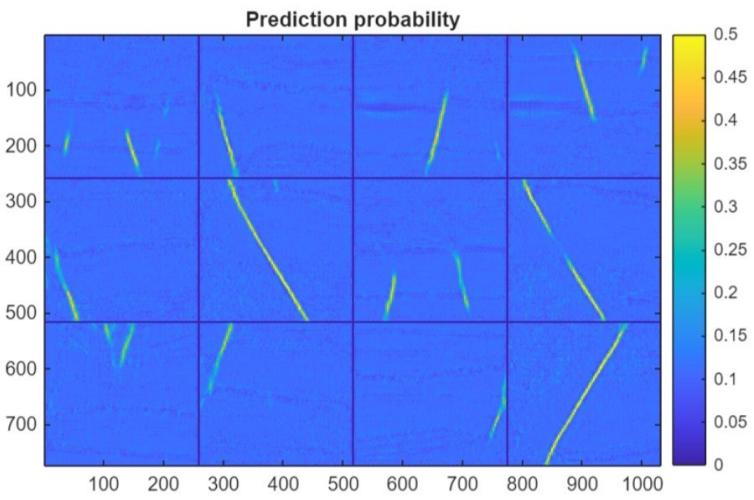
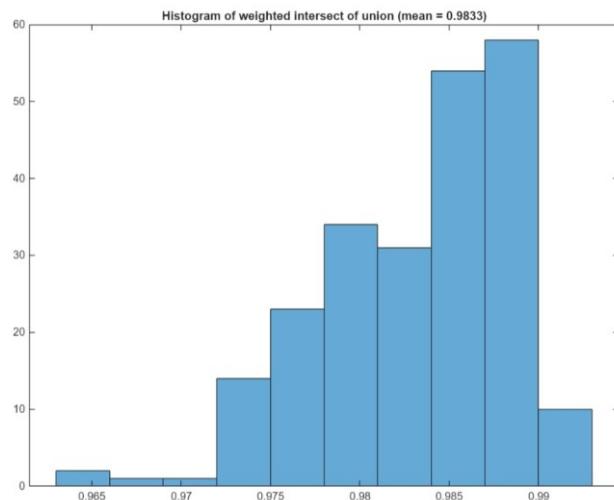
Non-fault Pixels

Fault Pixels



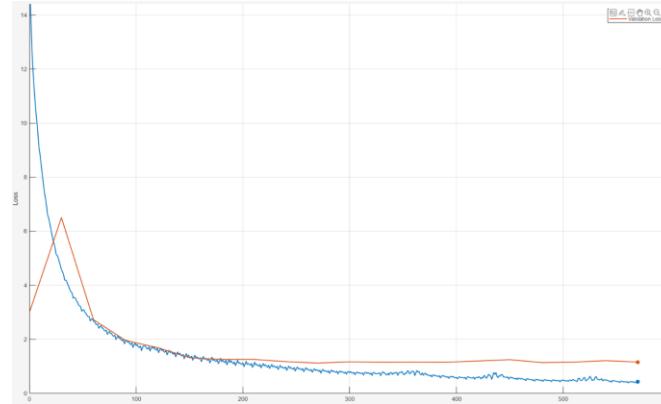
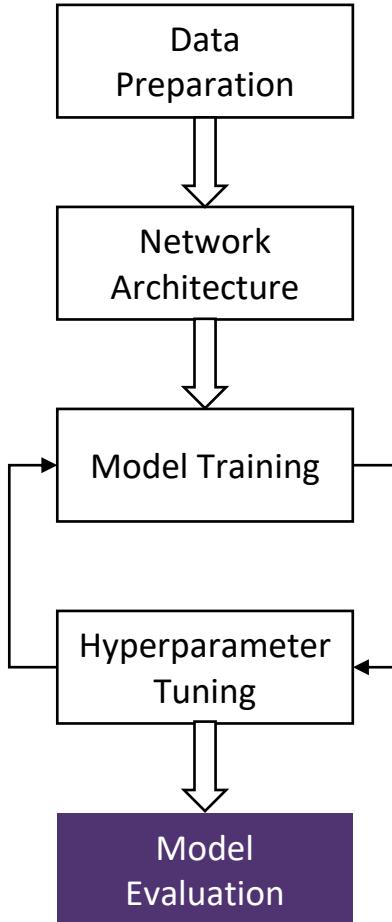
Binary cross function

$$L_{cls} = - \sum_{i=1}^N (y_i \log(p_i) (1-p_i)^\gamma \alpha_i + (1-y_i) \log(1-p_i) p_i^\gamma (1-\alpha_i))$$



# Model Evaluation

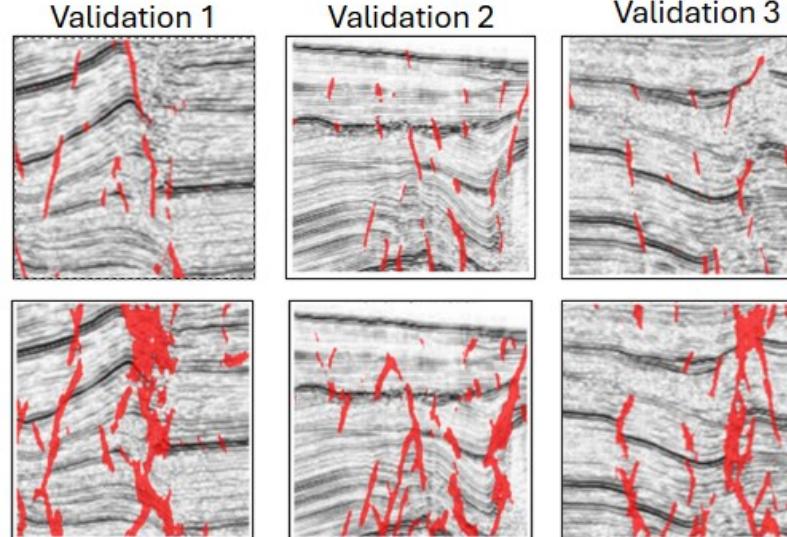
- Quantitative evaluation



observation for overfitting or underfitting

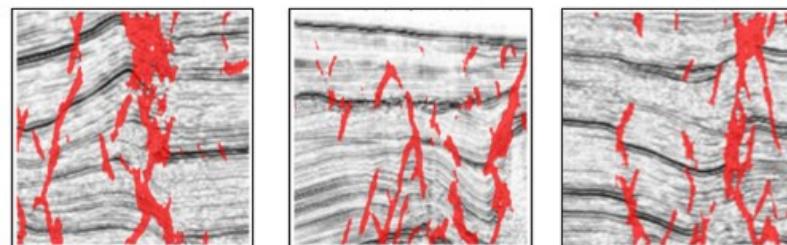
- Qualitative evaluation

Model A



Model B

visual inspection



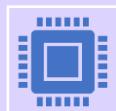
# Discussion



The model might overly predict any **discontinuities** in low quality seismic **image resolution**, such as the bottom of seismic lines.



Poor Signal Quality at Depth: The **signal-to-noise ratio decreased** in deeper parts of the seismic section, making fault detection harder.



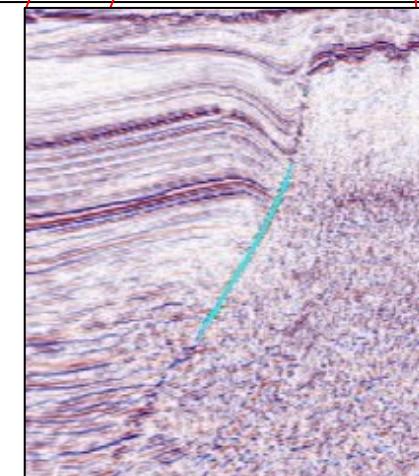
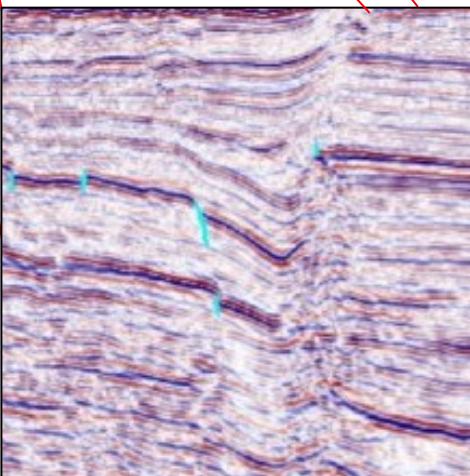
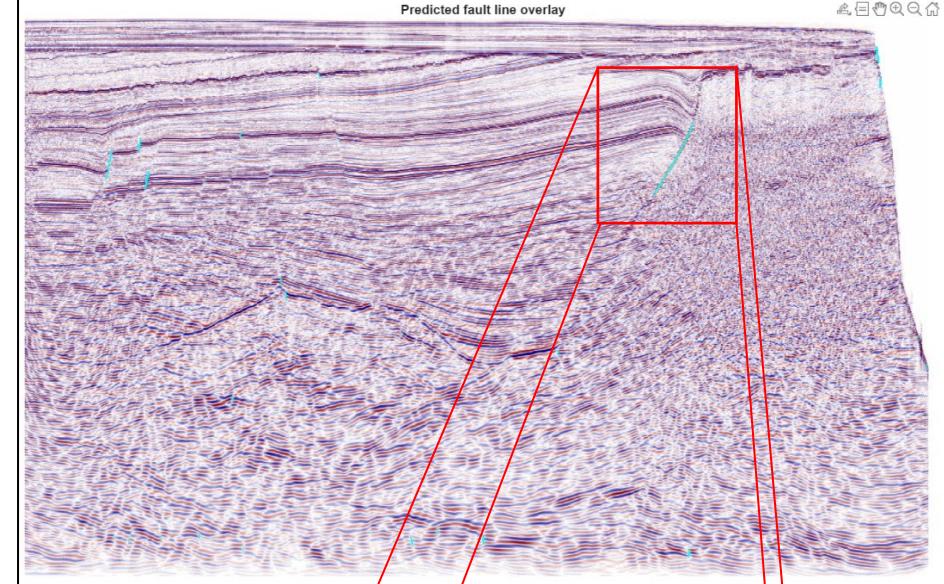
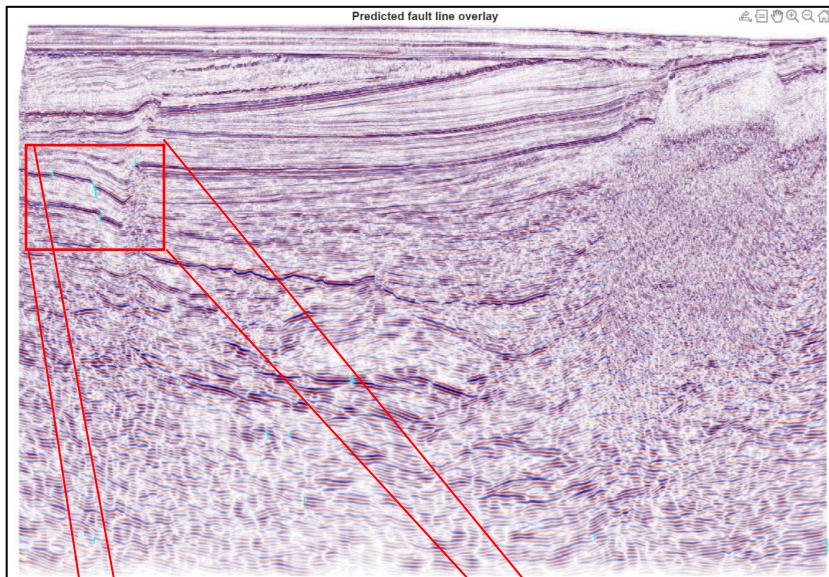
Limited Labeled Data: **Not enough labeled data** reduced the accuracy and performance of the model predictions.



Low Image Resolution: Training on .png images with around 4000 pixels limited the model's ability to produce clear and detailed results.

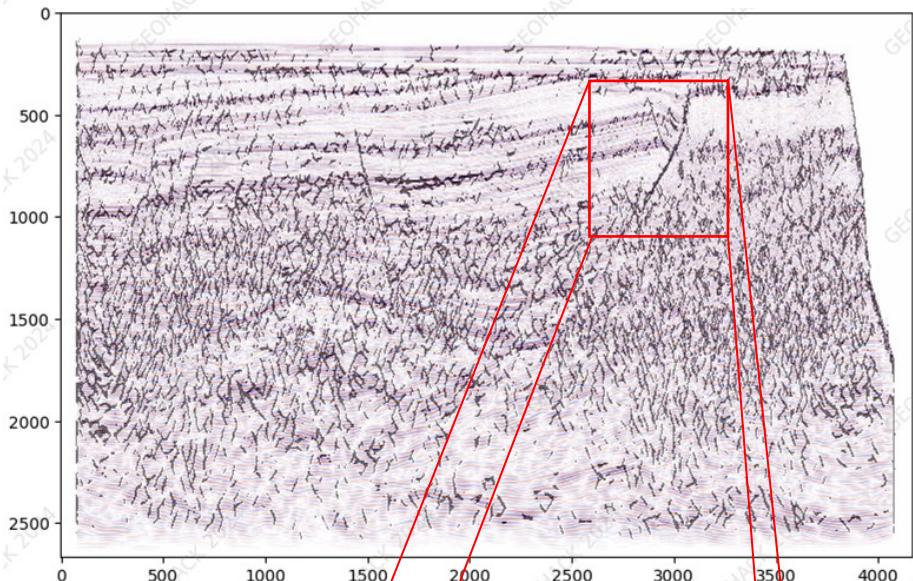
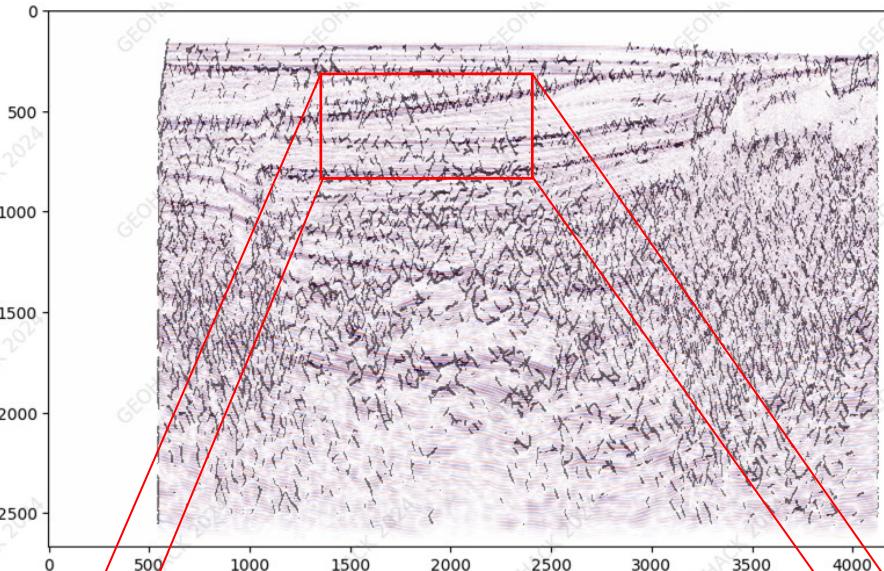
# Visualization of results on holdout data

- Model A predict new data

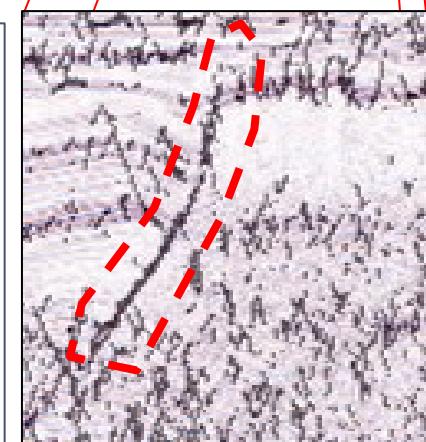
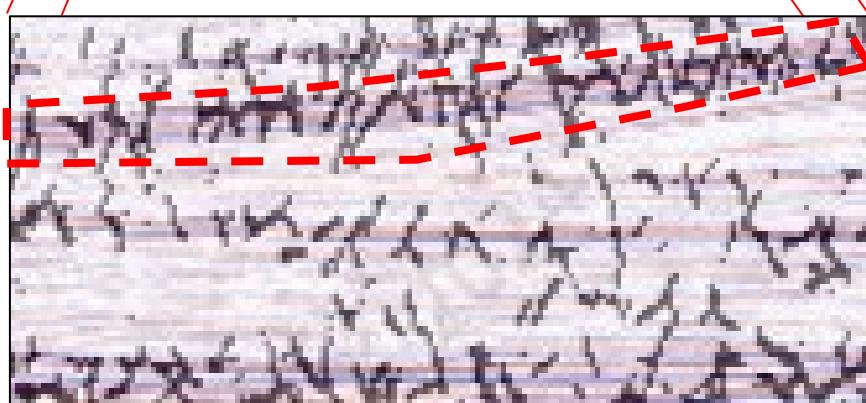
 MATLAB®

# Visualization of results on holdout data

- Model B predict new data



The model tends to predict 'fault' to stronger image resolution/bright amplitude range and any distorted and distorted area which resulted from covariance shift



# Summary and future recommendation

## Summary

1. AI (Deep Learning), especially **CNN** model **superior efficacy** over traditional manual methods which give quicker '*dirty interpretation*'.
2. The model excels particularly in processing seismic image, where it can **detect even minor faults** that manual interpretation may **overlook and save times**.

## Future recommendation

1. Have **more labelled training dataset**, incorporate synthetic and real dataset. Data augmentation?
2. Incorporate **additional pre-processing steps** to handle noise and enhance model performance.
3. Test the model on **different geological settings** to evaluate its generalizability.

[Open]

# Thank you!

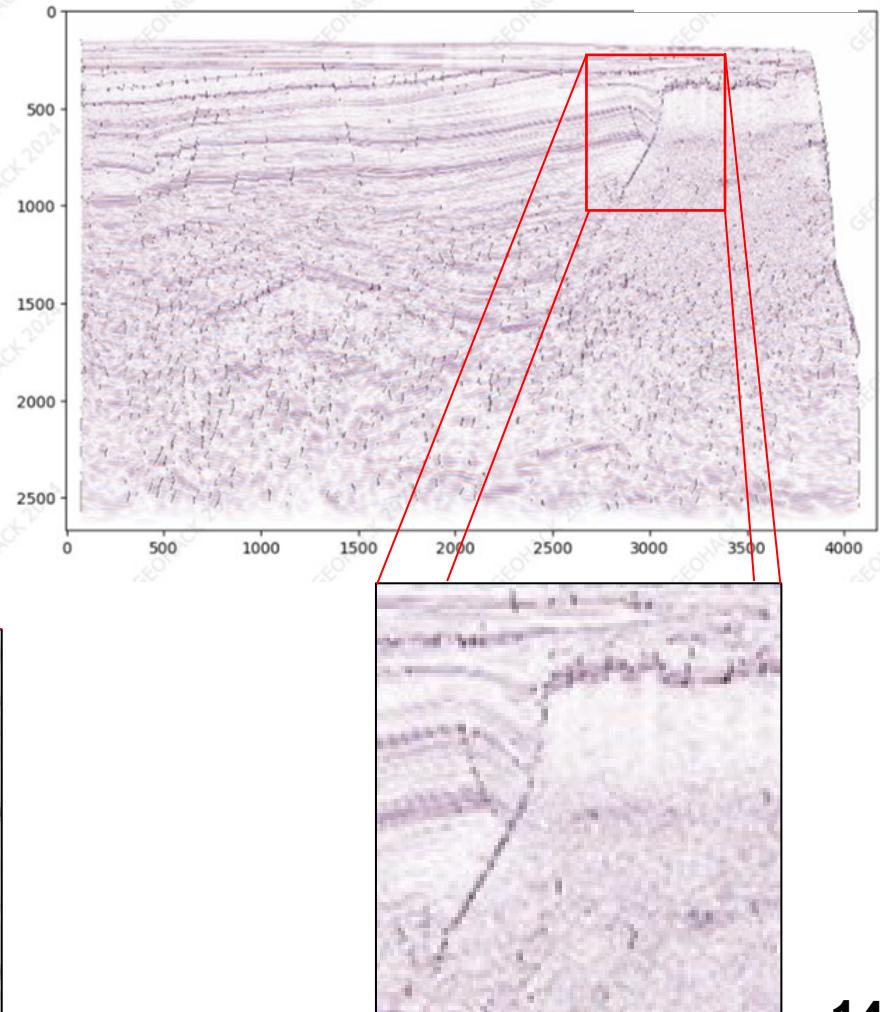
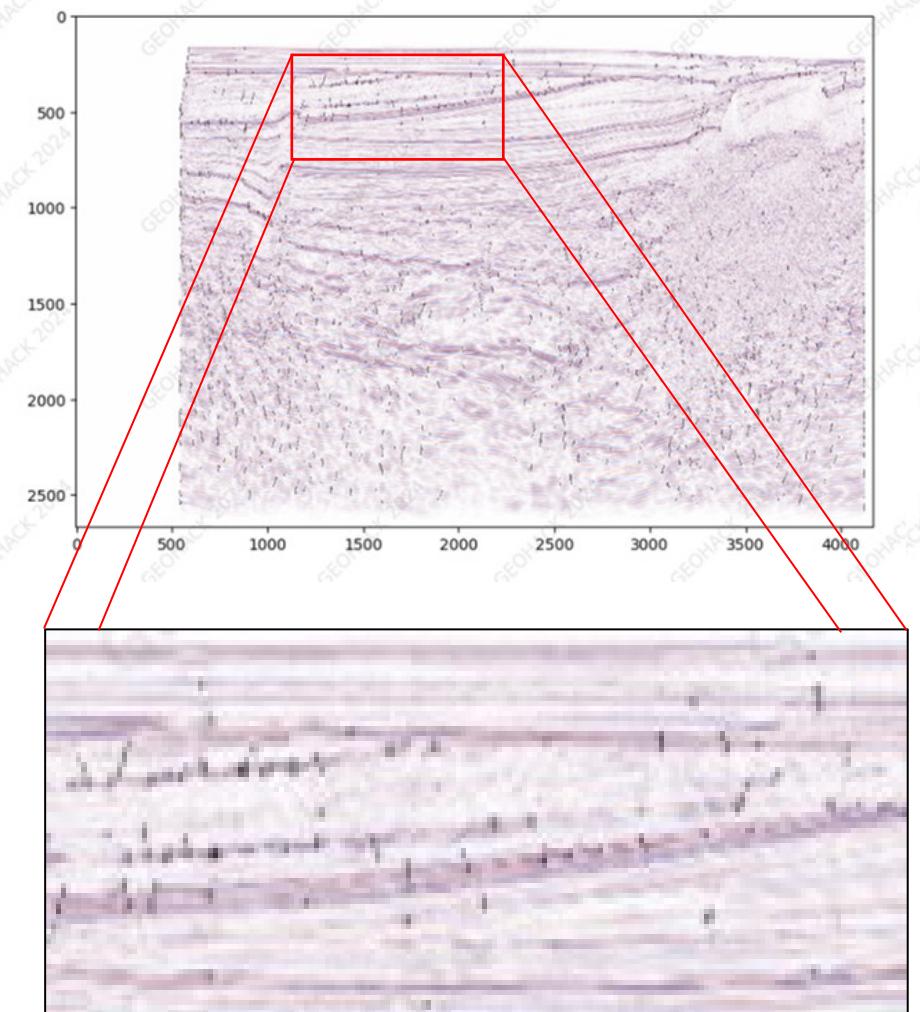
by  
the FAULT in Our STARS





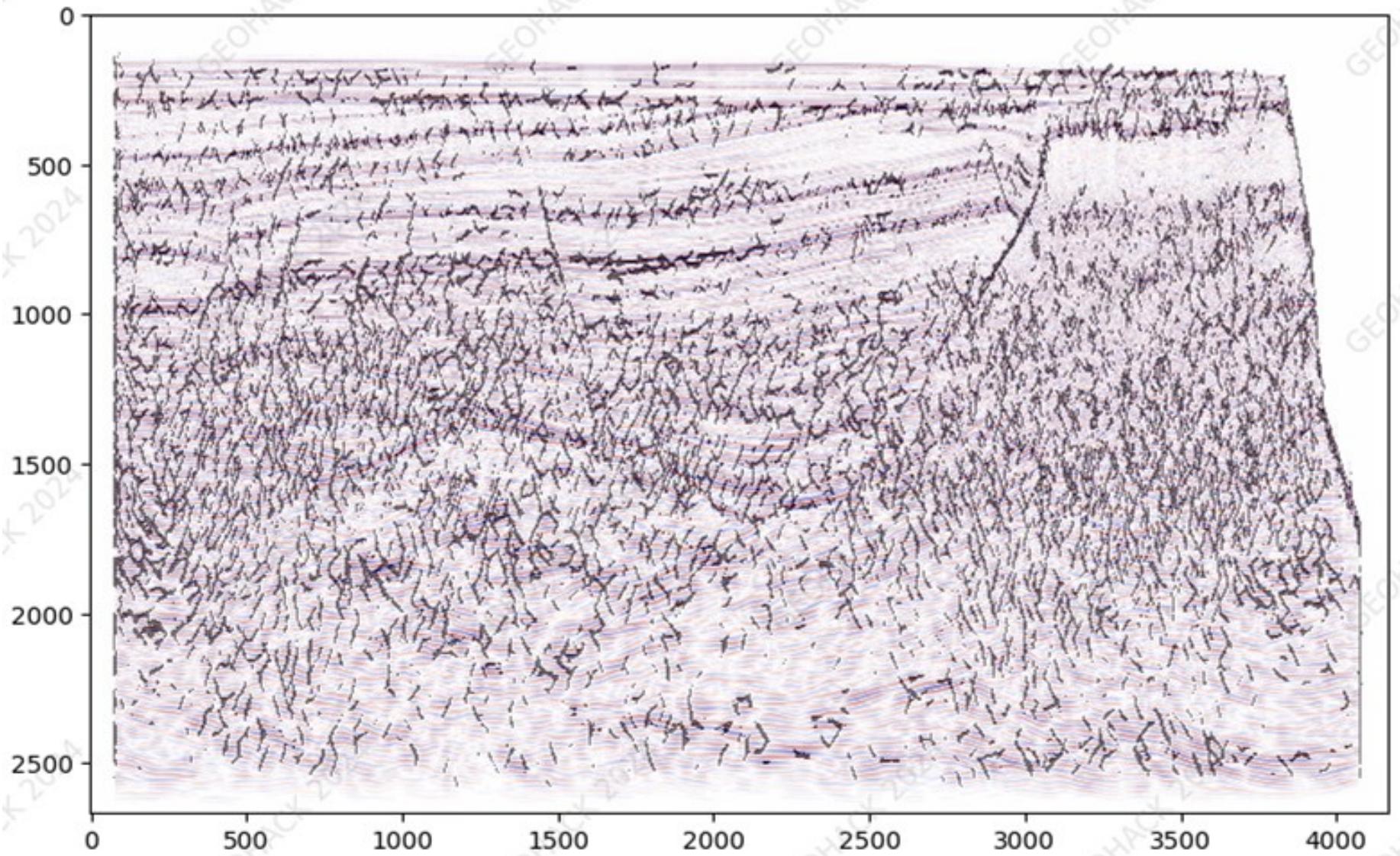
# Visualization of results on holdout data

- Model C predict new data



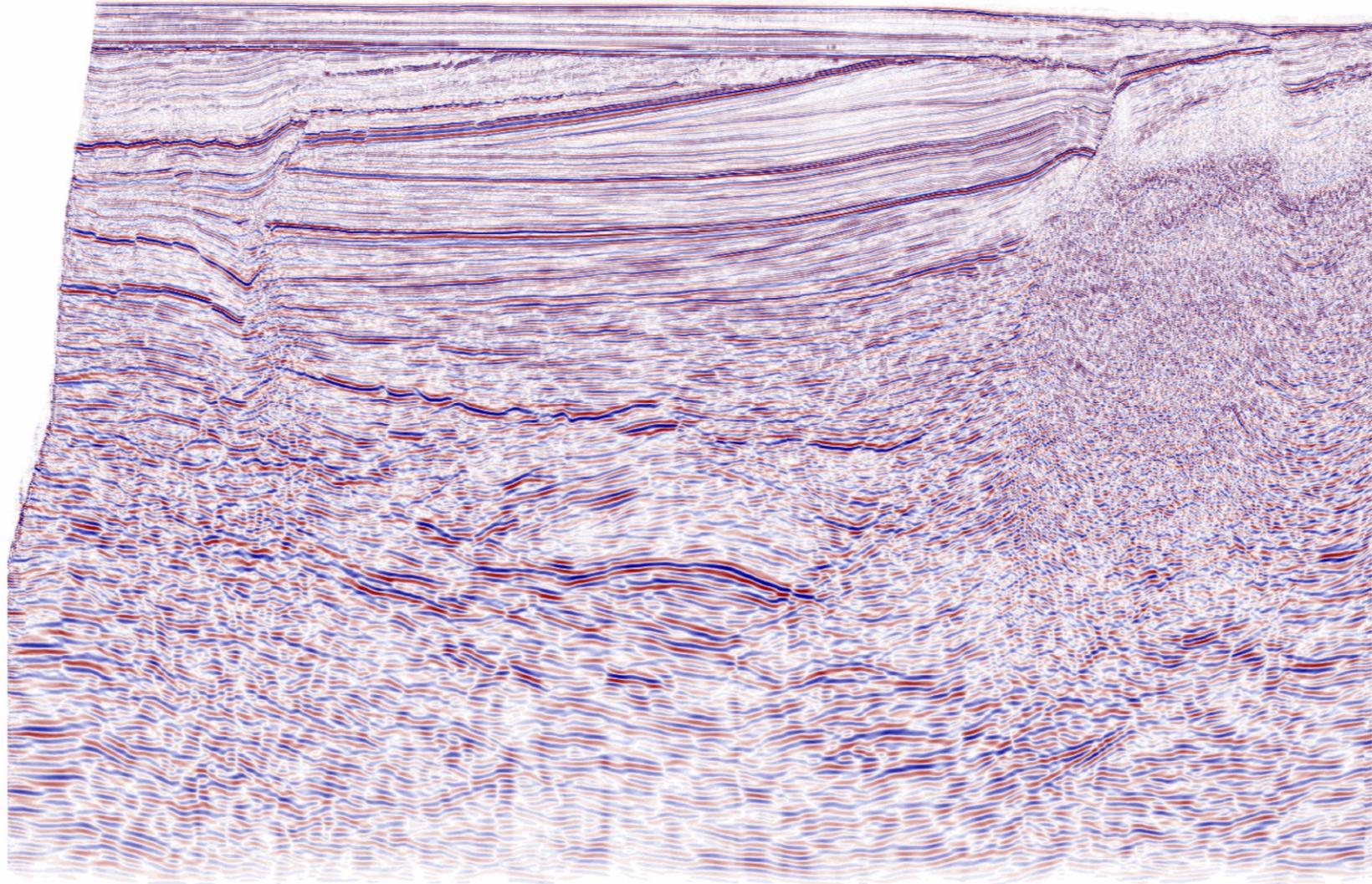


# Backup Slide





# Backup Slide



# Backup Slide

IoU Visualization: TP (Green), FP (Red), FN (Blue)

