UNIVERSITY OF WOLVERHAMPTON

PROPOSED TOPIC: AUTOMATED FAULT DETECTION IN SEISMIC DATA USING DEEP MACHINE LEARNING FOR OIL AND GAS EXPLORATION

NAME: MUHAMMAD BELLO ADAMU
STUDENT NUMBER: 2408698

Table of Contents

PROPOSED TOPIC: AUTOMATED FAULT DETECTION IN SEISMIC DATA USING DEEP MACHINE LEARNING FOR OIL AND GAS EXPLORATION

1.0 INTRODUCTION

1.1 DEFINITION OF KEYWORDS

1.1.1 Seismic Data:

Seismic surveying is undeniably the most crucial of all geophysical exploration methods due to its ability to detect a wide range of subsurface features, from large-scale to small-scale. In essence, seismic methods involve analyzing the shapes and physical properties of Earth's subsurface layers by studying the sound waves that travel through the Earth and return data (Mondol and Bjørlykke, 2010).

1.1.2 Fault:

A fault is a significant fracture in the Earth's crust resulting from the movement of one part of the crust against another. This phenomenon signifies the Earth's dynamic nature, serving as evidence of the powerful forces at work deep underground (Ibrahim, 2013).

1.1.3 Deep Machine Learning:

Machine learning refers to the ability of systems to learn from specific training data in order to automate the process of creating analytical models and solving relevant tasks. Deep learning, on the other hand, is a machine learning concept that is based on artificial neural networks (Janiesch, Zschech and Heinrich, 2021).

1.2 RESEARCH QUESTION:

What are the effective applications of deep learning methods in automated fault detection within seismic data, and what impact do these techniques have on enhancing exploration approaches within the oil and gas sector?

1.3 EXPLANATION:

The focus of this research is on utilizing deep learning techniques to automate the detection of faults in seismic data, specifically examining its impact on exploration strategies in the oil and gas industry. This study aims to explore the feasibility, effectiveness, and real-world applications of advanced machine learning algorithms for improving fault detection in seismic interpretation. Furthermore, the research emphasizes the significance of its findings for industry stakeholders and exploration activities in the UK Continental Shelf (UKCS).

1.4 JUSTIFICATION

Owing to the multitude of intricate fault structures present in seismic images, manual seismic interpretation is a laborious and time-intensive process (Islam et al., 2023).

In the last ten years, deep learning frameworks like convolutional neural networks (CNNs) have significantly advanced research in the upstream oil and gas industry. These frameworks have been applied in various areas including seismic processing/imaging, velocity model building, petrophysics, geological seismic interpretation, as well as in development, production, and supply chain logistics (Chen et al., 2022).

It is important to consider that faults detected from coherence may be impacted by other discontinuities such as noise and stratigraphic features, which are not necessarily related to faults

(Yuan et al., 2019). Therefore, to enhance the precision and effectiveness of coherence, it is necessary to implement preprocessing or postprocessing techniques.

1.5 RELEVANCE OF FAULTS IN OIL AND GAS EXPLORATION

Geologists focus their exploration efforts on areas where geological faults or folds have formed traps capable of retaining hydrocarbons during oil and gas exploration.

1.6 SIGNIFICANCE:

This research question delves into a crucial issue in seismic interpretation and exploration geophysics, providing valuable insights into the use of advanced machine learning techniques to enhance fault detection processes. The implications of this study are significant, as it has the potential to improve exploration efficiency, minimize exploration risks, and optimize the identification of hydrocarbon reserves in both the UK's offshore and onshore areas.

By centering our research around this specific question, we can undertake a thorough examination within the scope of this master's dissertation, while also setting the stage for possible expansions and advancements at the Ph.D. level.

1.7 AIM:

Investigating the feasibility and effectiveness of utilizing deep learning techniques for automated fault detection in seismic data and assessing their implications for improving exploration strategies in the oil and gas industry.

1.8 OBJECTIVES:

1.8.1 Review Literature and Methodologies:

Conduct a comprehensive review of existing literature on seismic interpretation, fault detection methods, and deep learning techniques.

Identify relevant methodologies and frameworks for automated fault detection using deep learning in the context of petroleum exploration.

1.8.2 Data Acquisition and Preprocessing:

Obtain seismic datasets from publicly available repositories, industry collaborations, or synthetic data generation tools.

Preprocess the seismic data to remove noise, standardize formats, and prepare it for model training and validation.

1.8.3 Model Development and Optimization:

Design and implement deep learning architectures tailored for automated fault detection in seismic images.

Experiment with network architectures, hyperparameters, and optimization algorithms to optimize model performance.

1.8.4 Training and Evaluation:

Train the developed models using labeled seismic data samples, emphasizing balanced datasets representing diverse geological settings.

Evaluate model performance using metrics such as precision, recall, F1 score, and receiver operating characteristic (ROC) curves.

1.8.5 Integration with Geophysical Constraints:

Incorporate geological constraints and domain knowledge into the deep learning framework to improve model interpretability and reduce false positives.

Explore methods for integrating structural attributes, stratigraphic information, and seismic attributes into the fault detection process.

1.8.6 Validation and Comparison:

Validate the automated fault detection system using independent test datasets and compare its performance against manual interpretations by domain experts.

Assess the efficiency gains and potential cost savings achieved through automation compared to traditional fault detection methods.

1.8.7 Implications for Exploration Strategies:

Analyze the implications of automated fault detection on exploration strategies within the UK's oil and gas industry, including its impact on decision-making, risk assessment, and resource optimization.

Discuss potential challenges, limitations, and opportunities for further research and implementation in real-world exploration projects.

By delineating clear aims and objectives, we can structure this research activity and focus this effort on achieving the desired outcomes within the constraints of this master's dissertation. These objectives provide a roadmap for conducting a systematic investigation into automated fault detection using deep learning techniques while addressing the broader implications for petroleum exploration in the UK.

1.9 KEY COMPONENTS

Utilization of Deep Learning Techniques: The research focuses on the application of deep learning methods, such as convolutional neural networks (CNNs), recurrent neural networks (RNNs), or other relevant architectures, for automated fault detection in seismic data.

Effectiveness and Feasibility: Exploring the efficacy and feasibility of employing deep learning techniques for fault detection involves considering various factors such as data quality, model performance, computational resources, and interpretability.

Implications for Exploration Strategies: Investigate the potential impact of automated fault detection on exploration strategies within the UK's oil and gas industry, including implications for decision-making, risk assessment, and resource optimization.

2.0 LITERATURE REVIEW

2.1 Example Forgetting: A Novel Approach to Explain and Interpret Deep Neural Networks in Seismic Interpretation (Benkert, Aribido and AlRegib, 2022)

JOURNAL: IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING

The use of deep neural networks has had a significant impact on the process of seismic interpretation. Their simple implementation and low interpretation costs make them an appealing component for the common interpretation pipeline. However, neural networks are often met with skepticism due to their tendency to produce semantically incorrect outputs when exposed to sections of the model that were not included in the training data.

This paper aims to address this issue by explaining the behavior of the model and improving its generalization properties through example forgetting. First, a method was introduced to effectively relate semantically malfunctioned predictions to their positions within the neural network representation manifold. This method tracks how models "forget" seismic reflections during training and establishes a connection to the decision boundary proximity of the target class.

Second, an analysis technique was used to identify frequently forgotten regions within the training volume and augment the training set with state-of-the-art style transfer techniques from computer vision. This approach demonstrated an improvement in segmentation performance on underrepresented classes while significantly reducing the forgotten regions in the F3 volume in the Netherlands.

2.2 Seismic fault detection using an encoder–decoder convolutional neural network with a small training set (Li et al., 2019)

JOURNAL: JOURNAL OF GEOPHYSICS AND ENGINEERING

In the field of seismic interpretation, detecting faults is a critical step that often requires a significant amount of manual labor and time. Convolutional neural networks (CNN) are state-of-the-art deep learning technology capable of outperforming humans in image recognition. However, traditional methods of using CNNs for prediction typically require a large dataset to train the network, which is impractical for many researchers and interpreters in geophysics who struggle to obtain sufficient quantities of labeled real data.

This paper introduces a method for seismic fault detection using a CNN that requires only a very small training set. The fault detection process is treated as a semantic segmentation task, training an encoder–decoder CNN, specifically a U-Net, to make pixel-by-pixel predictions on the seismic section to determine whether each pixel represents a fault or non-fault.

The experiments using this type of CNN yielded good prediction results on real data. With the proposed method, interpreters only need to pick and label a few 2D sections when interpreting a new seismic volume, after which the model can predict faults in any other section of the same volume, significantly improving interpretation efficiency.

To evaluate the performance of the proposed method, the paper introduces a fault detection accuracy index that describes the accuracy of the prediction results.

This paper demonstrates that using only seven seismic sections of a seismic volume to train a CNN can successfully predict faults in any other section of the same volume.

2.3 <u>Transformer assisted dual U-net for seismic fault detection</u> (Wang et al., 2023)

JOURNAL: FRONTIERS IN EARTH SCIENCE

The identification of seismic faults in seismic data is crucial for oil and gas exploration. Traditional manual methods are inadequate for processing large amounts of seismic data. The advancement of artificial intelligence technology has led to the popularity of deep learning techniques based on pattern recognition for identifying seismic faults.

While U-shaped neural networks (U-Net) have made progress, they still do not fully meet the requirements for fault prediction in complex structures. As a result, a new approach called Dual U-Net with Transformer, which combines a standard U-Net with a transformer U-Net to create a parallel dual U-Net model, has been proposed.

In order to improve the accuracy of fault prediction, six loss functions were compared using synthetic data. The binary cross entropy loss function was found to be the most robust based on three evolution metrics, including the Dice coefficient, Sensitivity, and Specificity.

An example comparing the prediction performance of different U-Net models on synthetic data has shown the superior performance of the Dual U-Net model, demonstrating its practical application value.

To further validate the practical feasibility of this method, real seismic data with a complex fault system was used. It was found that the proposed model is more accurate in predicting the fault system compared to well-developed U-Net models such as the classical U-Net and classical coherence cube algorithm, without transfer learning. This confirms the potential for widespread application of this proposed model.

2.4 <u>Application of a Pre-Trained CNN Model for Fault Interpretation in the Structurally Complex Browse Basin, Australia</u> (Islam et al., 2023)

JOURNAL: APPLIED SCIENCES

Detecting faults is a crucial part of analyzing subsurface data and characterizing reservoirs from 3D seismic images. The manual interpretation of seismic images can be time-consuming and labor-intensive due to the presence of numerous and intricate fault structures.

In a study focused on the Poseidon field in the Browse Basin, Australia, a pre-trained CNN model was utilized to forecast faults from the 3D seismic volume. This field exhibits a high level of structural complexity, particularly with regards to normal faulting in the targeted Plover Formations.

The goal of this research is to compare the fault prediction accuracy of machine learning algorithms with that of user-interpreted fault identification using seismic variance attributes.

The results of the study were reasonably satisfactory, with the CNN model generating an improved fault probability volume that outperformed the variance technology. As a result, it is suggested that this workflow could significantly reduce the time needed for fault prediction, particularly in areas with complex structural features.

2.5 <u>Augmented deep learning workflow for robust fault prediction over multiple tectonic regimes</u> (Chen et al., 2022)

JOURNAL: JOURNAL OF GEOPHYSICS AND ENGINEERING

In the last ten years, deep learning frameworks like convolutional neural networks (CNNs) have made significant progress in the upstream oil and gas industry. These advancements have been applied in various areas such as seismic processing, velocity model building, petrophysics, and geological seismic interpretation, as well as in development, production, and supply chain logistics.

CNN fault prediction focuses on image edge detection. To achieve better prediction results, three data-driven steps are recommended. Firstly, the seismic data should be pre-conditioned to maximize the signal-to-noise ratio. This involves using iterative dip-steered median filtering and principal component filtering to improve the sharpness and signal-to-noise ratio for edge detection. Secondly, a fault probability volume obtained through deep learning-based fault detection using a U-Net architecture is suggested. This approach uses synthetic seismic models as samples to enhance fault prediction, as generating labeled fault sets can be time-consuming. Finally, edge enhancement is performed on the inference results to improve precision and fault continuity. The study also includes a comparative analysis of related edge enhancement technologies.

The results show the proposed workflow's effectiveness through a comparative analysis of three different faulting modes (normal, reverse, and strike-slip) over three real seismic field datasets from China.

3.0 METHODOLOGY

3.1 THE MODEL

The U-Net model, a type of image segmentation model, will be employed for the identification of geological faults from seismic image patches. The following steps will be implemented for this purpose:

3.1.1 Preprocessing the seismic images:

Prior to inputting the seismic images into the U-Net model for fault identification, it is essential to preprocess the data to improve its quality and suitability. This may include tasks such as reducing noise, enhancing contrast, normalizing, and resizing to ensure that the data is consistent and compatible with the model's input requirements.

3.1.2 Loading the data and preparation:

The seismic image data is loaded into the computational environment, along with corresponding labels or ground truth annotations indicating the location of geological faults. The data is carefully prepared for training by organizing it into batches, splitting it into training and validation sets, and applying any necessary transformations or augmentations to increase dataset diversity and improve model generalization.

3.1.3 Training the U-Net:

The U-Net model undergoes training using meticulously prepared seismic image data in order to effectively recognize geological faults with precision. Throughout the training process, the model continuously fine-tunes its internal parameters by analyzing the input data and corresponding ground truth labels to minimize any disparities between predicted and actual fault locations. This iterative process entails inputting batches of seismic image patches into the model, evaluating loss functions to measure prediction errors, and adjusting model weights through backpropagation and gradient descent optimization.

3.1.4 Saving the trained model:

Upon successful completion of the training process and achieving satisfactory performance on the validation set, the U-Net model is saved to disk for future use. This enables easy deployment and inference on new seismic data without the necessity of retraining from the beginning.

3.1.5 Predicting and identifying the geological faults:

Utilizing the U-Net model, we can make predictions on seismic image patches that have not been previously seen to detect geological faults. The model takes raw seismic data patches as input and produces predictions or probability maps that show the probability of a fault being present at each pixel or voxel. These predictions are then adjusted and post-processed to generate binary fault maps that pinpoint the locations of detected faults.

3.1.6 Visualization:

Upon completion, the fault identification results are visualized and interpreted to evaluate the U-Net model's performance and to gain a better understanding of the geological structures within the seismic data. Visualization methods may involve overlaying predicted fault maps onto original seismic images, creating 3D reconstructions of fault networks, and comparing model predictions with ground truth annotations to conduct qualitative analysis and validation.

3.2 POSSIBLE SOURCES OF DATA:

3.2.1 Public Repositories and Open Data Platforms:

It is important to prioritize accessing seismic datasets from public repositories and open data platforms. Seek out datasets offered by government agencies, academic institutions, and international organizations that are freely available for academic research purposes. Platforms such as the UK Oil and Gas Authority's Open Data Hub, the British Geological Survey's GeoIndex, and international repositories like SEG Wiki and Open Seismic Repository are great places to start.

3.2.2 Academic Collaborations and Research Networks:

Consider seeking collaborations with academic researchers and institutions that possess relevant seismic datasets for this research. Contact professors, researchers, or graduate students in the fields of geophysics, petroleum engineering, or related areas to inquire about the possibility of accessing or sharing their data for collaborative research projects.

3.2.3 Synthetic Data Generation:

Generating synthetic seismic datasets using open-source software and modeling tools can be a valuable resource for algorithm development, testing, and validation, especially when access to real seismic data is limited. Tools such as OpendTect and SEISGEN are capable of producing synthetic seismic datasets based on geological models and parameters.

3.2.4 Open Challenges and Competitions:

Consider participating in open challenges or competitions that center around seismic interpretation and machine learning. These events typically grant access to benchmark datasets, code repositories, and online forums where participants can collaborate and share insights. Platforms such as Kaggle, SEG Advanced Modeling, and EAGE Student Challenges offer valuable opportunities to access datasets and engage with the research community.

3.2.5 Literature Review and Data Requests:

It is important to conduct a thorough literature review to identify relevant studies, datasets, and authors in this research area. Consider reaching out to the authors of published papers or researchers in the field to inquire about the availability of their seismic datasets for academic purposes. Many researchers are open to sharing their data upon request, particularly if it can contribute to further research and collaboration.

3.3 THE DURATION OF THIS RESEARCH PROJECT (SOLE RESEARCHER)

3.3.1 Month 1 - Week 1-2: Data Acquisition and Preprocessing:

Allocating the initial two weeks to procure seismic image data from pertinent sources, such as public repositories or industry partnerships. Following this, the obtained data will be preprocessed, encompassing tasks such as noise reduction, contrast enhancement, and normalization.

3.3.2 Month 1 - Week 3-4: Loading the Data and Preparation:

Loading the preprocessed seismic data into the computational environment. Then, Preparing the data for training by organizing it into batches, splitting it into training and validation sets, and applying transformations or augmentations.

3.3.3 Month 2 - Week 1-3: Training the U-Net:

Begin training the U-Net model using the prepared seismic image data and monitoring the training process, adjust hyperparameters if necessary, and evaluate model performance on the validation set. Also, Iteratively refine the model architecture and training strategy to improve performance.

3.3.4 Month 2 - Week 4: Saving the Trained Model:

Once the U-Net model has converged and achieved satisfactory performance, saving the trained model to disk for future use is next. Ensuring that all necessary model files and metadata are properly stored and documented.

3.3.5 Month 3 - Week 1-2: Predicting and Identifying the Geological Faults:

Using the saved trained model to make predictions on unseen or test seismic image patches. Processing the model predictions to generate binary fault maps highlighting detected fault locations.

3.3.6 Month 3 - Week 3: Visualization:

Visualizing the results of fault identification to assess model performance and gain insights into geological structures. Generate visualization outputs, such as overlaying fault maps onto original seismic images and creating 3D reconstructions of fault networks.

3.3.7 Month 3 - Week 4: Finalize Dissertation and Report Writing:

Dedicating the final week to finalizing the dissertation, including writing the introductory and methodological chapters, summarizing results, and discussing implications. Also, reviewing and editting the dissertation for clarity, coherence, and accuracy, ensuring that all sections are well-organized and meet academic standards.

4.0 References

Benkert, R., Aribido, O. J. and AlRegib, G. (2022) Example Forgetting: A Novel Approach to Explain and Interpret Deep Neural Networks in Seismic Interpretation, *IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING*, Piscataway, Ieee-Inst Electrical Electronics Engineers Inc, **60**, p. 5916812, [online] Available from: https://www.webofscience.com/api/gateway?GWVersion=2&SrcAuth=DOISource&SrcApp=WOS&KeyAID=10.1109%2Ftgrs.2022.3178112&DestApp=DOI&SrcAppSID=EUW1ED0D6Cufv6DfVOrGSgcCcSjI5&SrcJTitle=IEEE+TRANSACTIONS+ON+GEOSCIENCE+AND+REMOTE+SENSING&DestDOIRegistrantName=Institute+of+Electrical+and+Electronics+Engineers (Accessed 14 April 2024).

Chen, H., Tang, D., Liu, Y., Gu, H. and Zhou, C. (2022) Augmented deep learning workflow for robust fault prediction over multiple tectonic regimes, *JOURNAL OF GEOPHYSICS AND ENGINEERING*, Oxford, Oxford Univ Press, **19**(4), pp. 914–926, [online] Available from: https://www.webofscience.com/api/gateway?GWVersion=2&SrcAuth=DOISource&SrcApp=WOS&KeyAID=10.1093%2Fjge%2Fgxac056&DestApp=DOI&SrcAppSID=EUW1ED0D6Cufv6DfVOrGSgcCcSjI5&SrcJTitle=JOURNAL+OF+GEOPHYSICS+AND+ENGINEERING&DestDOIRegistrantName=Oxford+University+Press (Accessed 14 April 2024).

Ibrahim, S. (2013) Geological faults,.

Islam, M. M., Babikir, I., Elsaadany, M., Elkurdy, S., Siddiqui, N. A. and Akinyemi, O. D. (2023) Application of a Pre-Trained CNN Model for Fault Interpretation in the Structurally Complex Browse Basin, Australia, *APPLIED SCIENCES-BASEL*, Basel, MDPI, **13**(20), p. 11300, [online] Available from: https://www.webofscience.com/api/gateway?GWVersion=2&SrcAuth=DOISource&SrcApp=WOS&KeyAID=10.3390%2Fapp132011300&DestApp=DOI&SrcAppSID=EUW1ED0D6Cufv6DfVOrGSgcCcSjI5&SrcJTitle=APPLIED+SCIENCES-BASEL&DestDOIRegistrantName=MDPI+AG (Accessed 14 April 2024).

Janiesch, C., Zschech, P. and Heinrich, K. (2021) *Machine learning and deep learning*,.

Li, S., Yang, C., Sun, H. and Zhang, H. (2019) Seismic fault detection using an encoder-decoder convolutional neural network with a small training set, *JOURNAL OF GEOPHYSICS AND ENGINEERING*, Oxford, Oxford Univ Press, **16**(1), pp. 175–189, [online] Available from: https://www.webofscience.com/api/gateway?GWVersion=2&SrcAuth=DOISource&SrcApp=WOS&KeyAID=10.1093%2Fjge%2Fgxy015&DestApp=DOI&SrcAppSID=EUW1ED0D6Cufv6DfVOrGSgcCcSjI5&SrcJTitle=JOURNAL+OF+GEOPHYSICS+AND+ENGINEERING&DestDOIRegistrantName=Oxford+University+Press (Accessed 14 April 2024).

Mondol, N. and Bjørlykke, K. (2010) Seismic Exploration, In pp. 375–402.

Wang, Z., You, J., Liu, W. and Wang, X. (2023) Transformer assisted dual U-net for seismic fault detection, *FRONTIERS IN EARTH SCIENCE*, Lausanne, Frontiers Media Sa, **11**, p. 1047626, [online] Available from: https://www.webofscience.com/api/gateway?GWVersion=2&SrcAuth=DynamicDOIArticle&SrcApp=WOS&KeyAID=10.3389%2Ffeart.2023.1047626&DestApp=DOI&SrcAppSID=EUW1ED0D6Cufv6DfVOrGSgcCcSjI5&SrcJTitle=FRONTIERS+IN+EARTH+SCIENCE&DestDOIRegistrantName=Frontiers+Media+SA (Accessed 14 April 2024).

Yuan, Z., Huang, H., Jiang, Y., Tang, J. and Li, J. (2019) An enhanced fault-detection method based on adaptive spectral decomposition and super-resolution deep learning, *INTERPRETATION-A JOURNAL OF SUBSURFACE CHARACTERIZATION*, Tulsa, Soc Exploration Geophysicists - Seg, **7**(3), pp. T713–T725, [online] Available from: https://www.webofscience.com/api/gateway?GWVersion=2&SrcAuth=DOISource&SrcApp=WOS&KeyAID=10.1190%2Fint-2018-0180.1&DestApp=DOI&SrcAppSID=EUW1ED0D6Cufv6DfVOrGSgcCcSjI5&SrcJTitle=INTERPRETATION-A+JOURNAL+OF+SUBSURFACE+CHARACTERIZATION&DestDOIRegistrantName=Society+of+Exploration+Geophysicists (Accessed 14 April 2024).