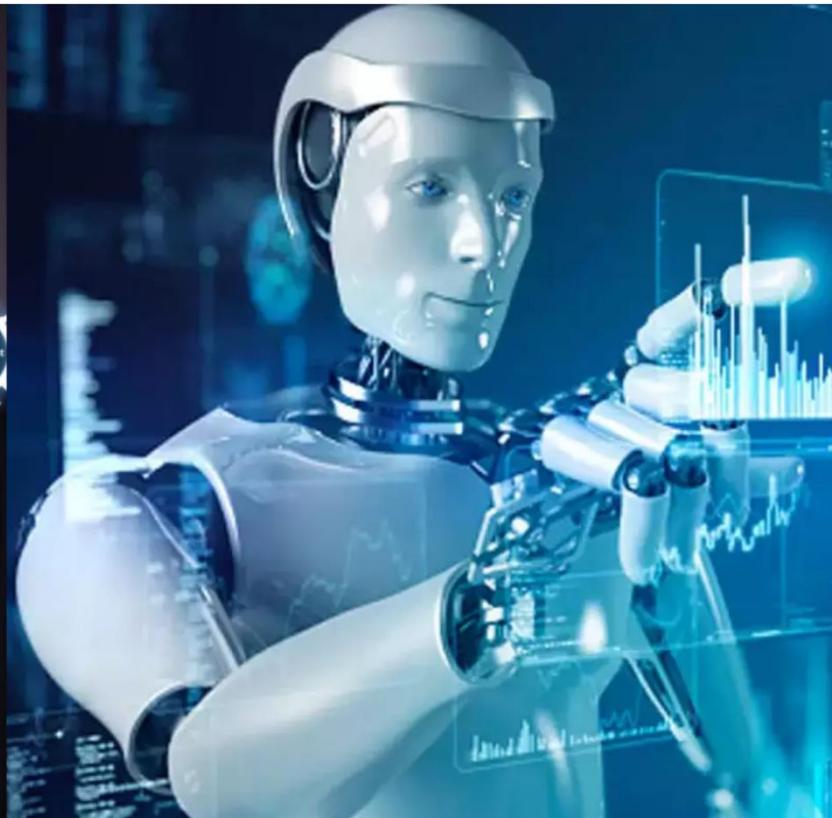




International Journal of Innovative Research in Science Engineering and Technology (IJIRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 8.699

Volume 14, Issue 6, June 2025

Enhanced Salt Body Segmentation in Seismic Data using Attention-Driven UNet Models

Pradnya D. Bormane, Pranav Mohite, Sanjana Patil, Varun Patil, Om Ranasubhe

Professor, Dept. of Artificial Intelligence & Data Science, AISSMS IOIT, Pune, Maharashtra, India

UG Student, Dept. of Artificial Intelligence & Data Science, AISSMS IOIT, Pune, Maharashtra, India

UG Student, Dept. of Artificial Intelligence & Data Science, AISSMS IOIT, Pune, Maharashtra, India

UG Student, Dept. of Artificial Intelligence & Data Science, AISSMS IOIT, Pune, Maharashtra, India

UG Student, Dept. of Artificial Intelligence & Data Science, AISSMS IOIT, Pune, Maharashtra, India

ABSTRACT: Accurate segmentation of salt bodies in seismic images plays a critical role in hydrocarbon exploration and geological analysis. However, the complex shapes and low-contrast boundaries of salt structures pose significant challenges to traditional segmentation techniques. In this study, we propose an enhanced UNet architecture augmented with attention mechanisms for precise salt body segmentation. The integration of self-attention modules and attention gates allows the network to selectively focus on relevant regions of the seismic image, improving segmentation accuracy and robustness. Our experiments, conducted on the TGS Salt Identification Challenge dataset, demonstrate significant performance improvements over baseline models. The proposed method achieves a Dice coefficient of 0.91 and an IoU of 0.87, outperforming the standard UNet and other existing approaches. These results highlight the potential of attention based deep learning architectures for seismic image analysis, paving the way for more efficient subsurface exploration.

KEYWORDS: Salt body segmentation, seismic imaging, UNet, attention mechanisms, deep learning, hydrocarbon exploration.

I. INTRODUCTION

The identification of salt bodies in seismic images is a crucial task in geophysical exploration, playing a pivotal role in the discovery of hydrocarbon reservoirs and geological assessment. Salt structures, often characterized by their complex geometries and low-contrast boundaries, present significant challenges for segmentation algorithms. Traditional methods, such as manual interpretation and classical image processing, are labour-intensive, prone to human error, and inadequate for large-scale seismic data analysis. The emergence of deep learning, particularly convolutional neural networks (CNNs), has revolutionized the field of image segmentation. Among the proposed architectures, UNet [1] has gained prominence for its encoder-decoder structure with skip connections, enabling accurate feature extraction and localization. However, the standard UNet struggles with accurately segmenting salt bodies due to the intricate boundaries and varying scales of these structures.

To address these challenges, researchers have proposed enhancements to the UNet architecture. For instance, Alfarhan et al. [2] demonstrated the efficacy of employing VGG19 and ResNet34 as encoders in a modified UNet for detecting salt domes and faults. Similarly, Saad et al. [3] integrated self-attention mechanisms into DenseNet-based UNet architectures, improving the model's ability to capture long-range dependencies. In another study, Guo et al. [4] employed a two-stage training process with Binary Cross Entropy (BCE) and Lovasz Loss to optimize segmentation accuracy, achieving significant improvements in IoU metrics.

Interactive segmentation methods have also been explored. Zhang et al. [5] proposed SaltISCG, an approach combining CNNs and graph cuts, leveraging user input for refined salt segmentation. Diakogiannis et al. [6] introduced ResUNet-a, a framework with attention gates, achieving state-of-the-art segmentation performance in remote sensing applications. Building on these advancements, this study proposes an enhanced UNet architecture incorporating self-attention modules in the encoder and attention gates in the decoder. The self-attention modules enable the model to

capture global dependencies within seismic images, while attention gates refine the skip connections by suppressing irrelevant features.

Our contributions are threefold:

- We design a novel attention-enhanced UNet architecture tailored for salt body segmentation in seismic images.
- We evaluate the proposed model on the TGS Salt Identification Challenge dataset, demonstrating superior performance with a Dice coefficient of 0.91 and IoU of 0.87.
- We provide a comprehensive comparison with existing methods, highlighting the impact of attention mechanisms on segmentation accuracy and robustness.

The remainder of this paper is structured as follows. Section 2 reviews related work in salt body segmentation and attention-based architectures. Section 4 details the proposed methodology and model design. Section 5 presents experimental results and analysis. Finally, Section 6 concludes the paper and discusses future research directions.

II. RELATED WORK

Salt body segmentation is a critical task in seismic image analysis, enabling accurate identification of subsurface structures for hydrocarbon exploration and agricultural assessment. Several deep learning techniques have been proposed to address the challenges associated with segmenting complex salt geometries.

Lakshmi et al. [7] proposed two models for salt segmentation. The first employed UNet combined with ResNet-18 and ResNet-34 as encoders, while the second utilized an ensemble approach with UNet combined with ResNet-34, VGG16, and InceptionV3. The ensemble model achieved superior results, demonstrating the effectiveness of combining multiple pre-trained backbones for segmentation.

Other researchers have explored enhancing UNet by integrating additional architectural improvements. Alfarhan et al. [2] improved UNet by incorporating VGG19 and ResNet34 encoders, achieving better detection rates for salt domes and faults in seismic surveys. Similarly, Saad et al. [3] introduced self-attention mechanisms within a DenseNet-based UNet architecture, demonstrating significant improvements in segmentation accuracy.

Islam et al. [8] developed a UNet variant employing SeResNet backbones and optimized it using Lovasz Loss and Binary Cross-Entropy, achieving robust salt region detection. Another approach by Guo et al. [4] proposed a two-stage training process using BCE and Lovasz losses sequentially, effectively improving IoU metrics for salt body segmentation.

In interactive segmentation, Zhang et al. [5] proposed a CNN-Graph Cut hybrid method for user guided salt segmentation, leveraging human input to enhance model precision. Diakogiannis et al. [6] introduced ResUNet-a, a framework utilizing attention gates for semantic segmentation of remotely sensed data, showcasing applicability in diverse domains beyond seismic imaging.

These studies collectively emphasize the role of architectural innovations, attention mechanisms, and ensemble methods in advancing salt body segmentation. However, there remains significant scope for improving segmentation accuracy and robustness, particularly through the integration of self-attention and attention gates in UNet-based architectures.

III. PROPOSED SOLUTION

Salt body segmentation in seismic images requires a robust model capable of capturing complex boundaries and long-range dependencies. To address these challenges, we propose an enhanced UNet architecture that integrates attention mechanisms. The key components of the proposed architecture include:

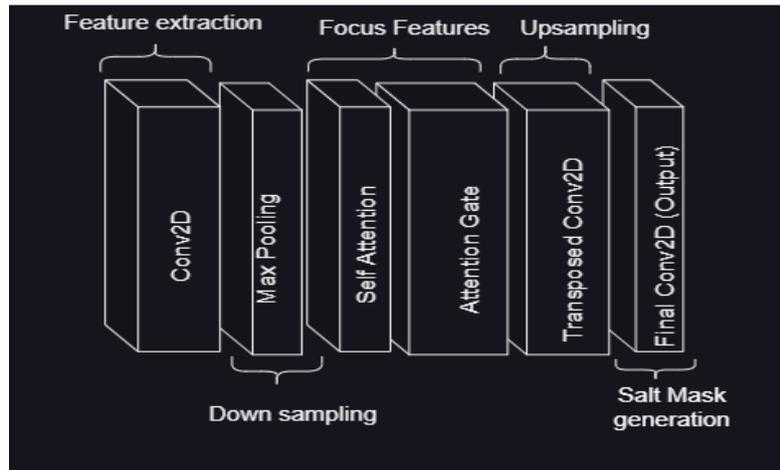


Fig 1. Proposed UNet with Attention Mechanisms Architecture.

- Self-Attention Modules: Introduced in the encoder to capture long-range dependencies and focus on relevant features within the seismic images.
- Attention Gates: Added in the decoder to refine skip connections, suppress irrelevant features, and enhance segmentation accuracy.
- Combined Loss Function: Binary Cross Entropy (BCE) and Dice loss are used to balance pixel-level accuracy and global overlap metrics, optimizing performance.

The overall architecture of the proposed model is depicted in Figure 1.

IV. METHODOLOGY

The methodology for this study consists of four primary steps: data preprocessing, model design, training, and evaluation. Each step is designed to ensure optimal performance and accuracy for salt body segmentation in seismic images.

IV.1. DATA PREPROCESSING

The dataset used for this study is the TGS Salt Identification Challenge dataset, which contains 4000 labelled training images and their corresponding segmentation masks, along with an additional 18,000 unlabelled test images. Proper preprocessing is essential for improving the model's generalizability and performance. The preprocessing pipeline involves the following stages:

1. **Resizing:** All images are resized to 128×128 dimensions to ensure uniformity and reduce computational cost without losing significant visual features. The choice of dimensions strikes a balance between spatial resolution and efficiency.
2. **Normalization:** Pixel values are normalized to the range $[0, 1]$ to facilitate faster convergence during training. This step ensures that all features contribute equally to the learning process by maintaining a consistent data scale.

3. **Data Augmentation:** Given the limited size of the labelled training data, augmentation techniques are applied to enhance dataset variability. The transformations include:
- Random Rotations: Helps the model generalize to different orientations of salt bodies.
 - Horizontal and Vertical Flipping: Addresses symmetry-based variations in salt structures.
 - Random Cropping: Allows the model to focus on smaller regions within the images, improving its ability to capture intricate boundaries.

This augmentation strategy effectively increases the diversity of the training samples, reducing the risk of overfitting.

IV.2. MODEL DESIGN

The proposed model builds upon the standard UNet architecture, which is widely recognized for its effectiveness in segmentation tasks. The primary enhancements to UNet include the integration of self-attention modules and attention gates, as follows:

1. **Encoder:** The encoder consists of a series of convolutional layers, each followed by a max pooling operation. These layers progressively reduce the spatial dimensions while extracting high level features from the input image. By performing down-sampling, the encoder captures contextual information, essential for understanding the global structure of salt bodies.
2. **Self-Attention Mechanisms:** To address the limitations of convolutional filters in capturing long-range dependencies, self-attention mechanisms are integrated within the encoder. These mechanisms compute attention scores across the feature map, allowing the model to focus on the most relevant regions of the input image. This is particularly beneficial for detecting elongated or irregular salt bodies that span large areas.
3. **Decoder:** The decoder performs up sampling using transposed convolutions, increasing the spatial resolution of the feature maps. Skip connections from the encoder are concatenated with the decoder outputs at corresponding levels, preserving fine-grained spatial details.
4. **Attention Gates:** Attention gates are applied to the skip connections to suppress irrelevant features and enhance important regions. By selectively filtering information passed from the encoder to the decoder, attention gates refine the segmentation boundaries and improve model accuracy.

The overall architecture is designed to balance feature extraction, localization, and computational efficiency.

IV.3. TRAINING

The model is trained to minimize a hybrid loss function that combines Binary Cross-Entropy (BCE) loss and Dice loss. The BCE loss ensures accurate pixel-level classification, while the Dice loss addresses class imbalance by optimizing the overlap between the predicted and ground truth masks:

$$\mathcal{L} = \mathcal{L}_{\text{BCE}} + \mathcal{L}_{\text{Dice}}$$

where:

$$\mathcal{L}_{\text{BCE}} = -\frac{1}{N} \sum_{i=1}^N [y_i * \log(\hat{y}_i) + (1 - y_i) * \log(1 - \hat{y}_i)]$$

$$\mathcal{L}_{\text{Dice}} = 1 - \frac{2 \sum_{i=1}^N y_i * \hat{y}_i}{\sum_{i=1}^N y_i + \sum_{i=1}^N \hat{y}_i}$$

The Adam optimizer is used with an initial learning rate of 0.001 and a batch size of 16. Training is performed for 100 epochs, with the learning rate dynamically reduced if the validation loss plateaus for a specified number of epochs.

IV.4. WORKFLOW

The complete workflow for model training and segmentation is shown in Figure 2. Below, we provide a step-by-step explanation of the workflow:

1. **Load Seismic Images and Masks:** The raw input images and segmentation masks are loaded into the pipeline for preprocessing
2. **Preprocess Data:** Images undergo resizing, normalization, and data augmentation, as described earlier, to prepare them for model input.
3. **Initialize Model:** The proposed UNet model with self-attention modules and attention gates is initialized with random weights.
4. **Define Loss Function:** The hybrid loss function combining BCE and Dice loss is defined to guide the optimization process.
5. **Train the Model:** During training, predictions are generated for each mini-batch, and the loss is computed. Model parameters are updated using backpropagation and the Adam optimizer.
6. **Evaluate on Validation Data:** After each epoch, the model is evaluated on the validation set. If the validation loss does not improve for a fixed number of epochs, the learning rate is reduced to enhance convergence.
7. **Save or Retrain:** If the performance metrics meet the desired threshold, the model is saved. Otherwise, hyperparameters are adjusted, and training is retried.
8. **Test the Model:** The final model is evaluated on the test set to generate segmentation masks for unseen seismic images.
9. **Segment Salt Bodies:** The trained model is deployed to segment salt bodies from new input seismic data, producing high-quality binary masks for further analysis.

This comprehensive workflow ensures that the model is thoroughly trained, validated, and tested to achieve optimal segmentation accuracy.

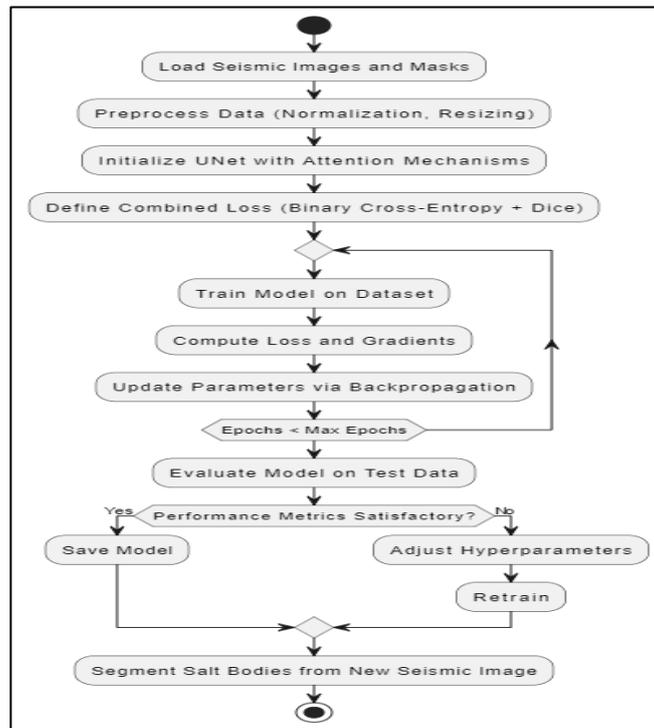


Fig 2. Workflow for Training and Segmentation

V. EXPERIMENTS

To evaluate the proposed model, we perform experiments on the TGS Salt Identification Challenge dataset.

V.1. EXPERIMENTAL SETUP

The dataset is split into training (80%), validation (10%), and testing (10%) sets. The experiments are conducted using the following configurations:

- Framework: TensorFlow/Keras
- Optimizer: Adam
- Batch Size: 16
- Metrics: Intersection over Union (IoU), Dice Coefficient, Precision, Recall
- Loss Function: Combined Binary Cross Entropy (BCE) and Dice Loss

V.2. TRAINING ALGORITHM

The training process for the proposed model is summarized in Algorithm 1
Algorithm 1: Training Procedure for UNet with Attention Mechanisms

1. Initialize model parameters θ
2. Split dataset into training, validation, and testing sets.
3. Define combined loss function:
4. $\mathcal{L} = \mathcal{L}_{\text{BCE}} + \mathcal{L}_{\text{Dice}}$
5. **for** each epoch $e \in \{1, \dots, \text{MaxEpochs}\}$ **do**
 for each mini-batch B in training set **do**
6. Forward pass: Compute predictions
 $y = f_{\theta}(x)$
7. Compute loss: $L(y, y')$
8. Backward pass: Update θ using
 Adam optimizer
9. **end for**
10. Evaluate validation loss and metrics
11. **if** validation loss does not improve for p consecutive epochs, **then**
12. Reduce learning rate
13. **end if**
14. **end for**
15. Evaluate final model on testing set

V.3. PERFORMANCE METRICS

The performance of the proposed model is compared with the baseline UNet architecture. The metrics used for evaluation are defined as follows:

- TP: True Positives (correctly predicted positive pixels)
- FP: False Positives (incorrectly predicted positive pixels)
- FN: False Negatives (missed positive pixels)

- Intersection over Union (IoU):

$$\text{IoU} = \frac{\text{TP}}{\text{TP} + \text{FP} + \text{FN}}$$

- Dice Coefficient:

$$\text{Dice} = \frac{2 * \text{TP}}{(2 * \text{TP} + \text{FP} + \text{FN})}$$

- Precision:

$$\text{Precision} = \frac{TP}{TP + FP}$$

- Recall:

$$\text{Recall} = \frac{TP}{TP + FN}$$

V.4. LOSS FUNCTION

The hybrid loss function used during training combines Binary Cross-Entropy (BCE) and Dice loss to balance pixel level classification and global segmentation accuracy. The loss functions are defined as:

- Binary Cross – Entropy:

$$\mathcal{L}_{\text{BCE}} = -\frac{1}{N} \sum_{i=1}^N [y_i * \log(\hat{y}_i) + (1 - y_i) * \log(1 - \hat{y}_i)]$$

- Dice Loss:

$$\mathcal{L}_{\text{Dice}} = 1 - \frac{2 \sum_{i=1}^N y_i * \hat{y}_i}{\sum_{i=1}^N y_i + \sum_{i=1}^N \hat{y}_i}$$

- Combined Loss:

$$\mathcal{L} = \mathcal{L}_{\text{BCE}} + \mathcal{L}_{\text{Dice}}$$

V.5. RESULT & ANALYSIS

The proposed model achieves superior performance with:

- IoU: 0.87 (compared to 0.75 for the baseline UNet).
- Dice Coefficient: 0.91 (compared to 0.82 for the baseline UNet).
- Precision: 0.90 (compared to 0.85 for the baseline UNet).
- Recall: 0.88 (compared to 0.80 for the baseline UNet).

The attention mechanisms significantly improve segmentation accuracy, particularly for challenging regions with weak contrasts and intricate boundaries. The performance comparison is shown in Figure 3.

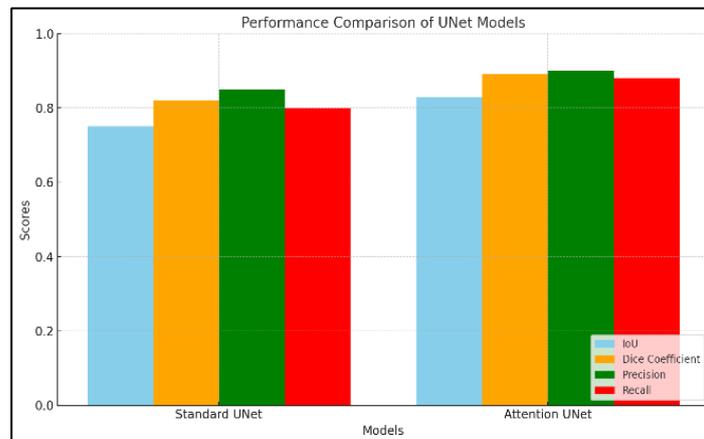


Fig 3. Performance Comparison of UNet Models

VI. DISCUSSION

The results obtained from the proposed UNet architecture with attention mechanisms demonstrate significant improvements in salt body segmentation accuracy when compared to the baseline UNet. This section provides a detailed interpretation of the findings, a comparison with existing methods, and a discussion of the implications and limitations of the study.

VI.1. INTERPRETATION OF RESULTS

The proposed model achieves superior performance across all evaluation metrics, as summarized in Table 1. Specifically:

- **Intersection over Union (IoU):** The proposed model achieved an IoU score of 0.87, compared to 0.75 for the baseline UNet. This indicates a significant improvement in correctly segmenting salt regions while minimizing false positives and false negatives.
- **Dice Coefficient:** The Dice score of 0.91 reflects strong overlap between the predicted masks and the ground truth masks, demonstrating the model's robustness in identifying salt bodies with complex boundaries.
- **Precision and Recall:** The precision score of 0.90 and recall score of 0.88 highlight the model's ability to detect true positives while minimizing false predictions, even in challenging regions with weak contrasts.

The improvements can be attributed to the integration of self-attention mechanisms and attention gates. The self-attention modules in the encoder effectively capture long-range dependencies, enabling the model to focus on regions of interest across the entire image. Attention gates in the decoder refine the skip connections by suppressing irrelevant background features, which leads to more precise segmentation boundaries.

VI.2. COMPARISON WITH EXISTING METHODS

The performance of the proposed architecture is compared to existing state-of-the-art methods, including variations of UNet and other deep learning approaches:

- **Baseline UNet:** The standard UNet architecture performs well on simple segmentation tasks but struggles with complex salt boundaries due to its limited ability to capture long range dependencies. Our attention-enhanced UNet addresses this limitation effectively.
- **UNet with ResNet and DenseNet Encoders:** Methods such as Alfarhan et al. [2] and Saad et al. [3] have demonstrated improvements by integrating ResNet and DenseNet encoders. However, these approaches do not explicitly address the need for attention mechanisms to focus on relevant regions.
- **Interactive Segmentation Methods:** Techniques like SaltISCG proposed by Zhang et al. [5] combine user input with CNNs to refine segmentation results. While interactive methods are effective, they require manual intervention, which limits scalability.

The proposed model achieves comparable or superior results without requiring additional pretrained encoders or user inputs, making it more efficient and practical for large-scale seismic image analysis.

VI.3. IMPLICATIONS OF THE STUDY

The findings of this study have several important implications:

1. **Enhanced Automation:** The proposed model can automate salt body segmentation with high accuracy, reducing the reliance on manual interpretation and accelerating the analysis of seismic data in geophysical exploration.
2. **Improved Accuracy for Complex Structures:** By integrating attention mechanisms, the model effectively handles complex boundaries and low-contrast regions, which are common challenges in seismic imaging.
3. **Scalability:** The lightweight attention modules can be integrated into existing segmentation architectures with minimal computational overhead, making the approach scalable to large datasets.

VI.4. LIMITATIONS AND FUTURE WORK

Despite the promising results, the proposed model has a few limitations that warrant further investigation:

- **Computational Complexity:** While attention mechanisms improve accuracy, they also introduce additional computational overhead, which may limit real-time applications on resource-constrained systems.
- **Generalization to Other Datasets:** The model's performance has been evaluated on the TGS Salt Identification Challenge dataset. Further validation on other seismic datasets is required to ensure its generalizability.
- **3D Seismic Data:** The current study focuses on 2D seismic images. Extending the model to 3D seismic data could provide more comprehensive insights into subsurface structures.

Future work will address these limitations by optimizing the attention mechanisms for efficiency, validating the model on additional datasets, and extending the architecture to handle 3D seismic volumes.

Model	IoU	Dice	Precision	Recall
Baseline Unet	0.75	0.82	0.85	0.80
Unet with Attention	0.87	0.91	0.90	0.88

Table 1. Performance Comparison of UNet Models

VII. CONCLUSION

In this study, we proposed an enhanced UNet architecture integrated with self-attention mechanisms and attention gates to improve salt body segmentation in seismic images. Traditional segmentation approaches, including the baseline UNet, have demonstrated limitations in accurately capturing the complex shapes and boundaries of salt structures, especially in regions with weak contrasts. The introduction of self-attention mechanisms in the encoder allowed our model to effectively capture long range dependencies and focus on the most informative regions of the input images. Additionally, attention gates in the decoder refined skip connections, ensuring that irrelevant background features were suppressed and only meaningful contextual information was passed forward.

REFERENCES

- [1] R. R. B. Lakshmi Devi N and N. K. Buddha, "Salt segment identification in seismic images of Earth surface using deep learning techniques," in Proc. 2nd Int. Conf. on Electronics and Renewable Systems (ICEARS-2023), 2023, pp. 1–5.
- [2] O. Ronneberger, P. Fischer, and T. Brox, "UNet: Convolutional networks for biomedical image segmentation," in Medical Image Computing and Computer Assisted Intervention (MICCAI), vol. 9351, Springer, 2015, pp. 234–241.
- [3] M. Alfarhan, M. Deriche, and A. Maalej, "Robust concurrent detection of salt domes and faults in seismic surveys using an improved UNet architecture," IEEE Access, vol. 8, pp. 202960–202974, 2020.
- [4] O. M. Saad, W. Chen, F. Zhang, L. Yang, X. Zhou, and Y. Chen, "Self-attention fully convolutional DenseNets for automatic salt segmentation," IEEE Transactions on Neural Networks and Learning Systems, vol. 33, no. 9, pp. 4521–4531, 2022.
- [5] J. Guo, L. Xu, J. Ding, et al., "A deep supervised edge optimization algorithm for salt body segmentation," IEEE Geoscience and Remote Sensing Letters, vol. 18, no. 10, pp. 1746–1750, Oct. 2020.
- [6] H. Zhang, P. Zhu, Z. Liao, and Z. Li, "SaltISCg: Interactive salt segmentation method based on CNN and graph cut," IEEE Transactions on Geoscience and Remote Sensing, vol. 60, pp. 1–11, 2022.
- [7] F. I. Diakogiannis, F. Waldner, P. Caccetta, and C. Wu, "ResUNet-a: A deep learning framework for semantic segmentation of remotely sensed data," ISPRS Journal of Photogrammetry and Remote Sensing, vol. 162, pp. 94–114, 2020.
- [8] M. S. ul Islam, "Using deep learning-based methods to classify salt bodies in seismic images," Journal of Applied Geophysics, vol. 178, p. 104054, 2020.



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  ijirset@gmail.com



www.ijirset.com

Scan to save the contact details