Salt Body Segmentation in Seismic Images using Mask R-CNN

Arsha P V
Department of Computer Science and Engineering
NSS College of Engineering
Palakkad, India
arshaabdulla92@gmail.com

Pillai Praveen Thulasidharan
Department of Computer Science and Engineering
Government Engineering College Idukki
Painavu, India
praveentpillai@gmail.com

Abstract—Salt body exploration has always been a major area of study in Geophysics and the associated area of seismic interpretation, owing to its commercial significance in the oil and petroleum industry. In recent years, with the advancement in the area of machine learning; deep learning models are used for the efficient detection and segmentation of salt body from seismic images. In this paper, a deep neural model with instance segmentation capability (Mask R-CNN) has been proposed for the detection and segmentation of salt body from a set of seismic images. The model accepts a seismic image as input, performs the required classification of the pixels and outputs an image with annotated marks, clearly showing the segmented salt regions. The model has been trained and tested using the Seismic image dataset provided by Kaggle. The results obtained have been compared with the already known standard CNN model in this area and the experimental results shows that the proposed model is better.

Index Terms—Salt Exploration, Mask R-CNN, Seismic Image, Geophysics, Image Analysis.

Fig. 1. Seismic Imaging Environmental Set-up.

I. INTRODUCTION

Salt body is one of the important geological structures that aids in the sealing of oil and natural gas reservoirs beneath the earth's surface and hence accurate salt body detection would greatly facilitate in the exploration of petroleum and natural gas. Seismic images, generated by an imaging technique known as Seismic Imaging, are the commonly used artifacts for the detection of salt body. Seismic imaging comprises of techniques for pictographically modelling the subsurface of earth for geological and other engineering studies. The majority of these techniques use the reflection and refraction characteristics of the acoustic signals to generate the reconstructed model of the earth's crust. A detailed seismic imaging environmental setup is illustrated in Fig. 1. The acquisition process starts by propagating seismic waves of known characteristics, generated by an acoustic source, through the earth's interior and is received by geophones placed at specific location on the surface of earth. The received signal is analyzed to understand the hidden structure of the earth's crust and the data are calculated to construct an image representation of earth's subsurface. Traditionally, salt body interpretation from seismic images has largely been carried out by experienced human experts. However with the increase in the amount of data to be analyzed and in order to reduce processing time;
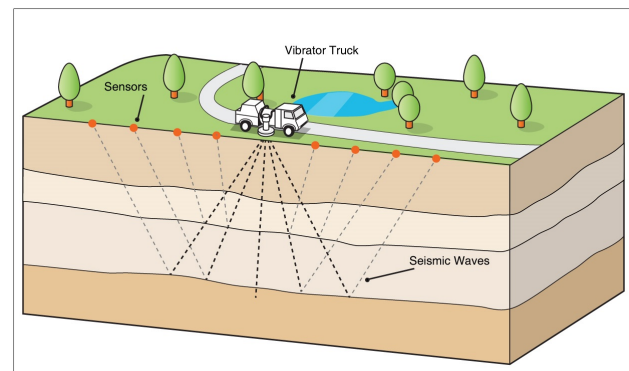
in recent years we have witnessed an increasing interest in developing computer-based systems that are intended to locate salt deposits within the seismic images.

Most of the salt body detection algorithms presented in the literature [5] [6] [7], use edge detection methods, Normalized Cuts Image Segmentation (NCIS) or texture attributes based methods. From the perspective of seismic attribute analysis; graph cut, edge detection and texture analysis are applicable for highlighting the boundaries of a salt deposits. These methods take into account the significant variation of seismic signals across the boundaries. The edge detection based techniques are computationally efficient and helpful in detecting the salt boundaries and horizon in seismic data. However, these techniques yield good results only when seismic data exhibit strong amplitude variations. The NCIS based algorithms used graph based models to segment the salt domes. These methods are computationally very expensive and therefore not suitable for real time seismic interpretation. The texture attributes based salt body detection methods were introduced to overcome the drawbacks of edge based detection methods that are heavily dependent upon the instantaneous amplitude only. However, in texture based methods, the choice of attributes and the size of window play a crucial role in the accurate detection of salt domes. Considering the geologic complexities; in most cases, the use of fewer salt attributes was insufficient for

accurate delineation of the boundaries of a salt body. For addressing such limitation, considering the insufficiency of attributes for reliable salt detection, researchers have suggested integrating multiple attributes through machine learning based techniques for improved detection accuracy and efficiency.

In this paper, Mask Region based Convolutional Neural Network (Mask R-CNN) [1] model is used for accurately delineating the area of salt bodies from seismic images. Mask R-CNN a deep neural network, is currently the new state of the art in terms of instance segmentation aimed to solve the object segmentation problem in machine learning or computer vision. In the proposed system, the seismic Image is passed onto the deep learning system, where a trained neural network, which is a Mask R-CNN, will process the image and performs classification to see whether there are any salt body elements existing in the frame and if present, generates masks around the salt body region. Our experiment results show that the proposed system outperforms the existing machine learning model available in this area.

II. PROPOSED SYSTEM

In this paper, we present a system for the detection and segmentation of salt body from Seismic Images using Mask R-CNN (Mask Region Based Convolutional Neural Network). The Mask R-CNN, is a segmentation model that allows us to identify the location of each of the class category within the given image at each pixel level. In the proposed system, the Mask R-CNN is able to classify each salt body pixels from the surrounding non-salt sediment pixels.

A. Architecture

The architecture of the proposed system is illustrated in Fig. 2. The proposed model intake seismic image as an input and provides the output image with the salt mask. First, the Mask R-CNN examines the input seismic images
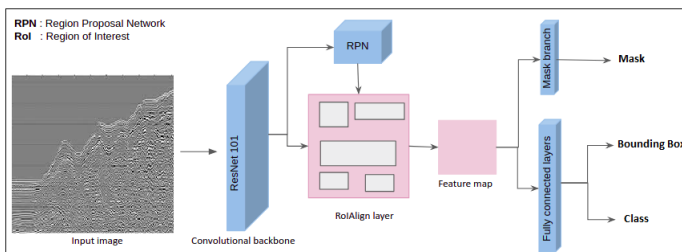

Fig. 2. Proposed System Architecture.

and generates proposals (area likely to contain salt), then it classifies the proposals and produces bounding box and mask around the salt body. Mask R-CNN consists of these components:

- **Backbone model**: Mask R-CNN is built on a backbone convolutional neural network architecture for feature extraction. In the proposed system, Resnet101 is used as the backbone network model, which is a standard

convolutional neural network that serves as a feature extractor. The early layers detect low-level features (salt body edges and corners) and later layers successively detect higher level features (structures). The output of this backbone will be a feature map that serves as an input for the next network layers.

- **Region Proposal Network (RPN)**: The RPN generates candidate regions known as anchors which may contain the salt area. These anchors are selected by traversing the image in a sliding window manner and based on the RPN predictions, top anchors that are prone to contain the salt area are further refined. After that, we have final proposals (region of interest) that we pass to the next stage.

- **Region of Interest Classification and Bounding Box**: In this step the algorithm takes the regions of interest (ROI) proposed by the RPN as an input and outputs a classification (salt, non-salt) and a bounding box to encapsulate the salt body area.

- **Segmentation Masks**: In the final step, the mask branch (convolutional network) that takes the positive regions and generates masks for salt area. Fig. 3 illustrates the input Seismic Image and final output image with salt mask. The model converts the input image into the corresponding output image marked with the salt mask.
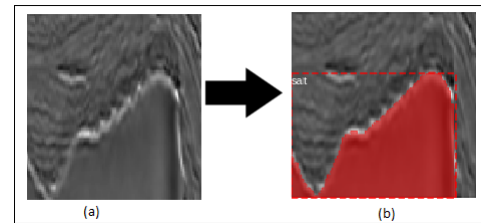

Fig. 3. (a) Input Image (b) Final Output Image.

III. IMPLEMENTATION DETAILS

A. Dataset

The dataset used by the system is obtained from the TGS salt identification challenge competition hosted on the Kaggle data science community [2]. The dataset consist of 4000 seismic images with the size of 101x101 for the various training and testing purpose. For each image, the corresponding binary image with the salt mask has been provided.

B. Implementation

The Implementation of the proposed system is done in Python. The experiment setup runs on Google Colaboratory, a cloud-based data science workspace using Keras Tensorflow backend and other python machine learning libraries. We have used the kaggle dataset to implement the proposed salt body detection model. The implementation is mainly based

on an existing Mask R-CNN model developed by Matterport repository [3], released under an MIT License. The network model is trained with a total of 80 epochs and 100 steps per each epoch with the batch size of 8. The detailed configuration information of the trained model is shown in Table 1.

IV. TRAINING

The Training was carried out using Kaggle Dataset, which provides seismic images with the salt mask. A total of 4000 seismic images of both salt body element and non-salt body element cases are used, which are splitted into two groups of 3500 for training and 500 for testing. The model is trained with Resnet101 backbone and with a batch size of 8 images. Each mini-batch has one GPU, each image has 16 sampled RoIs and 256 anchors for FPN. We use a top-down pyramid size of 128, a learning rate of 0.001 and a weight decay of 0.0001. The network model is trained for a total of 80 epochs and 100 steps per epochs. The training performance is evaluated by using Accuracy and Loss measures.

- **Accuracy:** Accuracy is the most common metric used to assess the performance of a model. It gives a measure of the correct predictions from the total available samples. Mathematically it is defined as the ratio of the predictions which are predicted correctly to the total number of samples given as input.

- **Loss:** Loss is defined as an error over the training dataset. Mask R-CNN uses a multi-task loss function. The total loss calculated as: $L = L_{cls} + L_{bbox} + L_{mask}$.
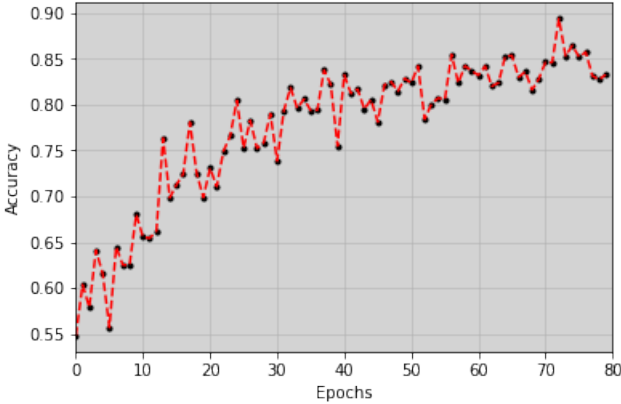

Fig. 4. Accuracy curve.

Fig. 4 illustrates the accuracy curve of the training process for 80 epochs. Similarly, Fig. 5 illustrates the loss curve of the training process for 80 epochs. Initially, the model starts from random noises, so the initial predictions are of wrong regions, thus the accuracy curve begins with low accuracy value and loss curve begins with high loss value. As the training continues, after about 70 epochs the model becomes capable of depicting the region of the valid seismic salt body

area. Finally, the model achieves an accuracy of 0.89 and a loss of 0.7. When the performance is not expected to further improve by more training epochs, the training is stopped at 80 epoch.
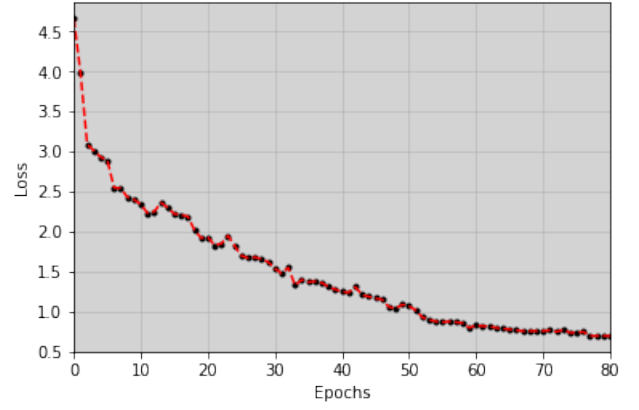

Fig. 5. Loss Curve.

V. RESULT AND ANALYSIS

A. Testing

At the end of the training, we have 80 distinct accurate models M-1, M-2 ... M-80; i.e. the first model derived from the first training epoch named as M-1, the second model derived from the second training epoch named as M-2 and so on. All the models are tested with 400 seismic images for examining its salt detection capability. For demonstrating the efficiency of these models, we compute different evaluation metrics. Four performance evaluation metrics such as Precision, Recall, F1 Score and Accuracy are used for determining the performance of the system. If Ω is the detected salt body region and Ω' is the salt body region in the ground truth image, then the evaluation metric is calculated using the below respective equations.

$$\mathbf{Precision} = \frac{|\Omega \cap \Omega'|}{|\Omega|} \quad (1)$$

$$\mathbf{Recall} = \frac{|\Omega \cap \Omega'|}{|\Omega'|} \quad (2)$$

$$\mathbf{F1\ score} = 2 * \frac{precision * recall}{precision + recall} \quad (3)$$

$$\mathbf{Accuracy} = \frac{|\Omega|}{|\Omega \cup \Omega'|} \quad (4)$$

While considering all 80 models, 72nd epoch model possess efficient Precision, Recall, F1-score and Accuracy. Fig. 6 illustrates the evaluation measures of models from Epoch 71,72,73,74 named as M-71, M72, M-73 and M-74. It is found that the evaluation metrics values are maximum at M-72. Therefore we choose the 72nd epoch model weights as our proposed model weights and hence further testing and discussion are carried out with this M-72 model.

TABLE I
MODEL CONFIGURATIONS

Configuration Parameter	Value	Configuration Parameter	Value	Configuration Parameter	Value
Backbone	resnet101	Backbone strides	[4, 8, 16, 32, 64]	Validatio steps	50
Batch size	8	BBox std dev	[0.1 0.1 0.2 0.2]	weight decay	0.001
Compute backbone shape	None	Detection max instance	3	Use mini mask	True
Detection in confidence	0.9	Detection NMS threshold	0.1	Use RPN rois	True
FPN classif FC-layer size	1024	GPU count	1	Train BN	False
Gradient clip norm	5.0	Images per GPU	8	Train rois per image	16
Image channel count	3	Image max dim	128	Steps per epochs	100
Image min dim	128	Image meta size	14	Topdown pyramid size	128
Image min scale	0	Image resize mode	square	RPN NMS threshold	0.7
Image shape	[128 128 3]	Learning momentum	0.9	RPN train anchors per image	256
Detection in confidence	0.9	Detection NMS threshold	0.1	RPN anchor strides	1
FPN classif FC-layer size	1024	GPU count	1	RPN bbox std dev	[0.1 0.1 0.2 0.2]
Gradient clip norm	5.0	Images per GPU	8	RPN anchor ratio	[0.5, 1, 2]
Image channel count	3	Image max dim	128	RPN anchor scales	(32, 64, 128, 256, 512)
Image min dim	128	Image meta size	14	Post NMS rois training	2000
Image min scale	0	Image resize mode	square	Roi positive ratio	0.33
Image shape	[128 128 3]	Learning momentum	0.9	Post NMS rois inference	1000
Learning rate	0.001	Loss weight	'rpn$_c$lass$_l$oss' : 1.0, 'rpn$_b$box$_l$oss' : 1.0, 'mrcnn$_c$lass$_l$oss' : 1.0, 'mrcnn$_b$box$_l$oss' : 1.0, 'mrcnn$_m$ask$_l$oss' : 1.0	Pre NMS limit	6000
• • • •		• • • •			
Mask pool size	14	Mask shape	[28,28]	Num Classes	2
Max GT instance	1	Mean pixel	[123.7 116.8 103.9]	Pool size	7
Mini mask shape	(56,56)	Name	find -salt		
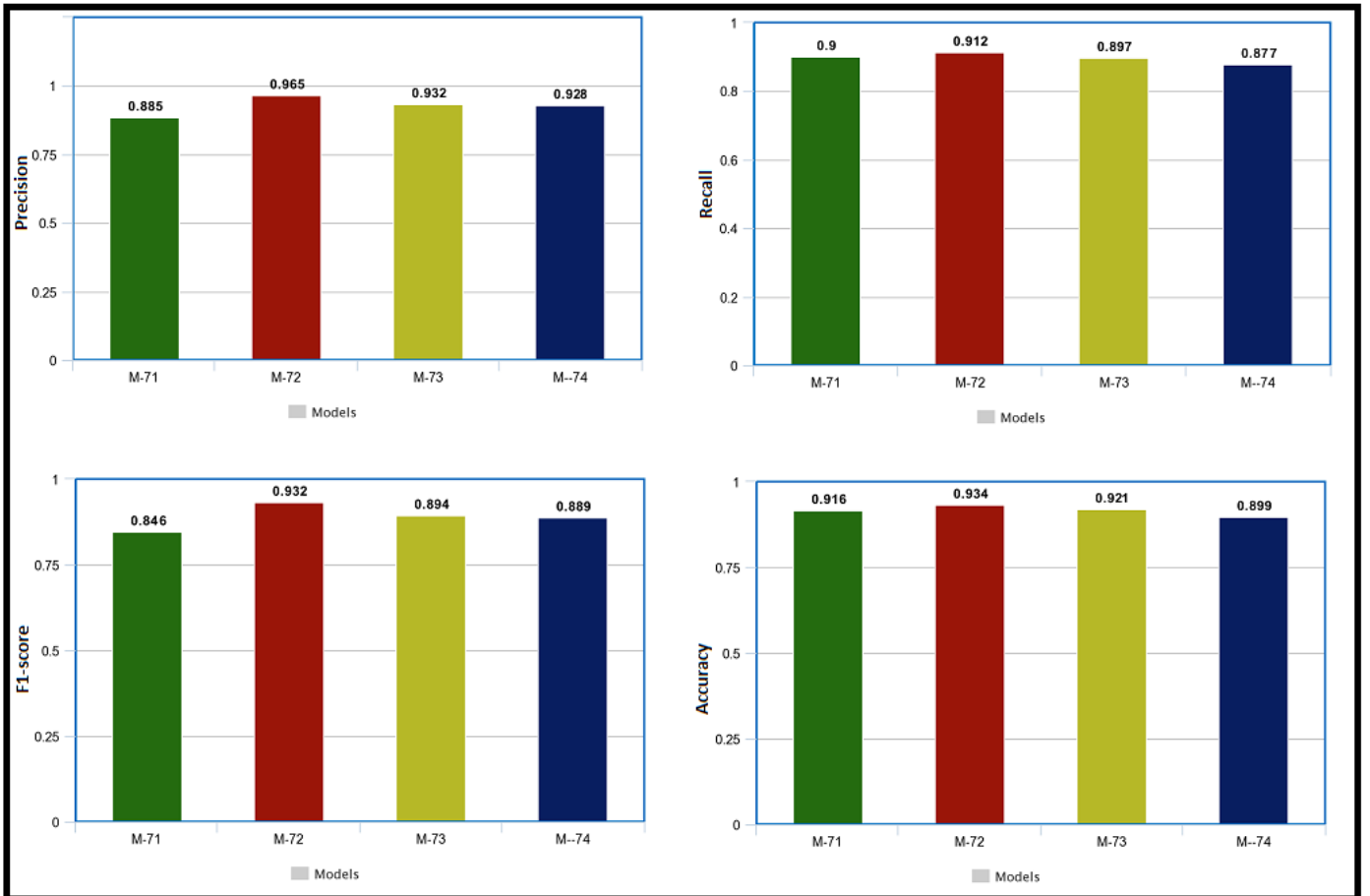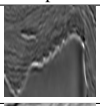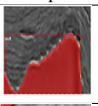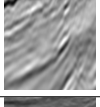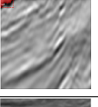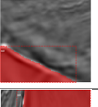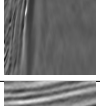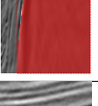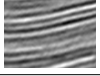
Fig. 6. Performance Analysis Based on Training Epochs.

	Input	Output	Precision	Recall	F1-score	Accuracy
1			0.965	0.906	0.927	0.935
2			0.999	0.967	0.982	0.966
3			0.998	0.887	0.987	0.934
4			0.984	0.950	0.967	0.971
5			1.00	1.00	1.00	1.00
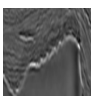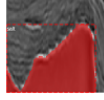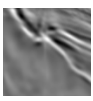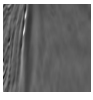
B. Results and Discussions

The proposed system is tested with a variety of seismic images to ensure the detection capability of the proposed model. Table 2 shows the experimental results of the proposed salt body detection model for some of the test images. The input image an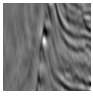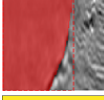d corresponding output image with the salt mask are illustrated by the first two entries in Table 2. The third, fourth, fifth and sixth entries illustrate the Precision, Recall, F1-score and Accuracy scores of the individual images respectively. The first test image is a seismic image with salt body areas in the lower middle portion. The model efficiently detects this salt body area and generates the output with the salt mask around the identified salt area. Even though the second test image contains only a small amount of salt body content in the left top corner, the model also discovers this salt content effectively. The model identifies and generates output images with valid salt regions for the third and fourth test images. The model effectively works well in the case of non-salt image content illustrated by the last test case in Table 2. The evaluation metrics scores of Precision, Recall, F1-score, Accuracy for each test image shows that the proposed model satisfies the goal of the salt body detection system.

The performance evaluation of the proposed model is evaluated by comparing it with the CNN [4] model, that too uses deep learning aspects for its functioning. The discussed CNN model is the simplest and one of the latest deep learning model used in the area of seismic interpretation. For evaluating, the proposed Mask R-CNN model and the CNN model are tested with a variety of seismic images obtained from the Kaggle dataset [2] and the results are analyzed on the basis of the evaluation metrics discussed in Section 6.1. The outputs received for a few of the tested images are illustrated in Table. 3; in which, for the Mask R-CNN model the segmented area is shown in 'Red', while for the CNN model, the segmented area is shown in 'Yellow'. The values for the respective evaluation

metrics are also shown along with the models. It is quite evident from the results that the proposed Mask R-CNN model has better evaluation metrics values as compared to the CNN model in all the aspects. To confirm this finding, both the models are tested on a set of randomly selected 400 images and the evaluation metrics (Recall, Precision, F1Score and Accuracy) obtained for each image is plotted and shown in Fig.7. The Fig. 7(a), illustrates the Recall values of the two models; similarly Fig. 7(b), Fig. 7(c) and Fig. 7(d), depicts the Precision, F1Score and Accuracy values respectively. The values for Mask R-CNN are denoted by 'Red' dots and that for the CNN model, they are shown as 'Green' dots. In each plot it can be seen that, majority of the red dots occupy a position at the upper portion of the plot denoting a higher value, while the green dots occupy a comparatively lower position denoting a smaller value. This confirms that the Mask R-CNN model generates better values as compared to the CNN model for all the four evaluation metrics. The average values obtained for the respective evaluation metrics during the testing of the 400 images is shown in Table. 4. From these experimental results, it can be concluded that the proposed Mask R-CNN model certainly performs better than the CNN model.

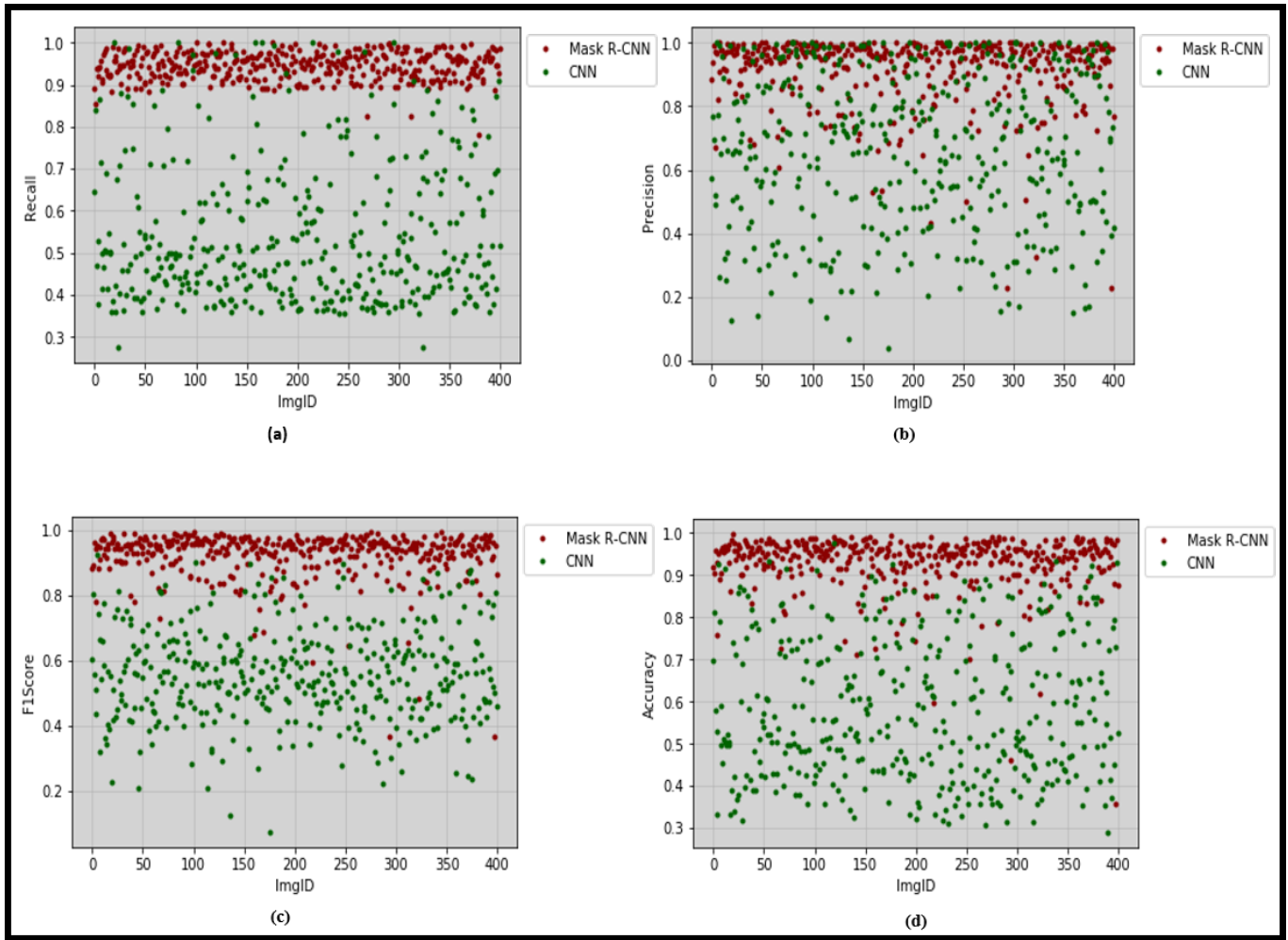| Input		Output	Precision	Recall	F1-score	Accuracy
	Mask R-CNN		0.965	0.906	0.927	0.935
	CNN		0.234	0.328	0.268	0.332
	Mask R-CNN		0.950	0.959	0.955	0.971
	CNN		0.767	0.763	0.765	0.816
	Mask R-CNN		0.965	0.906	0.927	0.935
	CNN		0.677	0.554	0.609	0.793
	Mask R-CNN		0.962	0.896	0.928	0.915
	CNN		0.342	0.404	0.370	0.571

Fig. 7. Mask R-CNN and CNN Evaluation Metrics.

TABLE IV
MASK R-CNN AND CNN EVALUATION METRICS AVERAGE

	Dataset	Recall	Precision	F1-score	Accuracy
Mask R-CNN	Kaggle Dataset	0.95	0.92	0.93	0.93
CNN	Kaggle Dataset	0.54	0.67	0.56	0.58

VI. CONCLUSION

An automatic salt body detection and segmentation system using Mask Region Based Convolutional Neural Network has been proposed in this paper. The proposed system has been tested extensively using Kaggle Dataset and the experimental results obtained has been demonstrated. It is observed, that our proposed model outperforms existing salt body detection methods by offering efficient performance metrics.

REFERENCES

[1] K. He, G. Gkioxari, P. Doll ar, and R. Girshick,"'Mask r-cnn," Proceedings of the IEEE international con-ference on computer vision, pp. 2961–2969, 2017.
[2] TGS, 2018. Tgs salt identification challenege. https://www.kaggle.com/c/tgs-salt-identification-challenege/data.
[3] Abdulla W., 2017. Mask r-cnn for object detection and instance segmentation on keras and tensorflow. https://github.com/matterport/Mask_RCNN.
[4] Waldeland A.U., Jensen A.C., Gelius L.J. and Solberg A.H.S., "Convolutional neural networks for automated seismic interpretation." The Leading Edge, vol. 37, issue. 7, pp. 529–537, July 2018.
[5] J. Lomask, B. Biondi, and J. Shragge, "Im-age segmentation for tracking salt boundaries," inSEG Technical Program Expanded Abstracts 2004,pp. 2443–2446, Society of Exploration Geophysicists,2004.
[6] J. Shi and J. Malik, "Normalized cuts and imagesegmentation," Departmental Papers (CIS), pp. 107,2000.
[7] A. D. Halpert, R. G. Clapp, and B. Biondi, "Seis-mic image segmentation with multiple attributes," inSEG Technical Program Expanded Abstracts 2009,pp. 3700–3704, Society of Exploration Geophysicists,2009.